PCCP

Do Molecular Dynamics Force Fields Accurately Model Ramachandran Distributions of Amino Acid Residues in Water?

Journal:	*Physical Chemistry Chemical Physics*
Manuscript ID	CP-ART-11-2021-005069.R1
Article Type:	Paper
Date Submitted by the Author:	21-Dec-2021
Complete List of Authors:	Andrews, Brian; Drexel University, Physics Guerra, Jose; University of Chicago Division of the Physical Sciences, Physics Schweitzer-Stenner, Reinhard; Drexel University, Chemistry Urbanc, Brigita; Drexel University, Physics

PCCP

ARTICLE TYPE

Cite this: DOI: 00.0000/xxxxxxxxxx

Do Molecular Dynamics Force Fields Accurately Model Ramachandran Distributions of Amino Acid Residues in Water?†

Brian Andrews,[a] Jose Guerra,[c], Reinhard Schweitzer-Stenner,[b] and Brigita Urbanc*[a]

Received Date
Accepted Date

DOI: 00.0000/xxxxxxxxxx

Molecular dynamics (MD) is a powerful tool for studying intrinsically disordered proteins, however, its reliability depends on the accuracy of the force field. We assess Amber ff19SB, Amber ff14SB, OPLS-AA/M, and CHARMM36m with respect to their capacity to capture intrinsic conformational dynamics of 14 guest residues x (= G, A, L, V, I, F, Y, D^P, E^P, R, C, N, S, T) in GxG peptides in water. The MD-derived Ramachandran distribution of each guest residue is used to calculate 5 J-coupling constants and amide I' band profiles to facilitate a comparison to spectroscopic data through reduced χ^2 functions. We show that the Gaussian model, optimized to best fit the experimental data, outperforms all MD force fields by an order of magnitude. The weaknesses of the MD force fields are: (i) insufficient variability of the polyproline II (pPII) population among the guest residues; (ii) oversampling of antiparallel at the expense of transitional β-strand region; (iii) inadequate sampling of turn-forming conformations for ionizable and polar residues; and (iv) insufficient guest residue-specificity of the Ramachandran distributions. Whereas Amber ff19SB performs worse than the other three force fields with respect to χ^2 values, it accounts for residue-specific pPII content better than the other three force fields. Additional testing of residue-specific RSFF1 and Amber ff14SB combined with TIP4P/2005 on six guest residues x (= A, I, F, D^P, R, S) reveals that residue specificity derived from protein coil libraries or an improved water model alone do not result in significantly lower χ^2 values.

Introduction

A classical model of a protein assumes that its amino acid sequence determines the unique native fold, which then dictates the function this protein will perform. The discovery and subsequent studies of intrinsically disordered proteins (IDPs), which by definition contain unfolded, disordered regions and do not adopt a unique native fold, have revealed that this classical view is overly simplistic. IDPs are known for their structural plasticity, which allows them to bind multiple partners and perform a myriad of functions[1]. Under stress conditions, some IDPs, such as amyloid β-protein, tau, and α-synuclein, can form toxic assemblies, which play central roles in Alzheimer's and Parkinson's diseases. A thorough understanding of structural plasticity of IDPs, which is encoded in their unfolded regions, and the capacity to undergo disorder-to-order transitions upon binding to other proteins, DNA, RNA, or cellular membranes, is missing. Due to their disordered nature and structural plasticity, IDPs resist crystallization and even information gathered by modern experimental methods, such single-molecule cryo-electron microscopy, is limited due to this inherent intrinsic disorder[2].

Molecular dynamics (MD) offers a plethora of powerful computational techniques that are ideally suited to provide atomistic-level details on conformational dynamics of IDPs in water. However, the validity of MD predictions depends on the accuracy of the underlying force field, i.e. interactions within and among protein molecules, water-water, and water-protein interactions. MD force fields are based on classical physics-based energy functions with many adjustable parameters, which are most often calibrated to experimental data and/or quantum mechanical calculations. When applied to IDPs, MD force fields tend to produce overly collapsed conformations[3,4]. This issue can be to some extent addressed by strengthening the protein-water interactions while keeping the water-water and protein-protein interactions intact[4,5] or by improving the water model so as to increase pro-

[a] Department of Physics, Drexel University, Philadelphia, Pennsylvania, USA. Tel: XX XXXX XXXX; E-mail: bu25@drexel.edu
[b] Department of Chemistry, Drexel University, Philadelphia, Pennsylvania, USA.
[c] Department of Chemistry, Stony Brook University, Stony Brook, New York, USA.

† Electronic Supplementary Information (ESI) available: Table S1: Experimental and calculated J-coupling constants. Table S2: Populations of mesostates defined in *Methods*. Table S3: Shannon entropy differences ΔS_I and ΔS_{II}. Figures S1-3: Time evolution of MD-derived J-coupling constants for isoleucine, arginine, and tyrosine. Figure S4: Comparison of reduced χ_J^2 values calculated using two different sets of Karplus parameters. Figures S5-32: Absolute differences between experimental and calculated J-coupling constants, reduced χ_J^2 values, reduced $\chi_{\vec{V}CD}^2$, and amide I' band profiles of guest residues x in GxG peptides. See DOI: 10.1039/cXCP00000x/

tein solubility[3,6,7]. It is unclear whether increased protein solubility alone, however, is sufficient to resolve the problem of force fields overly promoting protein self-assembly[8–10].

A reliable MD force field for studies of IDPs should capture intrinsic conformational preferences of amino acid residues in water, as reflected in their respective Ramachandran distributions of backbone dihedral angles. This issue has not drawn sufficient attention because amino acid residues in water have been assumed to sample all sterically allowed regions of the Ramachandran space. Moreover, conformational preferences have been expected to be similar for all residues other than glycine and proline. Most MD force fields implicitly support this conjecture by modeling the force field backbone dihedral potential energy using alanine-based peptides (as in CHARMM36[11], OPLS-AA/M[12], and Amber ff14SB[13]) and/or utilize non-glycine and non-proline residues within a set of IDPs (as in CHARMM36m[6,14]). The majority of MD force field-development efforts focused on reproducing the experimental J-coupling constants of alanine-based peptides[15–17]. Dihedral potential parameters are often derived from quantum mechanical (QM) energy calculations using short unfolded polypeptides *in vacuo* or in implicit water, whereby the selection of polypeptides and functional forms of dihedral energy functions are force field-dependent[6,12,13]. Best and collaborators introduced a common non-glycine and non-proline backbone CMAP potential refined against solution NMR data for weakly structured peptides, resulting in CHARMM36[11], which was carried over into the recent version of CHARMM, CHARMM36m[6]. Similar backbone CMAP potentials were introduced into the most recent Amber force field, Amber ff19SB, by Tian *et al.* who reported significantly improved backbone profiles of all 20 amino acid residues using residue-specific training implicit-solvent QM data[18]. The question addressed in this work is to which extent these mainstream MD force fields capture the intrinsic Ramachandran distributions of individual amino acid residues in water.

Substantial evidence accumulated over the past two decades shows that alanine has an unusually high propensity for polyproline II (pPII) state (see the review by Toal and Schweitzer-Stenner and references therein[19]). Note that in this context, pPII and other secondary structure propensities should be understood as local rather than associated with an entire or any significant region of the unfolded peptide under consideration. Older versions of MD force fields do not reproduce this high pPII content of alanine residue well. For example, Meral and collaborators reported that OPLS-AA/L[20,21] fails to capture intrinsic conformational preferences of most guest residues x in cationic GxG peptides in water[22]. Subsequently, the backbone dihedral potential functions of Amber, OPLS, and CHARMM force fields were modified so as to produce an increased pPII sampling of alanine-based short unfolded peptides. Despite significant progress in the MD force field development, discrepancies between experimental and MD-derived secondary structure propensities of amino acid residues still persist.

Whereas MD force fields are mostly optimized on extended polypeptides and/or folded proteins, experimental studies typically use short unfolded peptides, such as cationic tripeptides, GxG, to probe intrinsic conformational ensembles of guest residues x in water. It is then important to examine to which extent terminal charges affect the Ramachandran distributions of the guest residues in GxG peptides. Toal et al. demonstrated that terminal charges do not affect the conformational ensembles of central residues in AAA, GAG and GVG[23]. The central aspartic acid residue in GDG or DDD appears to be an exception in a sense that its conformational ensemble depends on pH. However, Milorey *et al.* recently reported that the pH-effects are mostly due to the protonation/deprotonation of the side chain of this residue[24]. Suitability of model GxG peptides as representative of intrinsic conformational dynamics of guest residues is further supported by an observation that $^3J(H^N\ H^{C_\alpha})$ constants of x-residues in GGxGG and GxG are very similar[25].

One issue, which is often brushed aside, is how many experimental constraints are needed to uniquely determine the Ramachandran distribution of guest amino acid residues x in GxG peptides in water. Zhang *et al.* recently compared six MD force fields, including Amber ff14SB, CHARMM36m, and the recent version of OPLS, OPLS-AA/M[12], with respect to their ability to capture available spectroscopic data that included a set of NMR scalar coupling constants, amide I' band profiles from infrared spectra (IR), Raman and vibrational circular dichroism (CD) spectra, on the central alanine residue in cationic GAG and AAA peptides[26]. To address the issue of how well such a set of spectroscopic data determine the Ramachandran distribution, Schweitzer-Stenner developed a Gaussian modeling approach, in which the Ramachandran distribution of the guest amino acid residue in GxG peptide (or any other tripeptide) is described as a superposition of two-dimensional Gaussian sub-distributions related to different conformational states, such as pPII, β-strand, right- and left-handed helices, and various turns, whereby the parameters of these subdistributions are optimized to best fit the experimental data[27]. Zhang and collaborators showed that such experiment-based Gaussian model of the central alanine in GAG and AAA in water outperforms all six MD force fields under study[26]. Of all MD force field/water model combinations that were evaluated, Amber ff14SB (combined with TIP3P water model) performed better than the other force fields[26]. A follow-up study, Andrews *et al.* asked whether alanine's preference for the pPII state stems from the backbone or alanine's side chain characteristics[28]. To this end, the experiment-based Gaussian modeling was applied to the central glycine residue in cationic GGG as a model of a peptide backbone, followed by the corresponding assessment of Amber ff14SB (with TIP3P), OPLS-AA/M (with TIP4P), and CHARMM36m (with the respective TIP3P model)[28]. The results revealed that pPII is driven by the tendency of water to maximize the number of hydrogen bonds between water and the functional groups of the central glycine, demonstrating that high pPII preference of amino acid residues stems from the water-backbone interactions[28]. Interestingly, Amber ff14SB was outperformed by the other two force fields with respect to its capacity to capture conformational dynamics of glycine residue in water[28]. Several recent studies revealed that CHARMM36m combined with its specialized TIP3P water model performs better than the other additive force fields, although there is still room for improvement[10,29].

The inability to accurately reproduce intrinsic conformational propensities of amino acid residues in unfolded peptides in water represents a potentially major weakness of MD. While capturing conformational propensities of residues in short peptides may not be consequential *per se*, short unfolded peptides are building blocks of longer unfolded regions of IDPs. Potential inaccuracies in these building blocks could propagate to larger systems, limiting the applicability of MD as a tool for studies of IDPs. Most recent MD force fields have not been validated against spectroscopic data for amino acid residues in short unfolded peptides in water other than glycine and alanine. In part, this is due to the lack of experimental data that have not been available until recently[30–34]. Here, we systematically assess four recent MD force fields: OPLS-AA/M (with TIP4P), CHARMM36m, Amber ff14SB (with TIP3P), and Amber ff19SB[18] (recommended to use with four-point OPC water model[35]) with respect to their capacity to capture conformational ensembles of 14 guest amino acids x in cationic GxG peptides in water in a way that is consistent with available J coupling constants and amide I' profiles. Just like in the case of the assessment of conformational dynamics of glycine and alanine residues[26,28,29], we utilize the experiment-based Gaussian modeling as a benchmark comparison. We envision that the comparison of experimental and MD results on guest residues x in GxG peptides in water is a starting point of a "bottom-up" approach to the assessment of MD force fields, which should significantly contribute to the future force field development.

Methods

Molecular dynamics simulations

Tripeptides GxG, where x = G, A, L, V, I, F, Y, protonated D (D^P), protonated E (E^P), R, C, N, S, or T, were constructed using the Visual Molecular Dynamics (VMD) software package[36]. Each tripeptide is immersed into a 64 nm^3 cubic box with periodic boundary conditions within GROMACS 5.1.2 package[37–43]. The following four combinations of force fields and water models are used: Amber ff19SB[18] with four-point OPC water model[35], Amber ff14SB[13] with TIP3P[44], OPLS-AA/M[12] with TIP4P[44], and CHARMM36m[6,11,45,46] with TIP3P. Additional simulations of six GxG peptides with guest residues x = A, I, F, D^P, R, S were acquired using two additional force field/water model combinations: Amber ff14SB with TIP4P/2005[47] and OPLS-AA/L[21]–based residue-specific force field RSSF1[48] with TIP4P/Ew[49]. In each simulation, the tripeptide under consideration is protonated at the N terminus (NH_3^+) and neutral at the C terminus. Within OPLS-AA/M and CHARMM36m as well as RSSF1, the C terminus is capped with a carboxyl group (COOH), whereas an amino group (NH_2) capping of the C terminus is applied in Amber force fields. Electrostatically neutral systems are obtained by adding positive (Na^+) or negative (Cl^-) ions as needed. The Verlet cutoff scheme[50] and a time step of 2 fs are used during the equilibration and production steps. Steepest descent is performed during the energy minimization for 100,000 time steps, followed by a 20 ps pressure equilibration step at 300 K and 1.0 bar. 300 ns-long trajectories are acquired during the production runs at temperature 300 K with the velocity rescale thermostat[51] and Berendsen barostat[52].

Analysis

MD-derived Ramachandran distributions. MD-derived Ramachandran distributions are calculated within GROMACS 5.1.2 using time frames (separated by 2 ps) within 50-300 ns of each MD trajectory. Ramachandran distributions are constructed by subdividing the Ramachandran space into 32400 bins ($2^\circ \times 2^\circ$) and calculating the respective local per-bin probabilities.

J-coupling constants and amide I' band profiles. The experimental data that we use to assess the MD-derived Ramachandran distributions of guest residues x in GxG peptides in water were reported earlier[30–33,53]. The experimental data set for each guest residue x encompasses the ϕ-dependent scalar coupling constants ${}^3J(H^N,H^{C\alpha})$, ${}^3J(H^N,C')$, ${}^3J(C,C')$, ${}^3J(H^{C\alpha},C')$, the ψ-dependent constant ${}^1J(N,C_\alpha)$, and the amide I' profiles in the respective IR, Raman, and vibrational circular dichroism (VCD) spectra (the prime sign indicates that the vibrational spectra are acquired for GxG peptides in D_2).

Gaussian modeling and Gaussian Ramachandran distributions. The Gaussian model, previously developed by Schweitzer-Stenner[27], is used as a benchmark model of the Ramachandran distribution for each guest residue x in GxG peptides. Briefly, Gaussian modeling is based on a linear Gaussian decomposition, whereby a Ramachandran distribution is modeled as a superposition of two-dimensional Gaussians associated with distinct local secondary structure states (such as pPII, right- and left-handed helix, turns, parallel and antiparallel β, and various turns). The Gaussian model parameters (relative statistical weights of Gaussian sub-distributions, their location in the Ramachandran space, and their corresponding widths) are then adjusted to best fit the available J-coupling constants and amide I' band profiles[27]. The optimization of the Gaussian model parameters is carried out by using the empirical Karplus equations[54–57] and an exitonic coupling formalism to calculate conformational averages of J-coupling constants and amide I' band profiles[27,30,32,58?]. In the current study, some Gaussian model parameters for guest residues x in GxG peptides are revisited and slightly modified. In the evaluation of the MD-derived Ramachandran distributions, the same algorithm described above is used to calculate force field-specific MD-derived J-coupling constants and amide I' band profiles. Consistent with MD-derived Ramachandran distributions, Gaussian Ramachandran distributions are also constructed by subdividing the Ramachandran space into 32400 bins ($2^\circ \times 2^\circ$) and calculating the respective local per-bin probabilities.

Reduced χ^2 functions. We use a reduced χ_J^2 function to quantitatively assess the ability of MD force fields and the Gaussian model to capture the conformational ensemble of guest residue x in GxG peptides in water:

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

where N is the number of J-coupling constants (in our case five), $J_{i,exp}$ are the experimental J-coupling constants, $J_{i,calc}$ are the calculated J-coupling constants obtained from MD-derived or Gaussian Ramachandran distributions, and s_i are the uncertainties due to the reported experimental errors[59] and the errors associated with the Karplus parameters[55], which are combined using the Gaussian error propagation method. Generally acceptable fits are associated with reduced χ_J^2 values below 2. However, this criterion is applicable only if all errors associated with experimental data are known. Unfortunately, this cannot be guaranteed in the current work because the error values are not available for all utilized J-coupling constant parameters. In such cases, we use the experimental uncertainties alone, which are typically smaller than the propagation errors due to fitting that determines the Karplus parameters, leading to an overestimation of χ_J^2 values.

In analogy to the reduced χ_J^2 function, a reduced χ_{VCD}^2 function is defined to evaluate the MD force fields and the Gaussian model with respect to their capacity to reproduce the experimental amide I' band profile of guest residue x:

$$\chi_{VCD}^2 = \frac{1}{N'} \sum_{k=1}^{N'} \frac{(\Delta\varepsilon_{exp,k} - \Delta\varepsilon_{calc,k})^2}{s_k^2} \quad (2)$$

where N' is the number of wavenumbers from 1600 to 1720 cm^{-1}, $\Delta\varepsilon_{exp,k}$ is the experimental value for all k considered, and s_k are experimental errors derived from an analysis of a spectral region dominated by noise. Note that due to the achiral nature of triglycine, there is no experimental VCD signal for glycine as the guest residue x in GxG; in our previous work, we calculated χ_{VCD}^2 for glycine by assuming that $\Delta\varepsilon_{exp,k}$ are zero for all k due to an inversion symmetry about the center of the corresponding Ramachandran distribution[28].

Definition of mesostates and mesostate populations. Mesostates are used as a visual guide for comparison of Ramachandran distributions derived from the expriment-based Gaussian model and MD simulations. The following mesostate definitions are used: (a) pPII ($-90^\circ < \phi < -42^\circ$, $100^\circ < \psi < 180^\circ$), (b) anti-parallel β-strand (aβ) ($-180^\circ < \phi < -130^\circ$, $130^\circ < \psi < 180^\circ$), (c) the transition region between aβ and pPII (βt) ($-130^\circ < \phi < -90^\circ$, $130^\circ < \psi < 180^\circ$), (d) right-handed α-helix ($-90^\circ < \phi < -32^\circ$, $-60^\circ < \psi < -14^\circ$). Each mesostate population is calculated by adding per-bin Ramachandran distribution probabilities within the corresponding mesostate region. For achiral triglycine, the $(-\phi, -\psi)$ conformations are identical to the corresponding (ϕ, ψ) conformations. Consequently, for the central glycine in triglycine, the mesostate populations are obtained by adding the respective left-handed and right-handed populations.

Shannon entropy. The Gaussian and MD-derived Ramachandran distributions are also used to calculate the Shannon entropy for each guest residue x in GxG in water:

$$S = -R \sum_{\phi,\psi} P(\phi,\psi) \ln P(\phi,\psi) \quad (3)$$

where R is the gas constant and P(ϕ, ψ) are the respective per-bin Ramachandran distribution probabilities ($2^\circ \times 2^\circ$ bins) that the guest residue x in GxG will occupy a state within the bin corresponding to backbone dihedral angles ϕ and ψ.

To facilitate the comparison of the Gaussian and MD-derived Ramachandran distributions, we also derive:

$$\Delta S_{I,GxG} = S_{x,MD} - S_{x,Gaussian} \quad (4)$$

where x is the guest residue in GxG under investigation. Similarly, within each category (Gaussian model or MD force field), the differences in the distributions between amino acid residues and alanine can be directly assessed by calculating:

$$\Delta S_{II,GxG} = S_{x,i} - S_{A,i} \quad (5)$$

where x is the guest amino acid, A is alanine, and i is the Gaussian model or an MD force field. Whereas ΔS_I allows for a direct quantitative comparison of MD-derived and Gaussian Ramachandran distributions, ΔS_{II} offers a measure of guest residue-specificity of the Ramachandran distribution, which is expected to be MD force field-dependent.

Results

In this study, we assess four MD force fields: OPLS-AA/M[12] combined with TIP4P water, CHARMM36m combined with the corresponding TIP3P[6], Amber ff14SB combined with TIP3P[13], and Amber ff19SB[18] combined with OPC water model[35], in terms of their capacity to capture intrinsic conformational dynamics of 14 guest residues x (x = G, A, L, V, I, F, Y, D^P, E^P, R, C, N, S, or T) in GxG peptides in water against the available experimental data sets. Hereafter, the term force field is used to broadly refer to a complete set of interaction parameters, including those involving water. Each guest residue-specific experimental data set includes five J-coupling constants — $^3J(H^N,H^{C_\alpha})$, $^3J(H^N,C')$, $^3J(H^N,C_\beta)$, $^3J(H^{C_\alpha},C')$, and $^1J(N,C_\alpha)$[59] — and amide I' profiles from IR and Raman spectra[53]. Note that the guest residues K and M[30] as well as H[60] in GxG peptides, for which only $^3J(H^N,H^{C_\alpha})$ and amide I' band profiles are available, are excluded from this study. The entire MD-derived Ramachandran distribution of the guest residue x in GxG peptides in water (see *Methods*) is used to calculate the five J-coupling constants and amide I' profiles to facilitate a direct comparison to experimental data. As a benchmark comparison, experiment-based Gaussian modeling is applied to each experimental data set. In the Gaussian model, the Ramachandran distribution is constructed as a linear combination of Gaussian sub-distributions, representing distinct secondary structure states, whereby the respective weights, heights, and widths are optimized to best fit the experimental data (see *Methods*)[27]. The resulting distribution produced by Gaussian modeling is hereafter referred to as the Gaussian distribution. As a visual guide, we also introduce four mesostates: pPII, βt, aβ, and α, corresponding to specific regions in the Ramachandran space (as defined in *Methods*). It is important to note that the assessment of the Gaussian and MD-derived Ramachandran distributions against experimental data is independent of the somewhat arbitrary definition of the mesostates.

Gaussian model better reproduces J-coupling constants and amide I' band profiles than MD force fields

We here quantitatively assess Gaussian modeling and four MD force fields with respect to their capacity to reproduce five J-coupling constants and VCD amide I' band profiles of guest residues in GxG peptides in water. To this end, we calculate the J-coupling constants and amide I' band profiles from guest-residue specific Gaussian Ramachandran distributions and the corresponding Ramachandran distributions derived from each MD force field. All experimental and calculated (Gaussian modeling- and MD-derived) J-coupling constants are reported in Table S1. Convergence of J-coupling constants in MD trajectories is evaluated by monitoring time evolution of the five J-coupling constants, averaged over 50 ns time intervals, for the three guest residues with the largest side chains: isoleucine (Fig. S1), arginine (Fig. S2), and tyrosine (Fig. S3) for a total simulation time of 500 ns per trajectory. Because the entire Ramachandran distribution of the guest residue contributes to the five J-coupling constants[26], time dependence of J-coupling constants is a reflection of time dependence of the Ramachandran distribution, providing a measure of sampling efficiency and, by implication, convergence. Figs. S1-3 demonstrate that while the J-coupling constants fluctuate with simulation time, they do not display significant drifts in their respective average values. The largest fluctuations in the MD-derived J-coupling constants are observed in Amber ff19SB for all three guest residues, whereas J-coupling constants' fluctuations observed in Amber ff14SB ($^3J(H^{C_\alpha},C')$ for Y in GYG, Fig. S3c) and CHARMM36m ($^1J(N,C_\alpha)$ for I in GIG, Figs. S1e and $^3J(H^{C_\alpha},C')$ for R in GRG, Fig. S2c) trajectories are smaller. Importantly, data on Figs. S1-S3 demonstrate that extending the total simulation time per trajectory to 500 ns is not required for convergence, in line with previously published studies[26,28,29,58]. MD-derived J-coupling constants and all amide I' profiles for all guest residues x in GxG peptides under study are thus calculated from MD trajectories using the time interval 50-300 ns.

The Gaussian model and MD force fields are further evaluated by using the two reduced χ^2-functions, χ^2_J and χ^2_{VCD}, defined in *Methods*. Fig. 1a shows χ^2_J values for 14 guest amino acid residues x in GxG peptides in water. The corresponding χ^2_{VCD} values are displayed in Fig. 1b for all guest residues x other than glycine, for which the reduced χ^2_{VCD} cannot be calculated due to unknown experimental values, as reported previously[28]. The comparison in Fig. 1 demonstrates that the Gaussian model produces convincingly low χ^2_J and χ^2_{VCD} values, mostly in an excellent agreement with experimental data. In comparison to the Gaussian model predictions, all four MD force fields result in significantly larger χ^2_J and χ^2_{VCD} values, demonstrating significantly less favorable agreement with experimental data.

Amber ff19SB results in the highest χ^2_J values for 10 of the 14 guest residues (L, V, I, F, Y, D^P, R, C, N, and S) and does not produce the lowest value for any guest residue (Fig. 1a). In comparison, Amber ff14SB produces the highest χ^2_J values for 2 guest residues (G and E^P) and the lowest χ^2_J values for 3 guest residues (A, L, and T). OPLS-AA/M produces the lowest χ^2_J value for 9 guest residues (V, I, F, Y, E^P, C, N, S, and T) and the next lowest in 2 other residues (G and D^P). Notably, OPLS-AA/M yields the highest χ^2_J and second highest χ^2_{VCD} of the four MD force fields for alanine residue in GAG, which is unexpected considering that OPLS-AA/M uses glycine- and alanine-based model systems to calibrate the backbone dihedral potentials[12]. CHARMM36m does not produce the highest χ^2_J value for any guest residue and results in the lowest χ^2_J value for 3 guest residues (G, D^P, and R), suggesting that it may be the most balanced of the four force fields with respect to NMR data.

One might argue that the choice of Karplus parameters affects the above comparisons. Whereas Fig. 1a is based on the Karplus parameter set reported by Wirmer and Schwalbe (Wirmer-Schwalbe)[56], we also calculated χ^2_J values using the Karplus parameters published by Ding and Gronenborn (Ding-Gronenborn)[57]. Note that switching between these two Karplus parameter sets in our case only affects the calculated $^1J(N,C_\alpha)$ values. The comparison of the χ^2_J values derived by using the Wirmer-Schwalbe and Ding-Gronenborn Karplus parameter sets is shown in Fig. S4a and b, respectively. Except for glycine and alanine, the Ding-Gronenborn set produces overall larger χ^2_J values. Aside from this increase in χ^2_J values, no significant changes are observed in the relative performance between the Gaussian model and MD force fields when the Wirmer-Schwalbe set is replaced by the Ding-Gronenborn set. In rare cases the choice of Karplus parameters alters the ranking of MD force fields with respect to the χ^2_J values. The change from Wirmer-Schwalbe to Ding-Gronenborn Karplus parameter set changes the force field ranking with respect to χ^2_J values for residues A, E^P, and S; however, these changes do not strongly affect the overall MD force field comparison, because the four force fields produce comparable χ^2_J for these residues (Fig. S4).

Fig. 1b shows χ^2_{VCD} values for each guest residue calculated from the Gaussian and MD-derived Ramachandran distributions. As in the case of NMR data, the Gaussian model also outperforms all MD force fields in reproducing the VCD amide I' band profile. Again, the largest χ^2_{VCD} values are associated with Amber ff19SB. Of the four force fields, Amber ff19SB exhibits the highest χ^2_{VCD} values for 11 guest residues (A, L, V. I, F, D^P, E^P, R, C, N, and T) and produces the lowest χ^2_{VCD} value for serine (S). Amber ff14SB exhibits the highest χ^2_{VCD} value for tyrosine (Y) and the lowest χ^2_{VCD} values for six guest residues (G, A, L, F, C, and N). OPLS-AA/M yields the highest χ^2_{VCD} value only for serine (S) and produces the lowest χ^2_{VCD} values for two guest residues (I and T). CHARMM36m results in the highest χ^2_{VCD} values for two guest residues (G and S) and the lowest χ^2_{VCD} values for four guest residues (V, Y, D^P, and R).

Because χ^2_J values are more sensitive to the dihedral angle ϕ, whereas χ^2_{VCD} is sensitive to both dihedral angles, we added up both χ^2 values to get a better sense of relative performances of the MD force fields (Fig. 1c). The comparison of the combined χ^2 values indicates that conformational ensembles of three guest residues (D^P, C, T) are particularly poorly captured by all four MD force fields. Amber ff14SB, OPLS-AA/M, and CHARMM36m exhibit comparable performance on the remaining 11 guest residues, whereas Amber ff19SB shows larger deviations across most guest residues. In the following, we examine the

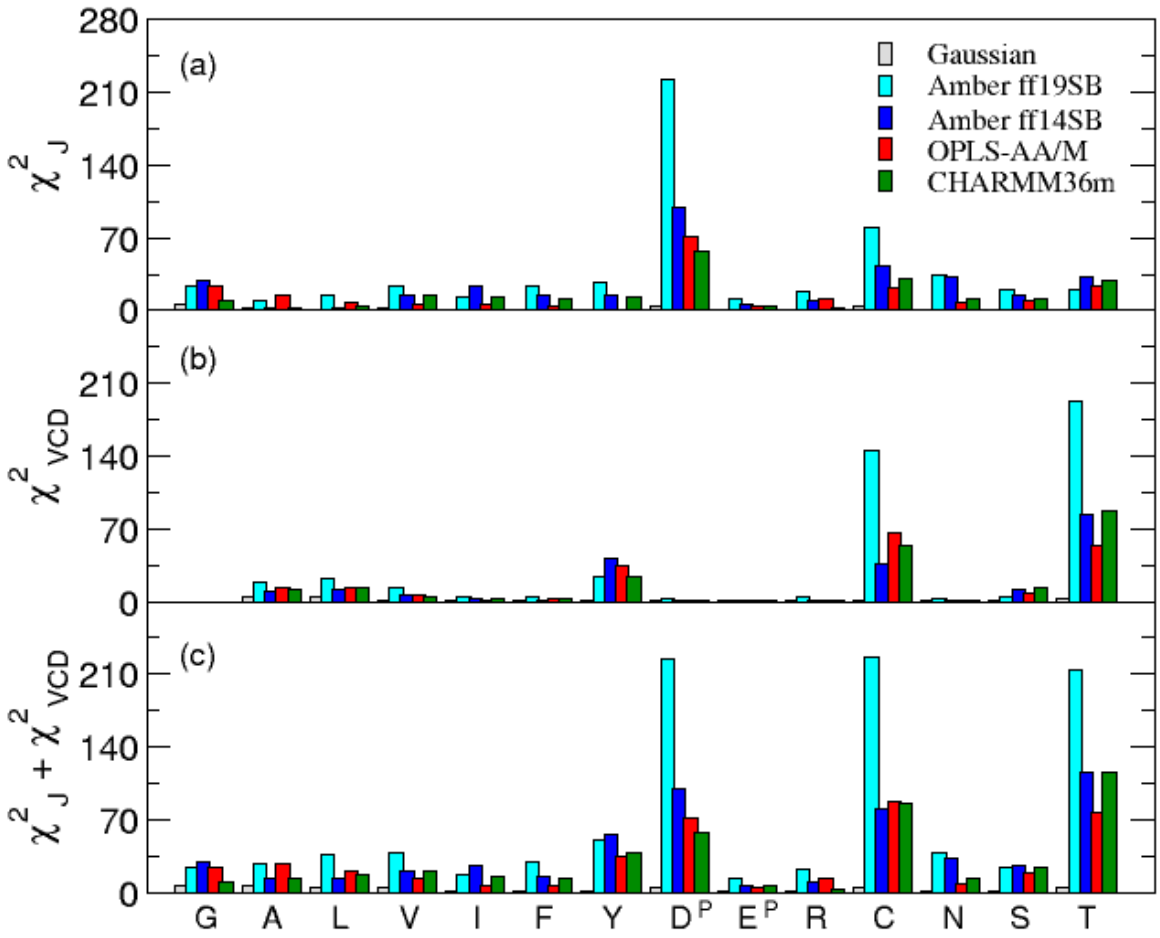

Fig. 1 Assessment of the Gaussian model and MD force fields with respect to their ability to reproduce the experimental data for the guest residue in cationic GxG in water using (a) χ^2_J and (b) χ^2_{VCD} functions. Values presented for GGG and GAG, aside from those for Amber ff19SB, are taken from two previous studies[26,28]. Reduced χ^2_{VCD} values for GGG are not available for the central glycine in GGG, as reported previously[28].

conformational ensembles of different groups of guest residues in more detail and demonstrate that χ^2 comparison alone may be overly simplistic and may not provide sufficient information to aid MD force field development efforts.

Assessment of MD force field-specific conformational preferences of guest residues x in GxG peptides in water

The quantitative analysis described in the following provides an evaluation of the recent MD force fields with respect to intrinsic conformational dynamics of 14 amino acid residues in water. Populations of distinct mesostates, albeit arbitrarily defined, can aid in identifying the qualitative strengths and weaknesses of each force field when compared to the Gaussian model predictions. The benchmark Gaussian distributions serve as visual guides in the assessment of the MD-derived Ramachandran distributions. Populations of the mesostates, defined in *Methods*, are calculated from the Gaussian and MD-derived Ramachandran distributions and reported in Table S2. We classify the guest residues into aliphatic (G, A, L, V, I), aromatic (F, Y), ionizable (D^P, E^P, R), and polar (C, N, S, T) groups and describe the results for each group individually.

Aliphatic amino acid residues. Here, we examine the central G in GGG and guest residues A, L, V, and I in GxG, which are characterized by side chains consisting exclusively of hydrocarbons.

Ramachandran distributions and mesostate populations. Fig. 2 shows the Ramachandran distributions produced by the Gaussian model and MD simulations with four force fields (panel a) and the corresponding populations of mesostates (panel b), as defined in *Methods*. The Gaussian distributions in Fig. 2a exhibit significant differences within this group of guest residues. The main trend is that pPII conformations which dominate the conformational ensembles of guest glycine and alanine, are less abundant for guest residues L, V, and I in favor of increased β-strand populations. In contrast, Ramachandran distributions in Fig. 2a derived from Amber ff14SB, OPLS-AA/M, and (to a lesser extent) CHARMM36m are similar to each other and do not exhibit significant changes in the abundance of pPII conformations. Of the four MD force fields, Amber ff19SB produces the most noticeable guest residue-specific Ramachandran distributions. All MD-derived Ramachandran distributions, including those obtained within Amber ff19SB, deviate from the respective Gaussian distributions. The β-strand basin in MD-derived distributions is dominated by aβ conformations

in contract to βt conformations, which are predominant in the Gaussian model. MD force fields other than OPLS-AA/M produce overly abundant right-handed helical conformations. These features are reflected in the respective mesostate populations in Table S2. Of the four force fields, only OPLS-AA/M produces βt and right-handed helical populations comparable to Gaussian model predictions (Fig. 2b and Table S2), however, the Ramachandran distribution difference among A, L, V, and I are very minor, inconsistent with guest-residue specific Gaussian distribution features (Fig. 2a). Fig. 2b shows that while MD force fields still underestimate the pPII populations of guest glycine and alanine, they actually overestimate the pPII populations in guest leucine, valine and isoleucine, most likely due to a lack of sensitivity to the guest residue. An exception is Amber ff19SB, which at least for leucine and valine residues accounts for the respective pPII populations in agreement with the Gaussian model predictions.

Glycine. Three MD force fields (Amber ff14SB, OPLS-AA/M, and CHARMM36m) were recently evaluated with respect to the capacity to reproduce experimental data for the central glycine residue in GGG in water[28]. Fig. 1 and Fig. S5g-h demonstrate that Amber ff19SB simulations of the central glycine in GGG result in a somewhat lower χ_J^2 value than Amber ff14SB, but produce a strongly increased χ_{VCD}^2 value, indicating that Amber ff19SB insufficiently accounts for the achiral character of the central glycine residue in GGG. Absolute errors with respect to the respective experimental J-coupling constants are shown in Fig. S5. The vibrational CD amide I' profile (Fig. S5i) calculated from Amber ff19SB data is comparable in magnitude to that derived from CHARMM36m, but produces a positive couplet, which contrasts that of CHARMM36m. Fig. S6 demonstrates that Amber ff19SB, just like Amber ff14SB. reproduce the amide I' profiles sufficiently well. Fig. 2b and Table S2 demonstrate that Amber ff19SB produces less βt and aβ and a comparable amount of pPII and α content as the Gaussian model. ΔS_I values in Table S3 demonstrate that of all four MD force fields, Amber ff19SB produces a Ramachandran distribution with the smallest relative Shannon entropy with respect to the Gaussian Ramachandran distribution (-1.02 J/mol/K). This result suggests that the mesostates in Amber ff19SB have comparable widths and weights as in the Gaussian model, however, as the Shannon entropy is invariant to locations of the mesostates in the Ramachandran space, small ΔS_I may still produce high χ^2 values as both the J-coupling constants and amide I' profiles are sensitive to mesostate locations. Amber ff19SB also results in the relative Shannon entropy between the central glycine and alanine, $\Delta S_{II} = 2.86$ J/mol/K, which is the most comparable to the Gaussian model prediction, $\Delta S_{II} = 3.24$ J/mol/K).

Alanine. Similarly as for glycine, several MD force fields, including Amber ff14SB, OPLS-AA/M, and CHARMM36m, were previously assessed with respect to their ability to capture conformational ensembles of the central alanine in GAG and AAA in water[29]. We here add Amber ff19SB to this comparison. Of the four MD force fields, Amber ff19SB produces the second highest χ_J^2 and highest χ_{VCD}^2 for guest alanine residue in GAG, whereas Amber ff14SB produces the lowest values of these functions. Absolute errors with respect to the experimental J-coupling constants are displayed in Fig. S7. Fig. S7i shows an almost completely diminished VCD amide I' profile compared to the Gaussian model prediction and amide I' profiles derived from the other three force fields. The amide I' profiles produced by Amber ff19SB in Fig. S8 are similar to those produced by Amber ff14SB. Similar to the results for the central glycine in GGG, a notable decrease in pPII and βt populations are observed in Amber ff19SB relative to Amber ff14SB (Fig. 2b and Table S2). Just like for glycine, of the four force fields, Amber ff19SB produces the lowest absolute value of ΔS_I (-0.61 J/mol/K) for the guest alanine in GAG (Table S3).

Leucine. Whereas Gaussian Ramachandran distributions of alanine and leucine bear some similarities, the distribution of leucine contains slightly less pPII and more βt that the Gaussian distribution of alanine (Fig. 2a). All MD force fields other than Amber ff19SB overestimate the pPII population of leucine residue in GLG (Fig. 2b). The absolute differences of J-coupling constants with respect to their experimental values are displayed in Fig. S9a-e. Among the MD force fields, Amber ff19SB and Amber ff14SB produce the highest and the lowest χ_J^2 and χ_{VCD}^2 values, respectively (Fig. S9g-h). In comparison to the Gaussian modeling, the VCD amide I' profiles are underestimated by MD force fields (Fig. S9i). In contrast, the IR and Raman profiles of amide I' are sufficiently well reproduced by the Gaussian model as well as the four force fields (Fig. S10). Table S3 shows that Amber ff14SB produces the lowest ΔS_I (0.08 J/mol/K) and the closest ΔS_{II} to that of the Gaussian model (0.75 J/mol/K and 1.25 J/mol/K for Amber ff14SB and Gaussian model, respectively).

Valine. Whereas guest alanine and leucine share some similarities of their respective Gaussian Ramachandran distributions, the Gaussian Ramachandran distribution of guest valine in GVG clearly deviates from the other two (Fig. 2a). In comparison to alanine, valine shows significantly reduced pPII and increased β-strand content, whereby β-strand conformations in the Gaussian distribution of valine are dominated by βt. The MD force fields do not reproduce this trend very well. All force fields other than Amber ff19SB significantly overestimate the pPII population (Fig. 2b). The absolute differences of J-coupling constants with respect to their experimental values are displayed in Fig. S10a-e and the χ^2 values for J-coupling constants and VCD amide I' profiles are shown in Fig. 11g-h. OPLS-AA/M produces the lowest χ_J^2 value of the four force fields under study and the second lowest χ_{VCD}^2 value; CHARMM36m results in the lowest χ_{VCD}^2 value. VCD amide I' profiles are underestimated by all four force fields (Fig. 11i), whereas the IR and Raman amide I' profiles are adequately captured (Fig. 12). Shannon entropy results displayed in Table S3 indicate that OPLS-AA/M yields the lowest ΔS_I value (-0.17 J/mol/K) as well as the ΔS_{II} value (-5.15 J/mol/K), which is the closest to the one predicted by Gaussian modeling (-3.49 J/mol/K).

Isoleucine. The Gaussian Ramachandran distribution of isoleucine in GIG is characterized by significantly reduced pPII content (similar to the case of valine) and increased β-strand content, which is dominated by parallel β-strand (pβ) rather than βt conformations (Fig. 2a). The prominent sampling of the parallel β-strand region is unique to isoleucine as no other guest residue in this study samples this region. Moreover, unlike in the case

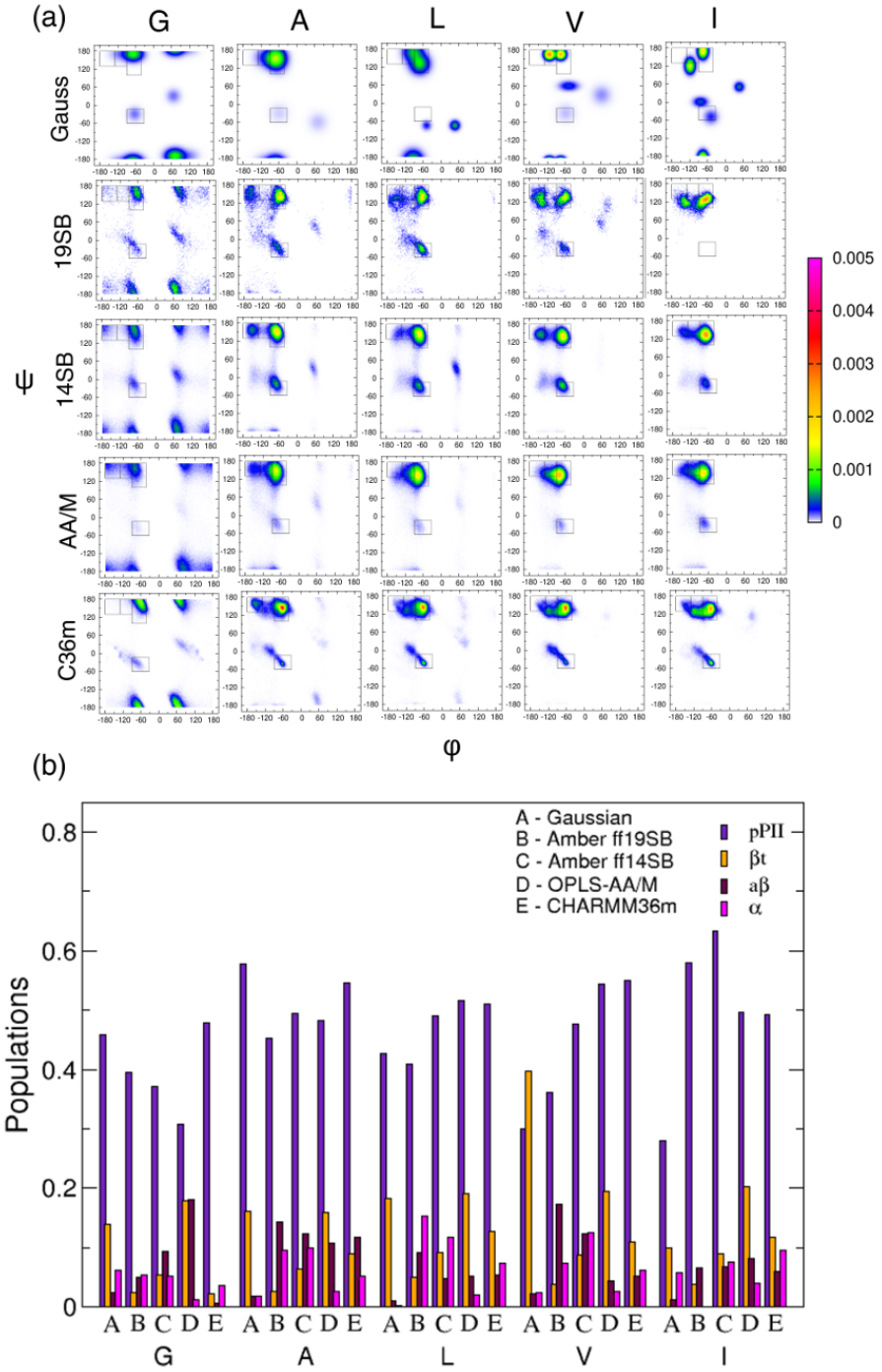

Fig. 2 (a) Gaussian and MD-derived Ramachandran distributions and (b) mesostate populations of aliphatic guest residues x = G, A, L, V, I in GxG peptides in water. Ramachandran distributions contain rectangular regions corresponding to mesostates defined in *Methods*. Values for glycine and alanine (except those derived within Amber ff19SB) are taken from two previous studies[26,28].

of other aliphatic residue, the pPII and pβ basins are clearly separated. In addition, isoleucine residue samples type I/II' β-turn and left-handed helical turn conformations to a greater extent than any other aliphatic guest residue. The mesostate populations in Fig. 2b for the Gaussian model show that isoleucine has the lowest pPII and βt contents of all aliphatic residues. All four MD force fields overestimate the pPII content of isoleucine (Fig. 2b) and none of them accounts for isoleucine's unique propensity for pβ conformations (Fig. 2a). Fig. S13a-e shows the absolute differences between the calculated and experimental J-coupling constants. Of the four MD force fields, OPLS-AA/M produces the lowest and Amber ff19SB results in the largest values of χ_J^2 and χ_{VCD}^2 (Fig. S13g-h). All force fields, and Amber ff19SB more so that the other three force fields, underestimate the VCD amide

I' profile (Fig. S13i), but reproduce the IR and Raman amide I' data adequately well (Fig. S14). Interestingly, although OPLS-AA/M reproduces the experimental data better than the other three force fields, the respective Ramachandran distribution is rather distinct from the Gaussian distribution as it lacks sampling in the pβ and the two turn regions (Fig. 2a). While the comparison of χ^2 values indicates that the Gaussian distribution reproduces the experimental data better than any MD force field, we cannot exclude a theoretical possibility that there is more than one distribution that reproduces the experimental data as well as the Gaussian distribution described above. Data reported in Table S3 demonstrate that OPLS-AA/M produces the lowest magnitude ΔS_I value (-0.25 J/mol/K) and the ΔS_{II} value (-2.74 J/mol/K), which is most similar to that predicted by the Gaussian model (-1.00 J/mol/K).

Aromatic amino acid residues. A comparison between phenylalanine and tyrosine residues is valuable, because these two residues have aromatic side chains, but only in tyrosine the hydroxyl group is attached to the aromatic ring, rendering tyrosine more hydrophilic than phenylalanine.

Ramachandran distributions and mesostate populations. Ramachandran distributions and mesostate populations for two aromatic residues, phenylalanine and tyrosine, are displayed in Fig. 3a and b, respectively. Unlike in the case of the aliphatic group, where chemically similar side chains produced quite distinct Gaussian distributions, only minor differences are observed between the Gaussian distributions of F in GFG and Y in GYG, in agreement with comparable values of experimental J-coupling constants (Table S1). Whereas Gaussian distribution for phenylalanine exhibits pPII and β conformations with overlapping basins, these two basins are more separated in the case of tyrosine. pPII populations as predicted by the Gaussian model are similar for both guest residues (Fig. 3b). Interestingly, a recent study of single and double protonated histidine residue in GHG in water revealed an equidistribution in the pPII and β-strand regions similar to the ones for F and Y[60]. The Gaussian distributions of F and Y also feature minor basins of conformations on the right side of the Ramachandran plot ($\phi > 0$).

Phenylalanine. Consistent with the Gaussian model predictions, MD-derived Ramachandran distibutions for F in GFG display the pPII and β-strand basins on the left-hand side of the Ramachandran plot, but the β-strand basin is shifted to the more negative ϕ values relative to the Gaussian distribution for all force fields (Fig. 3a). Only in OPLS-AA/M and CHARMM36m the two basins overlap, in agreement with the Gaussian model. In comparison to Gaussian model, all MD force fields slightly overestimate the pPII and strongly underestimate the βt content (Fig. 3b). Amber ff19SB captures the pPII populations in the best agreement with the Gaussian model, whereas OPLS-AA/M results in the highest βt population among the four force fields. The absolute differences of J-coupling constants with respect to their experimental values are displayed in Fig. S15a-e. Among the four force fields, OPLS-AA/M and Amber ff19SB produce the lowest and the highest χ^2_J values, respectively (Fig. S15g). Fig. S15h shows the χ^2_{VCD} values produced by the four force fields. Whereas Amber ff14SB

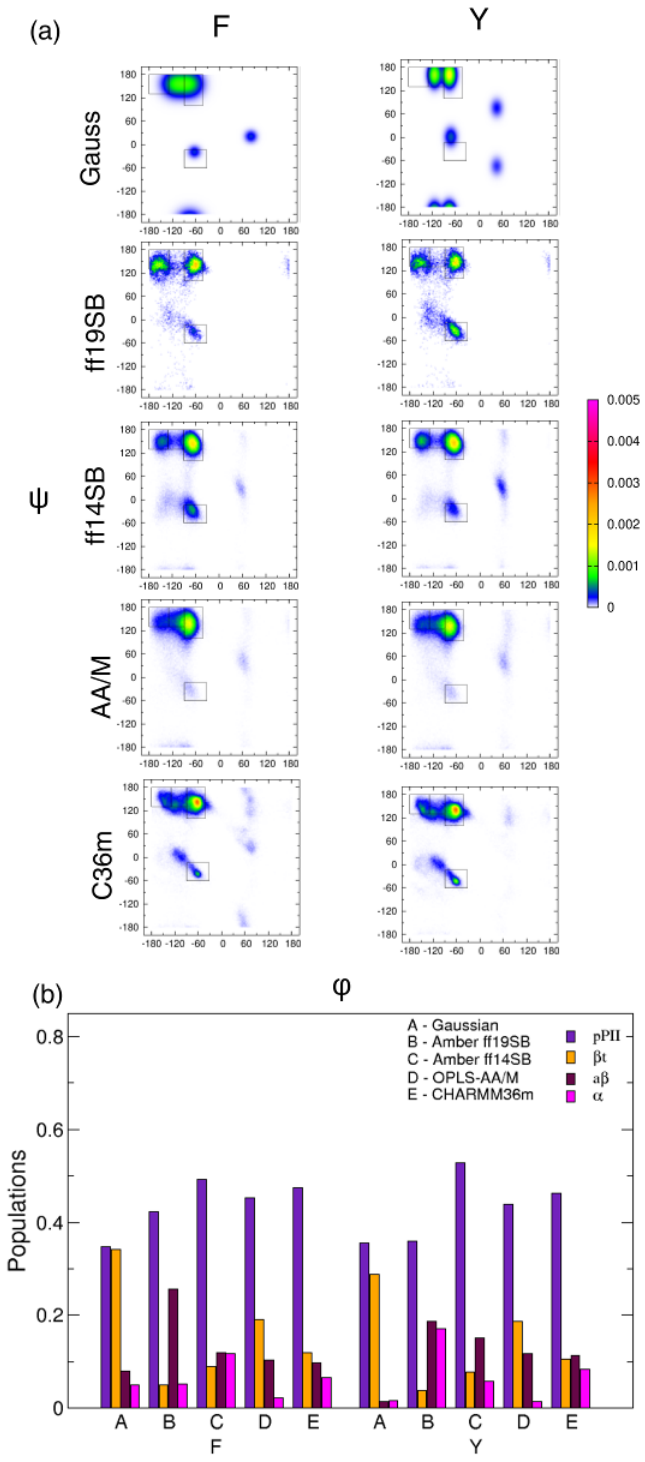

Fig. 3 (a) Gaussian and MD-derived Ramachandran distributions and (b) mesostate populations of aromatic guest residues x = F, Y in GxG peptides in water. Ramachandran distributions contain rectangular regions corresponding to mesostates defined in *Methods*.

yields the lowest χ^2_{VCD} value, the other force fields, in particular Amber ff19SB, underestimate the experimental VCD amide I' profiles (Fig. 15i). The IR and Raman amide I' data are overall

well reproduced by MD force fields, whereby OPLS-AA/M deviates from the experimental data more than the other three force fields (Fig. S16). Shannon entropy differences in Table S3 (row 6) indicate that Amber ff19SB produces that largest absolute value of ΔS_I (-2.39 J/mol/K), suggesting the largest deviation from the Gaussian distribution. The lowest absolute value of ΔS_I is associated with OPLS-AA/M (0.42 J/mol/K) and CHARMM36m produces a comparable value (0.50 J/mol/K). Amber ff19SB yields the largest absolute value of ΔS_{II} (-2.05 J/mol/K), i.e. the Shannon entropy difference between the Ramachandran distributions of phenylalanine and alanine. Notably, ΔS_{II} calculated within the Gaussian model has a significantly smaller absolute value (-0.25 J/mol/K). Of all MD force fields, OPLS-AA/M results in the ΔS_{II} value of -1.33 J/mol/K, which deviates the least from the Gaussian model prediction.

Tyrosine. MD-derived Ramachandran distibutions for Y in GYG exhibit well-separated pPII and β-strand basins on the left-hand side of the Ramachandran plot that do not differ much from Ramachandran distributions for F in GFG, (Fig. 3a). Whereas Amber ff19SB captures the pPII population of Y in GYG in agreement with Gaussian modeling, all force fields strongly underestimate the βt population, whereby OPLS-AA/M yields the highest βt content (Fig. 3b). Gaussian model predicts significantly lower right-handed helical content for Y in GYG than for F in GFG and this difference is not captured in MD. Consequently, all MD force fields other than OPLS-AA/M overestimate the right-handed helical content for Y in GYG (Fig. 3b). Fig. S17a-e shows the absolute differences of J-coupling constants with respect to their experimental values. Among the MD force fields, only OPLS-AA/M produces a reasonably low χ_J^2 value (Fig. S17g), whereas all force fields result in very large χ_{VCD}^2 values (Fig. S17h) as they do not capture the VCD amide I' profiles at all (Fig. S17i). The experimental VCD amide I' profile of Y in GYG (Fig. S17i) is significantly different from that of F in GFG (Fig. S15i), however, these differences may not have a structural origin. The VCD amide I' profile of tyrosine residue is peculiar because the couplet is heavily positively biased. This can only be accounted for by assuming an intrinsic magnetic transition dipole moment for the N-terminal peptide group. According to Nafie and collaborators such intrinsic moments can be caused by electric currents along loops formed by covalent and hydrogen bonds[61]. We took into account the same magnetic transition moment for the N-terminal peptide group of Y in GYG in Gaussian modeling, which is why the Gaussian model reproduces the experimental VCD amide I' profile reasonably well. Fig. S18a and c demonstrates the Gaussian and MD-derived Ramachandran distributions capture the isotropic Raman and IR amide I' profiles adequately, although OPLS-AA/M shows the most prominent deviations from experimental data. The Gaussian model outperform the MD force fields with respect to their capacity to reproduce the anisotropic Raman amide I' profile (Fig. S18b). Shannon entropy differences calculated for tyrosine show that all MD force fields produce comparable and low ΔS_I values ($\sim$0.50 J/mol/K). The value of ΔS_{II} produced by Amber ff19SB is in the best agreement with the Gaussian model prediction (-0.76 J/mol/K and -0.91 J/mol/K, respectively). In contrast, the value of ΔS_{II} calculated for OPLS-AA/M (-2.99 J/mol/K) deviates the most from the value predicted by the Gaussian model.

Ionizable amino acid residues. Here the results for protonated aspartic acid, protonated glutamic acid, and arginine guest residues in respective GxG peptides are described. Notably, only arginine has a charged side chain at acidic pH, which mimics the experimental conditions.

Ramachandran distributions and mesostate populations. The Gaussian distributions for ionizable residues in Fig. 4a reveal a distinct distribution of D^P, whereas distributions of E^P and R look quite similar. All MD-derived Ramachandran distributions strongly deviate from Gaussian distributions as reflected in large χ^2 values (Fig. 1). The Ramachandran distributions derived within OPLS-AA/M and Amber ff14SB do not exhibit much residue specificity. Only Amber ff19SB qualitatively captures the distinct distributions of D^P and E^P, while producing similar distributions of E^P and R. The MD-derived mesostate populations alongside their Gaussian model predictions are displayed in Fig. 4b.

Protonated aspartic acid residue. Previous studies reported that D^P in GD^PG in water exhibits a peculiar Gaussian Ramachandran distribution, which is dominated by β-strand rather than pPII content and is characterized by an above the average populations of asx-turn (the upper right quadrant of the Ramachandran plot) and type I/II' β-turn ($i+2$) (located slightly above the canonical right-handed helical basin in the Ramachandran plot). The asx-turn formation involves hydrogen bonding between the side chain of D^P and the C-terminal amide proton in GD^PG. None of the four MD force fields captures the asx-turn conformations. The Ramachandran distributions derived from both Amber force fields exhibit prominent sampling within the left-handed helical basin. Instead of type I/II' β-turn ($i+2$) conformations, MD sampling within all force fields in this general region tends to favor right-handed helical conformations. While Amber ff19SB captures the pPII population in agreement with the Gaussian model, it strongly underestimates the βt population (Fig. 4b). The other three force fields clearly overestimate the pPII content. Only OPLS-AA and CHARMM36m result in βt populations comparable to those derived within the Gaussian model. The J-coupling constants are not well reproduced by any MD force field (Figs. S19a-e and Fig. 1). CHARMM36m and Amber ff19SB produce the lowest and highest χ_J^2 values, respectively (Fig. S19g). The origin of large χ_J^2 values can be inferred from Fig. S19a-e. Amber ff19SB results in a particularly large deviation from the experimental $^1J(N,C_\alpha)$ value (Fig. S19e), which indicates that the pPII and β-strand basins are located at too low ψ values. Because the Karplus parameters fitting errors for this J-coupling constant are not available, the respective uncertainty used in the χ_J^2 calculation is purely due to experimental errors and thus smaller, which increases the relative contribution of deviations from this experimental coupling constant to χ_J^2. The other large contribution to χ_J^2 values stems $^3J(H^{C_\alpha},C')$ (Fig. S19d) and reflects the lack of conformations in the upper right quadrant (within the asx-turn basin) of the Ramachandran plot, in particular within OPLS-AA/M and CHARMM36m (Fig. 4a). Amber ff19SB in addition samples in a region with $\phi < -100^\circ$ and ψ-values close

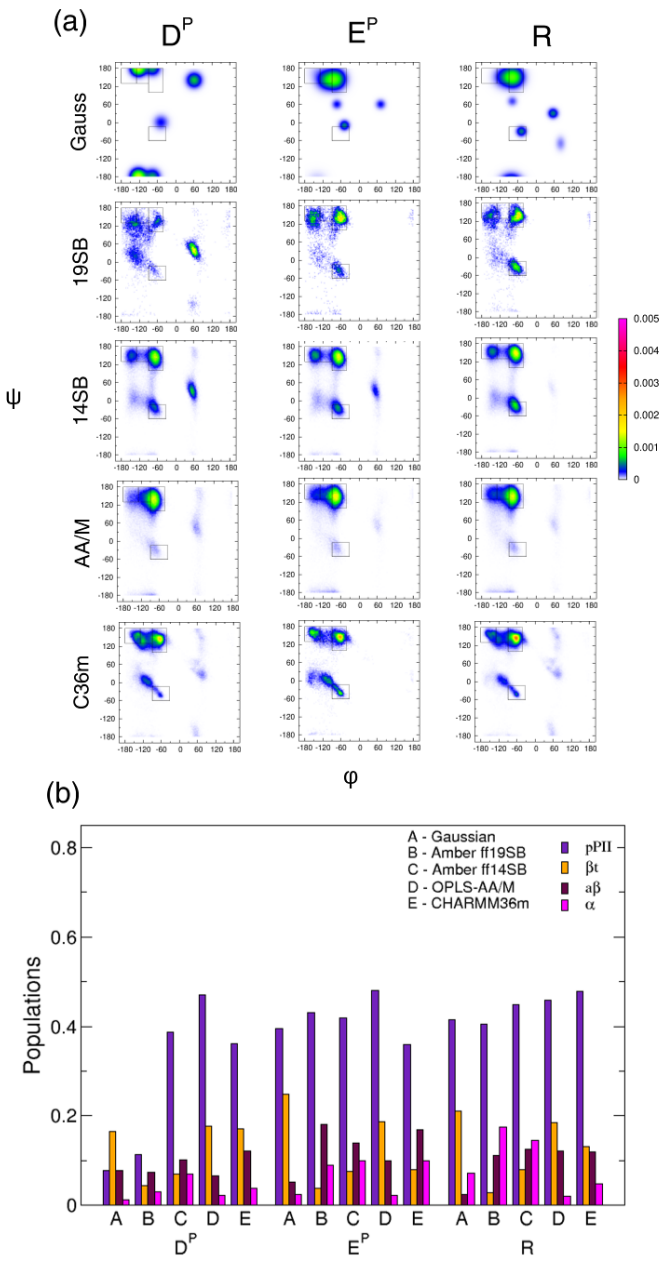

Fig. 4 (a) Gaussian and MD-derived Ramachandran distributions and (b) mesostate populations of ionizable guest residues x = D^P, E^P, R in GxG peptides in water. Ramachandran distributions contain rectangular regions corresponding to mesostates defined in *Methods*.

to 0°, which produces a large deviation from the experimental $^3J(H^N,C')$ value (Fig. S19b). Fig. S19h,i demonstrates that with an exception of Amber ff19SB, the other three MD force fields reproduce the VCD amide I' profile well (Fig. S19i). However, this agreement with experimental data is an artefact caused by the overestimation of the pPII content, which compensated for the lack of asx-turn conformations. Fig. S20 shows that peaks in the isotropic Raman and IR amide I' profiles derived from the Gaussian model and MD force fields are slightly shifted to lower wavenumbers compared to experimental data. Only the Gaussian model reproduces the isotropic Raman and IR amide I' profiles adequately, whereas Amber ff19SB and OPLS-AA/M exhibit the largest deviations from experimental data (Fig. S20). The Shannon entropy difference ΔS_I has the lowest absolute value within OPLS-AA/M (0.33 J/mol/K), reflecting the most similarity with the Gaussian model, and the highest absolute value within Amber ff14SB (2.58 J/mol/K). The value of ΔS_{II} deviates the most from the corresponding Gaussian model prediction (0.58 J/mol/K) for Amber ff14SB (2.58 J/mol/K), and the least for OPLS-AA/M (-0.58 J/mol/K).

Protonated glutamic acid residue. The Gaussian Ramachandran distribution of the guest E^P in GE^PG is notably distinct from that of D^P in GD^PG (Fig. 4a). The Gaussian modeling reveals that pPII mesostate is somewhat more populated than all β-strand mesostates combined. In addition, turn-forming conformations: type I/II' β-turn $i+2$ and inverse γ-turn ($\phi < 0$) and left-handed helical conformations ($\phi > 0$) are sampled to a significant extent. Notably, none of the four MD force fields produces sampling in the inverse γ-turn region and only Amber ff14SB samples the left-handed helical basin. All MD force fields produce comparable pPII populations to that predicted by the Gaussian model (Fig. 4b). While OPLS-AA/M and CHARMM36m produce βt populations in qualitative agreement with the Gaussian model, both Amber force fields underestimate βt and overestimate both aβ and right-handed helical populations (Fig. 4b). Of the MD force fields, OPLS-AA/M captures the J-coupling constants the best (Fig. S21a-e), resulting in the lowest χ_J^2 value (Fig. S21g). Apparently, a relatively good correspondence between Gaussian model and MD-derived populations is not sufficient to yield also low χ_J^2 values. All force fields except for Amber ff19SB well reproduce the VCD amide I' profiles, resulting in low χ_{VCD}^2 values (Fig. S21h). Measured and calculated Raman and IR amide I' profiles in Fig. S22 show that CHARMM36m-derived isotropic Raman amide I' profiles and Amber ff14SB-derived anisotropic Raman and IR amide I' profiles deviate the most from experimental data. Amber ff14SB produces a diminished peak in the lower wavenumber region compared to the other profiles (Fig. S22). In terms of Shannon entropy difference ΔS_I, OPLS-AA/M (-0.74 J/mol/K) exhibits the most resemblance to the Gaussian model predictions and Amber ff19SB (-3.32 J/mol/K) deviates from the Gaussian model the most. With respect to ΔS_{II} values, Amber ff19SB (-2.01 J/mol/K) deviates the most from the Gaussian model prediction (0.70 J/mol/K) and CHARMM36m produces the value (0.41 J/mol/K), which is best aligned with the Gaussian model.

Arginine residue. The Gaussian Ramachandran distribution of R in GRG resembles that of E^P in GE^PG (Fig. 4a). The mesostates have comparable populations of pPII and β-strand. Nonetheless, the pPII basin is more populated in the Gaussian distribution of R at the expense of combined β-strand conformations, which is reflected in the respective mesostate populations (Fig. 4b). The Gaussian distribution also features left-handed helical and inverse γ-turn populations (Fig. 4a), which are not captured by MD force fields. While OPLS-AA/M best reproduces the βt content, it underestimated the right-handed helical content (Fig. 4b). In comparison, both Amber force fields overestimate the right-handed helical population. Fig. S23a-e shows the absolute deviations of

calculated J-coupling constants from their experimental values. CHARMM36m produces the lowest χ_J^2 value of all MD force fields (Fig. 1 and Fig. S23g). Of the four force fields, CHARMM36m produces the lowest χ_{VCD}^2, followed by Amber ff14SB (Fig. S23h). The VCD amide I' profiles in Fig. 23i demonstrate that the Gaussian model best reproduces the experimental profiles, even the asymmetric ones, whereas MD-derived VCD amide I' profiles are all underestimated, in particular within Amber ff19SB. Gaussian and MD-derived Raman and IR amide I' profiles in Fig. S24 all exhibit quite large deviations from experimental data. Amber ff19SB, Amber ff14SB, and CHARMM36m deviate the most from the experimental isotropic amide I' profile while OPLS-AA/M shows the largest deviation from the experimental anisotropic Raman and IR profiles (Fig. S24). With respect to the Shannon entropy calculations, the lowest and highest absolute values of ΔS_I, which is a measure of Ramachandran distribution's similarity to the Gaussian distribution, are associated with Amber ff14SB (-0.91 J/mol/K) and Amber ff19SB (-2.49 J/mol/K), respectively. The Gaussian model predict $\Delta S_{II}$of arginine residue, which measures similarity of the guest residue to alanine residue, to be 1.50 J/mol/K. CHARMM36 (0.58 J/mol/K) deviates the least and OPLS-AA/M (-1.33 J/mol/K) deviates the most from this Gaussian prediction.

Polar amino acids. In this category of polar (but not ionizable) guest residues, we examine MD force fields with respect to conformational dynamics of C, N, S, and T in GxG peptides in water. This category contains residues with the most diverse side chains. Cysteine has a thiol side chain with a very low polarity; it is often classified as a hydrophobic amino acid. The side chain of asparginine is a carboxamide with more hydrogen bonding capacity than serine and threonine. Serine and threonine are derivatives of alanine and valine with one hydrocarbon group in the side chain replaced by a hydroxyl group.

Ramachandran distributions and mesostate populations. Gaussian distributions for all four guest residues in this group showcase rather balanced pPII and βt conformational basins, although some of the pPII basins lie outside the pPII mesostate region defined in *Methods* (Fig. 5a). The pPII basin is the most populated in the Gaussian distribution of N, while βt basin dominates the Gaussian distribution of T. These preferences are reflected in mesostate populations in Fig. 5b. Fig. 5a also reveals that all four residues sample various turn regions of the Ramachandran space: type I/II' β-turn ($i+2$), type I'/II β-turn ($i+2$), type III' (left-handed helix), and asx-turn to a significant extent.

The Ramachandran distributions obtained from MD simulations in Fig. 5a show that none of the four force fields adequately reproduces the Gaussian distributions of these four residues. MD-derived pPII mesostate populations are mostly overestimated (Fig. 5b). The exceptions are pPII populations of C in GCG and N in GNG obtained within Amber ff19SB. With respect to βt populations, only OPLS-AA/M captures these conformations, whereas the other three force fields (in particular Amber ff19SB) favor aβ over βt populations. Ramachandran distributions obtained within OPLS-AA/M exhibit almost no residue specificity; only threonine's distributions is slightly different from the other

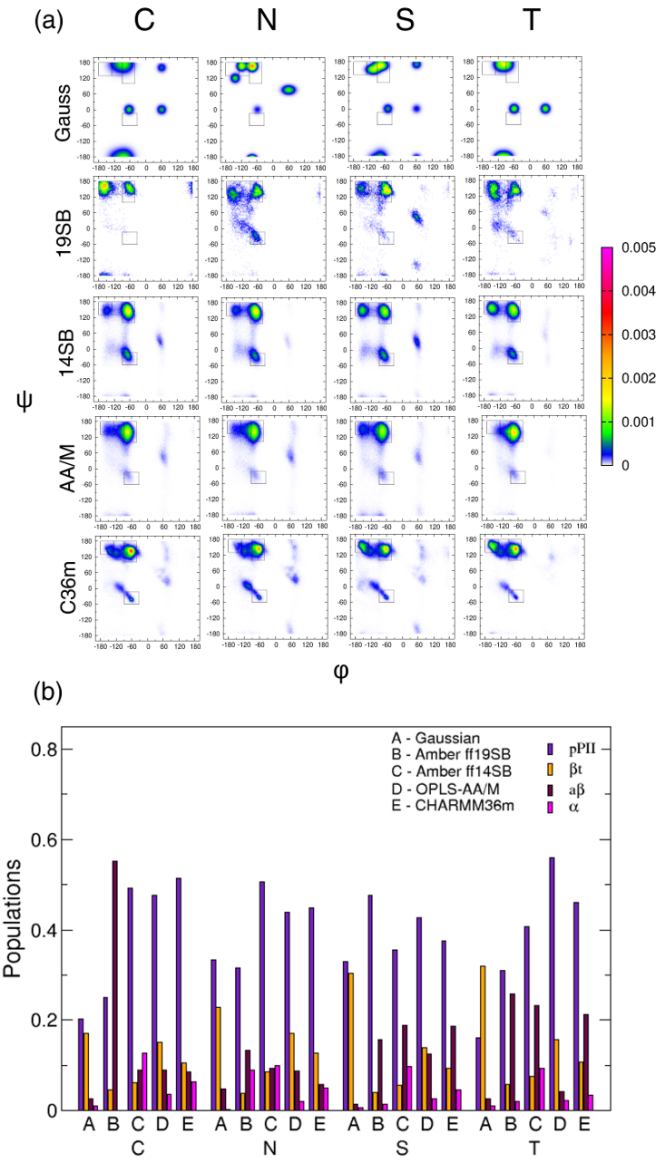

Fig. 5 (a) Gaussian and MD-derived Ramachandran distributions and (b) mesostate populations of polar guest residues x = C, N, S, T in GxG peptides in water. Ramachandran distributions contain rectangular regions corresponding to mesostates defined in *Methods*.

three. Ramachandran distributions derived within Amber ff14SB and CHARMM36m show more guest residue specificity, whereas Amber ff19SB shows the most. However, the MD-derived residue-specific differences of Ramachandran distributions do not resemble those predicted by the Gaussian model. MD-derived Ramachandran distributions (other than those of C in GCG and S in GSG obtained within Amber ff19SB), display some sampling of the type I/II' β-turn region that only partially overlaps with the right-handed helix basin. Importantly, MD force fields mostly fail to capture the sampling in the right side of the Ramachandran space ($\phi > 0$), specifically in turn-forming regions of type I'/II β-turn ($i+2$), type III' (left-handed helix), and asx-turn.

Cysteine residue. The Gaussian distribution for guest C in GCG

produces the third lowest pPII population among the residues under study, in part due to a shift of the pPII basin in the positive ψ direction (Figs. 5) relative to that of A in GAG (Figs. 2). This shift is not sufficiently accounted for by MD force fields, leading to large χ^2 values (Fig. 1). Of the four MD force fields, only Amber ff19SB accurately captures a decrease in the pPII population (Fig. 5b). The Gaussian distribution exhibits significant sampling in the type I/II' β-turn ($i+2$), type I'/II β-turn ($i+2$), and asx-turn regions. The latter two conformational basins (on the right side of the Ramachandran plot) are not produced by MD simulations (Fig. 5a). The absolute values of deviations of the calculated J-coupling constants from their experimental values are displayed in Fig. S25a-e. The lowest and highest χ_J^2 values are generated by OPLS-AA/M and Amber ff19SB, respectively (Fig. S25g). The VCD amide I' profiles are the best reproduced by Amber ff14SB, whereas Amber ff19SB yields the largest χ_{VCD}^2 value (Fig. S25h,i). Raman and IR amide I' profiles in Fig. S26 indicate that the Gaussian and MD-derived Raman profiles are shifted slightly to lower wavenumbers relative to the experimental profiles (Fig. S26a-b). Gaussian and MD-derived IR amide I' profiles (Fig. S26c) have overestimated peak heights, whereby the profile obtained from the Gaussian distribution deviates from the experimental data the least. The Shannon entropy difference, ΔS_I, assumes the lowest absolute value for OPLS-AA/M (0.14 J/mol/K) and the highest for Amber ff19SB (-6.95 J/mol/K). The Shannon entropy difference, ΔS_{II}, measuring the similarity of cysteine to alanine, assumes the value of 1.41 J/mol/K within the Gaussian model. The ΔS_{II} values obtained for Amber ff19SB (-4.94 J/mol/K) and Amber ff14SB (0.67 J/mol/K) deviate the most and the least, respectively, from the Gaussian model prediction.

Asparginine residue. The Gaussian distribution of N in GNG differs from the Gaussian distributions of the other three polar guest residues the most (Fig. 5). While the pPII basin of N is the most sampled among the polar guest residues under study, the Gaussian distribution is characterized by a comparable sampling of βt conformations. Note that the Gaussian distribution of N is unique due to prominent sampling of the extended β-strand region. Turn-forming conformations within type I/II' β-turn ($i+2$) and left-handed helical regions of the Ramachandran space also contribute to the conformational ensemble of this guest residue. Of the four force fields, OPLS-AA/M results in the lowest χ_J^2 values as shown in more detail in Fig. S27a-e, where the absolute values of the difference between the calculated and experimental J-coupling constants are depicted (Fig. S27g). While the Gaussian model best reproduces the experimental VCD amide I' profiles (Fig. S27h,i), the MD-derived profiles show significantly larger deviations, whereby Amber ff14SB and Amber ff19SB yield the lowest and the highest χ_{VCD}^2 values, respectively (Fig. S27h). The Amber ff19SB-derived profile produces an opposite couplet compared to experimental data. Raman and IR amide I' profiles in Fig. S28 indicate that a sharp peak in the experimental isotropic Raman amide I' profile around 1690 cm^{-1} is not captured by the Gaussian model or any MD force field. The IR profile is much better reproduced by the Gaussian model than by the MD force fields (Fig. S28c). Otherwise, the profiles are reproduced fairly well, whereby OPLS-AA/M deviates the most from experimental data. Notably, a combined value of χ_J^2 and χ_{VCD}^2 is the lowest for OPLS-AA/M, however, CHARMM36m is not lagging far behind (Fig. 1c). Shannon entropy calculations reveal that the lowest absolute value of ΔS_I is associated with Amber ff19SB (4.51 J/mol/K), although all force fields produce large deviations from the Gaussian model. All ΔS_{II} values strongly deviate from the Gaussian model prediction of -3.91 J/mol/K. Here, CHARMM36m (2.07 J/mol/K) deviates the most, whereas Amber ff14SB and OPLS-AA/M produce the lowest deviation (0.25 J/mol/K) from the Gaussian model prediction.

Serine residue. The Gaussian distribution of guest S in GSG is dominated by comparable sampling of pPII and βt basins, followed by sampling of type I/II' β-turn ($i+2$), type I'/II β-turn ($i+2$), and asx-turn regions, similar to the case of C in GCG (Fig. 5a). The MD-derived pPII content is overestimated, even in the case of Amber ff19SB (Fig. 5b and Table S2). As observed for other guest residues, βt and aβ populations are strongly underestimated and overestimated, respectively, by all force fields (Fig. 5b and Table S2). Absolute deviations of the calculated J-coupling constants from the experimental data are displayed in Fig. S29a-e. OPLS-AA/M produces the lowest χ_J^2 value, while the χ_J^2 value from CHARMM36m is not much larger (Fig. S29g). The Gaussian-derived VCD amide I' profile reproduces experimental data the best, although the negative peak is shifted to lower wavenumbers (Fig. S29h, i). The MD-derived VCD amide I' profiles do not reproduce well the peak magnitudes (Fig. S29i). MD-derived amide I' profiles of S in GSG are unique in showing a strong negative bias for wavenumbers where the experimental data shows a positive peak. Of the four force fields, only Amber ff19SB produces the profile with a negative couplet and results in the lowest MD-derived χ_{VCD}^2 value. Raman and IR amide I' profiles in Fig. S30 demonstrate that the Gaussian and MD-derived Raman amide I' profiles reproduce the experimental data well. When assessing the degree of similarity between the MD-derived and Gaussian distributions by Shannon entropy calculations, we observe that the ΔS_I value of lowest absolute value is produced by Amber ff19SB (0.81 J/mol/K), while the other three force fields exhibit larger deviations from the Gaussian model. The Shannon entropy difference of serine's to alanine's distribution, ΔS_{II}, is predicted by the Gaussian model to be -2.24 J/mol/K. In terms of ΔS_{II}, Amber ff19SB and CHARMSS36m deviate the least (-0.84 J/mol/K) and the most (2.99 J/mol/K) from the Gaussian model prediction.

Threonine residue. The Gaussian distribution of central T in GTG shows predominant sampling in the βt rather than pPII basin of the Ramachandran space (Fig. 5a), which is reflected also in the relative mesostate populations (Fig. 5b and Table S2). In addition, the Gaussian distribution of T in GTG reveals prominent sampling of type I/II' β-turn ($i+2$) and type I'/II β-turn ($i+2$) regions (Fig. 5a). While Amber ff19SB reproduces the lowest pPII population of the four force fields, all MD force fields overestimate the pPII populations by at least 10% and strongly underestimate the βt population (Fig. 5b and Table S2). The left-handed type I'/II β-turn ($i+2$) region is not sampled by any MD force field (Fig. 5a), which is in part responsible for poor reproduction of J-coupling constants and VCD amide I' band profiles (Fig. S31). Of

the four force fields, OPLS-AA/M produces the lowest χ^2_J as well as χ^2_{VCD} values (Fig. S31g,h). The Gaussian model reproduces the VCD amide I' profiles better than MD force fields. The MD-derived VCD amide I' profiles do not capture the experimental peak magnitudes (Fig. S31i). The profile derived from Amber ff19SB shows a positively biased signal in the lower wavenumber region, leading to a large χ^2_{VCD} value. Raman and IR amide I' profiles in Fig. S32 indicate that the Gaussian model-derived isotropic Raman amide I' profile best reproduces the experimental data, particularly in the lower wavenumber regions, however, the experimental anisotropic Raman profile proved difficult to reproduce within the Gaussian model. Amber ff19SB deviates the most from the experimental IR amide I' profiles (Fig. S32). The Shannon entropy difference, ΔS_I, assumes the lowest absolute value for OPLS-AA/M (1.00 J/mol/K), suggesting the most similarity with the Gaussian distribution. Similarly, ΔS_{II} values calculated for OPLS-AA/M (-3.24 J/mol/K) and Amber ff14SB (0.17 J/mol/K) exhibit the smallest and largest deviations, respectively, from the Gaussian model value (-2.74 J/mol/K).

Amber ff19SB best captures residue-specific pPII content predicted by the experiment-based Gaussian model

Previous experimental studies shown that the conformational dynamics of guest residues x in GxG peptides in water is dominated by a balance of pPII/β-strand populations. Importantly, the specifics of Ramachandran distribution are strongly side chain-dependent; turn-forming conformations contribute prominently to the dynamics of polar and ionizable guest residues with side chains that possess hydrogen bonding capacity, such as D^P, N, T, and S[31–34]. Interestingly, some sampling of turn-like conformations is observed also for aliphatic guest residues L, V, and I, although their respective side chains are not capable of hydrogen bond formation. The results described above demonstrate that Amber ff14SB, OPLS-AA/M, and CHARMM36m strongly underestimate guest residue specificity of the Ramachandran distributions and do not reproduce turn-forming conformational preferences. This may not be surprising, considering that dihedral potentials in Amber ff14SB and OPLS-AA/M are optimized for short alanine-based peptides, leading to an overestimation of the pPII content of most guest residues (other than alanine, glycine, and arginine) as well as an underestimation of the pPII content of the central alanine in GAG and AAA[26]. Unlike Amber ff14SB and OPLS-AA/M, CHARMM36m and Amber ff19SB both utilize CMAPs in the parameterization of the backbone dihedral potentials and account for more amino acid-specificity in the backbone dihedral potentials. Amber ff19SB, while overall performing the worst of the four force fields with respect to χ^2 functions, produces by far the most distinct guest residue-specific Ramachandran distributions. In particular, Amber ff19SB best captures decreased pPII populations of V, Y D^P, C, and N relative to A (Table S2).

To elucidate the amino acid-specificity aspect of MD force fields, we explore to which extent MD-derived mesostate populations (pPII, combined βt and aβ, and right-handed helix) correlate with the corresponding populations predicted by the Gaussian model. Correlations plots in Fig. 6 indicate only a few correlations between Gaussian model- and MD-derived populations. Gray solid lines in each graph in Fig. 6 serve as a visual guide of ideal correlations and black dot-dashed lines correspond to linear fits of MD data. Linear regression analysis of the pPII populations in Fig. 6a reveals some interesting correlations. As shown in Table 1, the Pearson correlation coefficient r for pPII populations is moderately high for Amber ff19SB ($r = 0.63$) and CHARRM36m exhibits some, albeit weaker, correlations ($r = 0.36$), indicating some agreement with Gaussian model predictions. In contrast, the pPII populations within OPLS-AA/M are weakly negatively correlated with those predicted by the Gaussian model ($r = -0.35$). Amber ff14SB exhibits a moderately high correlation of the combined β-strand population with the Gaussian model value ($r = 0.60$), followed by a weaker correlation in the case of CHARMM36m ($r = 0.41$).

Table 1 Pearson r correlation coefficients between the populations of the pPII, β (= aβ + βt), and α states of the Gaussian model and MD derived Ramachandran distributions.

Force Field	pPII	β	α
ff19SB	0.63	0.16	0.11
ff14SB	0.11	0.60	-0.02
OPLS-AA/M	-0.35	0.22	0.07
CHARMM36m	0.36	0.41	0.08

It is important to take into consideration that the combined β-strand mesostate used in this analysis does not distinguish between aβ and βt populations. Because MD force fields largely overestimate the aβ at the expense of βt content, which is significantly more representative of the experimental data, these correlations are of limited value, because the distinction between βt and aβ is of high relevance in this assessment study. The other MD-derived mesostate populations are not correlated with those predicted by the Gaussian model. Fig. 6c shows α populations, which are still vastly overestimated by most MD force fields, by a factor of $\sim$ 10 in some cases, compared to those in the Gaussian model predictions. Table S2 demonstrates that α populations are overestimated in Amber ff19SB, Amber ff14SB, and CHARMM36m (for guest residues A, L, V, Y, R, S). OPLS-AA/M generally reproduces the α content better than the other three force fields, but fails to account for experimentally observed residue-specificity of the α content.

While the above correlation analysis provides insights that can aid MD force field development, it could be argued that drawing conclusions based on mesostate populations may not be too reliable due to the some freedom in definition of the mesostates. We thus also consider the Shannon entropy differences ΔS_I, which offer a quantitative comparison between the MD-derived and experimental data that does not depend on the definition of the mesostates. As ΔS_I directly compares the MD-derived and Gaussian Ramachandran distributions for each residue, we calculated the root mean square of the sum over all guest residue ΔS_I values

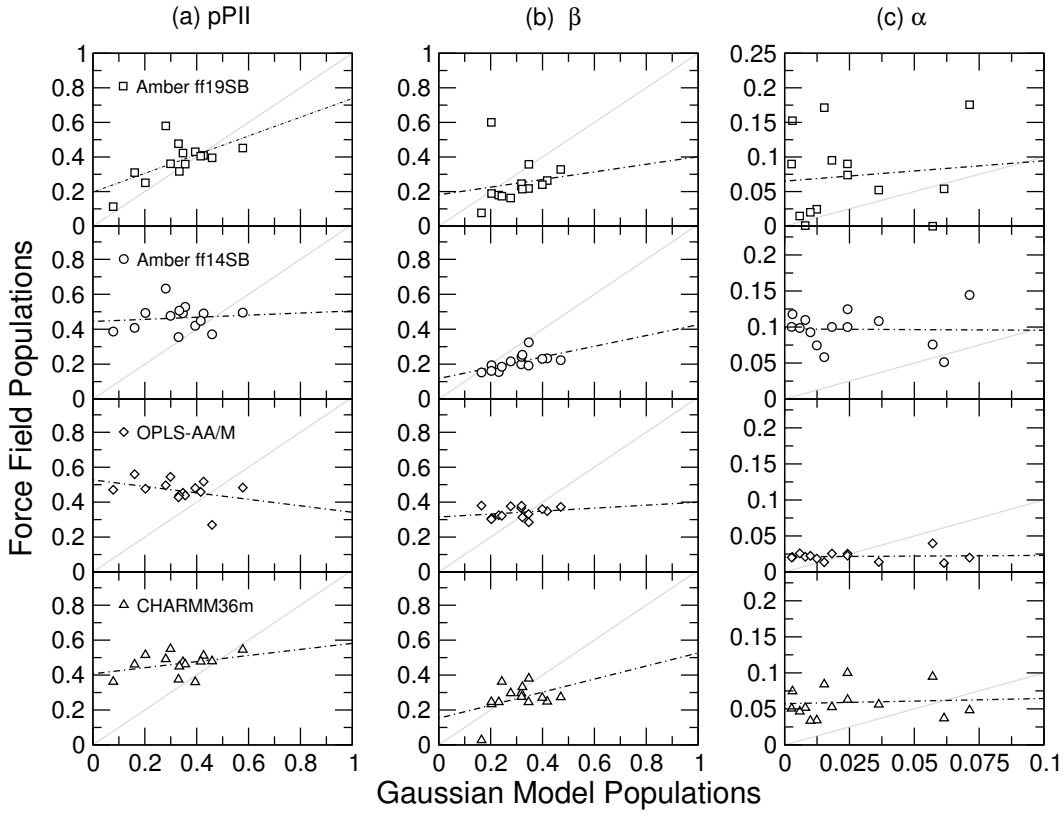

Fig. 6 A comparison of the MD-derived to the Gaussian-derived populations of (a) pPII, (b) combined transitional and antiparallel β-strand, and (c) right-handed helical mesostates (as defined in *Methods*) of the guest amino acid residuesx in GxG peptides. Ideal correlations between the Gaussian model and force field predictions would fall onto a solid gray line with a slope of one. Black dashed lines represent best linear fits. The corresponding Pearson correlation coefficients are reported in Table S4.

(Table 2). The results in Table 2 (second column) demonstrate that CHARMM36m results in the lowest value, indicating the most similarity to Gaussian model prediction with respect to this quantity. A linear regression analysis was then performed comparing MD-derived $\Delta S_I I$ values to those predicted by the Gaussian model. The results of this analysis, shown in Table 2 (third column), indicate some correlations in this quantity for OPLS-AA/M and Amber ff14SB. The interpretation of the analysis of the Shannon entropy differences is complicated by the observation, however, that this quantity is invariant to the location of mesostate basins in the Ramachandran space, and thus cannot be used instead of the two χ^2 functions, which provide the most concise assessment of the MD force fields with respect to spectroscopic data.

Table 2 The root mean square sum of Shannon entropy differences ΔS_I and the Pearson correlations coefficient r of a per-residue comparison between ΔS_{II} values derived from MD (MD) and Gaussian model (GM).

Force Field	$\sqrt{\frac{1}{N}\sum_i \Delta S_{I,i}^2}$	$r(\Delta S_{II,i,MD}, \Delta S_{II,i,GM})$
ff19SB	3.19	0.02
ff14SB	2.84	0.49
OPLS-AA/M	2.61	0.51
CHARMM36m	2.26	0.23

It is difficult to gauge an overall performance of the Gaussian model and the four MD force fields from Fig. 1. For this reason, we calculate the average values and standard deviations of χ_J^2 using two sets of Karplus parameters, by Wirmer and Schwalbe[56] (WS) and Ding and Gronenborn[57] (DG), and χ_{VCD}^2 over all guest residues. The results are reported in Table 3. The Gaussian model clearly outperforms all force fields by one order of magnitude. Replacing the WS set of Karplus parameters by the DG set slightly increases all MD-derived χ_J^2 values. OPLS-AA/M overall outperforms the other three force fields, but CHARMM36m produces χ_J^2 and χ_{VCD}^2 values that are only slightly larger than those for OPLS-AA/M. Amber ff19SB produces on average χ^2 values that are at least twice as large as the respective OPLS-AA/M values.

Moving toward a more accurate MD force field: Residue specificity versus water model?

Taking into consideration that three of the four MD force fields examined above display a severe lack of residue-specificity of the respective Ramachandran distributions, we here asked if an implementation of coil library into the backbone dihedral and side chain torsional potentials would allow for an improved agreement with the spectroscopic data. Jian, Zhou, and Wu developed

Table 3 The average ($\langle\rangle$) and standard deviation (σ) of χ_J^2 values for Wirmer-Schwalbe (WS) and Ding-Gronenborn (DG) sets of Karplus parameters

and χ_{VCD}^2 values calculated over all guest residues under study obtained within the Gaussian model and four MD force fields. Note that the guest glycine in GGG is excluded from the calculation of $\langle\chi_{VCD}^2\rangle$ and σ_{VCD}.

Gaussian/MD	$\langle\chi_{J,WS}^2\rangle \pm \sigma_{J,WS}$	$\langle\chi_{J,DG}^2\rangle \pm \sigma_{J,DG}$	$\langle\chi_{VCD}^2\rangle \pm \sigma_{VCD}$
Gaussian	1.86 ± 1.89	1.84 ± 2.40	1.46 ± 1.42
ff19SB	38.84 ± 55.56	47.85 ± 78.81	34.03 ± 61.21
ff14SB	24.11 ± 24.82	24.22 ± 27.37	15.75 ± 24.26
OPLS-AA/M	15.33 ± 17.85	17.37 ± 22.07	15.64 ± 21.73
CHARMM36m	15.66 ± 14.83	15.69 ± 15.52	16.46 ± 25.58

a residue-specific modification of OPLS-AA/L force field, RSFF1 (combined with TIP4P/Ew water model[49]), using statistical analysis of protein coil libraries derived from high-resolution protein structures that excluded all amino acid residues involved in secondary structure (hydrogen bond) formation[48]. To this end, we selected six guest amino acid residues across the four categories: A, I, F, D^P, R, and S, acquired the respective MD trajectories of cationic GxG peptides in water, and derived the Ramachandran distributions, shown in Fig. 7 (left column). These distributions, which indeed exhibit residue specificity, are then used to calculate the five J-coupling constants and VCD amide I' profiles to facilitate a direct comparison to experimental data and the other MD force fields displayed in Figs. S7, S13, S15, S19, S23, and S29 for guest residues A, I, F, D^P, R, and S, respectively. The χ_J^2 (for the two sets of Karplus parameters) and χ_{VCD}^2 values derived from RSFF1 simulations are compared to the corresponding values obtained from other MD force fields and Gaussian model in Fig. 8. These results indicate that RSFF1-derived χ^2 values are comparable to those derived from OPLS-AA/M and CHARMM36m simulations (Table 4), suggesting that the protein coil library-based calibration of backbone dihedral and side-chain torsional potentials alone does not significantly improve intrinsic conformational dynamics of amino acid residues in water.

It is important to note that Ramachandran distributions derived from coil libraries do not exactly reflect intrinsic propensities of amino acid residues because the amino acid residue of interest is flanked by various neighboring amino acid residues in the coil-library proteins. Coil library-based statistical analysis of backbone conformations of the amino acid residue is based on averaging over many conformational ensembles of the residue of interest, which does not eliminate the influence of nearest neighbor interactions[62]. Even if only GxG sequences in segments of coil-library proteins that do not exhibit helical and β-sheet structures are included in the statistical analysis, the respective Ramachandran distribution is expected to differ from the Ramachandran distribution of guest residue x in GxG tripeptides due to the presence of potential non-local interactions within the loops and turns of structurally-resolved coil-library proteins, which alters its conformational ensemble[33].

Previous studies suggested that the water model play an important role in the MD force field development[10,22,26,58]. Meral

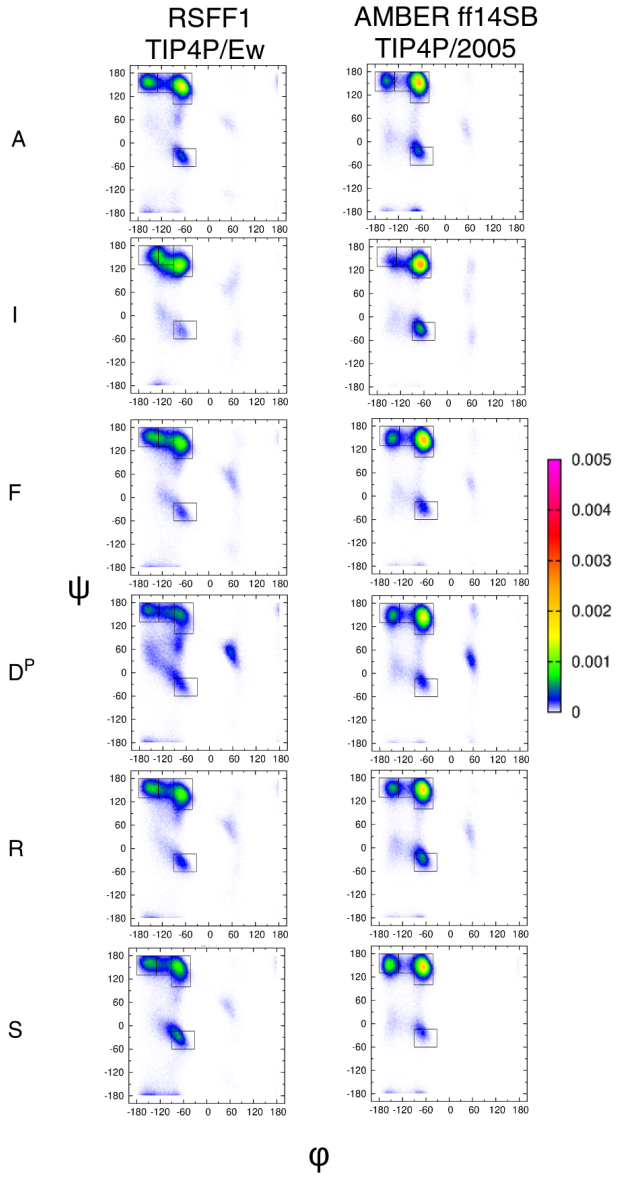

Fig. 7 MD-derived Ramachandran distributions and (b) mesostate populations of aromatic guest residues x = F, Y in GxG peptides in water. Ramachandran distributions contain rectangular regions corresponding to mesostates defined in *Methods*.

et al. tested OPLS-AA/L combined with TIP3P or SPC/E to assess the intrinsic conformational dynamics of 15 guest residues in GxG peptides, trialanine, and alanine dipeptide[22], Zhang *et al.* examined OPLS-AA/L and OPLS-AA/M, each combined with 4 water models with respect to conformational ensembles of central alanine in GAG in AAA[26], and Milorey *et al.* evaluated Amber ff14SB and CHARMM36m combined with TIP3P or OPC for guest arginine in GRG[58]. None of these substitutions lead to a major improvement in the reproduction of experimental data for short oligopeptides. Interestingly, while examining the issue of protein solubility on villin headpiece protein domain, Andrews *et al.* found that replacing TIP3P in Amber ff14SB by TIP4P/2005 significantly increases villin headpiece solvation in agreement with

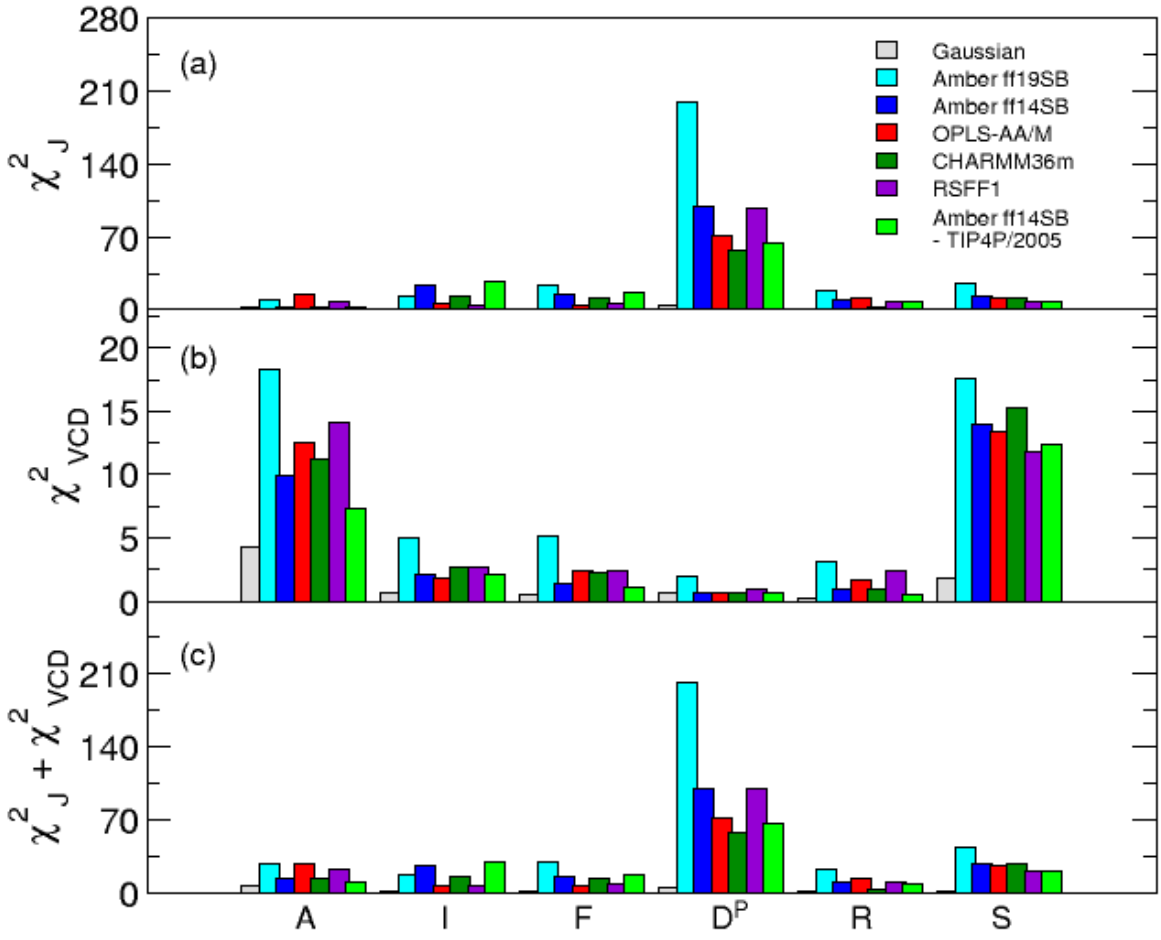

Fig. 8 Assessment of the Gaussian model and MD force fields with respect to their ability to reproduce the experimental data for the guest residues A, I, F, D^P, R, and S in cationic GxG in water using (a) χ^2_J and (b) χ^2_{VCD} functions. Values presented for GAG, aside from those for Amber ff19SB, are taken from a previous study[26].

Table 4 The average ($\langle\rangle$) and standard deviation (σ) of χ^2_J values for Wirmer-Schwalbe (WS) and Ding-Gronenborn (DG) sets of Karplus parameters and χ^2_{VCD} values calculated over six residues (A, I, F, D^P, R, S) under study obtained within the Gaussian model and six MD force fields. Amber ff14SB combined with TIP4P/2005 is denoted as ff14SB**.

Gaussian/MD	$\langle\chi^2_{J,WS}\rangle \pm \sigma_{J,WS}$	$\langle\chi^2_{J,DG}\rangle \pm \sigma_{J,DG}$	$\langle\chi^2_{VCD}\rangle \pm \sigma_{VCD}$
Gaussian	1.36 ± 1.45	1.46 ± 1.23	1.36 ± 1.51
ff19SB	48.24 ± 74.25	71.95 ± 119.43	8.53 ± 7.39
ff14SB	27.24 ± 35.90	32.82 ± 39.01	4.86 ± 5.70
OPLS-AA/M	19.98 ± 25.63	22.72 ± 32.97	5.43 ± 5.87
CHARMM36m	16.54 ± 20.55	18.79 ± 19.74	5.52 ± 6.18
RSFF1	21.87 ± 37.27	17.25 ± 20.71	5.73 ± 5.71
ff14SB**	20.95 ± 23.05	21.54 ± 23.18	4.01 ± 4.80

experimental observations[10]. Encouraged by these findings, we here asked if substituting TIP3P by TIP4P/2005 in Amber ff14SB simulations improves the agreement with spectroscopic data on GxG peptides for the same six guest residues x explored by RSFF1: A, I, F, D^P, R, and S. The MD-derived Ramachandran distributions obtained within Amber ff14SB combined with TIP4P/2005 are presented in Fig. 7. While a visual inspection of Ramachandran distributions does not show any specific residue specificity, the comparison of the corresponding J-coupling constants and amide I' profiles in Figs. S7, S13, S15, S19, S23, and S29 for guest residues A, I, F, D^P, R, and S, respectively, reveals a clear improvement over the predictions of Amber ff14SB/TIP3P. These findings are summarized also in Fig. 8, where all six MD force fields and Gaussian model are assessed with respect to χ^2_J and χ^2_{VCD} values. While these findings are encouraging and in favor of water model-centered force field development, Table 4, which shows the average χ^2_J and χ^2_{VCD} values and variances demonstrates that this improvement is not large when compared to the Gaussian model, indicating again that water model improvement alone, while critical, is not sufficient for significant increase in force field accuracy.

Conclusions and Discussion

In this work we assessed four mainstream MD force fields (Amber ff19SB, Amber ff14SB, OPLS-AA/M, and CHARMM36m) with respect to their capacity to reproduce previously reported spectroscopic data comprising of five J-coupling constants and amide I' band profiles derived from VCD and Raman spectra of 14 guest

amino acid residues x in GxG peptides in water[30–33,53]. Short model peptides GxG are well suited to capture intrinsic conformational preferences of the guest amino acids x in water[19]. Gaussian model, in which the relative weights, widths, and locations of Gaussian sub-distributions corresponding to distinct mesostate basins in the Ramachandran distribution are optimized with respect to the experimental data[27], is used as a benchmark in this assessment study. The limitations of Gaussian modeling stem from the assumption that the mesostate basins can be modelled as a sum of 2D Gaussian sub-distributions, optimized with respect to weights, locations, and widths along ϕ- and ψ axes of the Ramachandran space. The sensitivity of parameters used in the Gaussian modeling was addressed in several earlier studies[24,26,28]. In Gaussian modeling, the basins are ellipsoidal with the two main axes aligned along ϕ and ψ axes. Andrews *et al.* examined the effect of a rotation of the right-handed helical basin in the Ramachandran space, which adds $\phi - \psi$ correlations among conformations within this basin, and concluded that such rotations do not have any identifiable effect on the calculated J-coupling constants and amide I' profiles[28].

The results of our assessment study, which uses reduced χ^2 functions, χ^2_J and χ^2_{VCD}, as quantitative evaluation measures, demonstrate that the Gaussian model reproduces the J-coupling constants and amide I' profiles better than any MD force fields under investigation. None of the four MD force fields captures the experimental data sufficiently well. A comparison of Gaussian and MD-derived Ramachandran distributions of guest residues x in GxG reveals systematic deviations in the latter, i.e. overestimation of the aβ at the expense of βt population (which is the most pronounced in the two Amber force fields), overly sampled right-handed helical conformations (in all force fields other than OPLS-AA/M), and a notorious lack of sampling of various types of turn conformations, which is the most prominent in the case of ionizable and polar guest residues. The lack of turn-forming conformations is specifically visible in MD-derived Ramachandran distributions of guest residues D^P, C, and T, which leads to the largest deviations from experimental data reflected in large χ^2_J and χ^2_{VCD} values. The Gaussian model produces χ^2_J and χ^2_{VCD} values that are on average one order of magnitude smaller than those produced by the four MD force fields. On average, OPLS-AA/M (closely followed by CHARMM36m) produces the lowest χ^2_J and χ^2_{VCD} values, while Amber ff19SB produces the largest χ^2_J and χ^2_{VCD} values.

The above findings indicate a major issue that plagues the MD force fields, which is the lack of amino acid residue specificity of the resulting Ramachandran distributions, indicative of a lack of side chain dependence of the backbone conformational ensemble. This issue is best illustrated by the variability of the pPII population, which is typically underestimated, when compared to the Gaussian model predictions, for the guest residues with the highest pPII content and overestimated for the residues with the lowest pPII content. Although overall associated with the largest χ^2_J and χ^2_{VCD} values, Amber ff19SB exhibits the highest level of guest-residue specificity in the Ramachandran distributions. This is reflected in positive correlations between Amber ff19SB-derived and Gaussian model pPII populations, which are much smaller in CHARMM36m, even smaller in Amber ff14SB, whereas OPLS-AA/M displays negative correlations. While combined β-strand populations produced by MD force fields display some correlations with Gaussian model-derived counterparts, they do not provide much insight as the MD force fields are overly biased toward the aβ rather than βt populations. Interestingly, CHARMM36m displays some guest-residue specificity and does not produce extreme χ^2 values. While Amber ff14SB and OPLS-AA/M do not show much amino acid specificity, OPLS-AA/M overall performs relatively well except for glycine and alanine residues. With respect to the two χ^2 values used in our assessment, the strengths of OPLS-AA/M include a lesser bias toward aβ (samples more βt conformations in most cases) and less sampling of the right-handed helix region, consistent with the Gaussian model predictions. Notably, the pPII basins predicted by OPLS-AA/M are shifted to the negative ϕ values relative to the Gaussian model predictions, which may be related to the negative correlations between the respective pPII populations (in particular for aliphatic residues in Fig. 2).

Amber ff19SB and CHARMM36m use CMAP energy landscapes, rather than uncoupled 1-D functionals used in the other two force fields, to optimize backbone dihedral potentials and derive parameters for more amino acid residues than the other two force fields, which allows for more amino acid residue-specific features to emerge in Ramachandran distributions. This may explain why Amber ff19SB better reproduces the residue-specific pPII content in the Ramachandran distributions of the investigated guest residues. However, the poor reproduction of experimental values by Amber ff19SB demonstrates that an increased number of parameters alone, while important and necessary, is not sufficient for reproduction of the experimental data. In this context, it is noteworthy to consider several lines of evidence indicating that the hydration of backbone and side chain is the major determinant of the conformational preferences of amino acid residues in the absence of significant nearest-neighbor interactions[22,26,28]. Amber ff19SB utilized a continuum implicit water model in the QM energy calculations, which were used to parameterize the backbone dihedral and side chain torsional potentials[18]. However, recent DFT calculations clearly demonstrate that using explicit water, which can form hydrogen bonds with the peptide's functional groups, are necessary to capture the pPII/β-strand equilibria and produce the experimentally observed pPII content of alanine residue[63,64]. The biggest question arising from these observations is why the explicit water model within the MD force field (which is based on potentials derived from implicit-water QM calculations) is unable to account for experimentally-observed residue-specific conformational ensembles.

Improvements of the residue-specific backbone CMAP potentials (in combination with side-chain torsional potentials) can potentially eliminate the issues of residue specificity and result in the balance between the βt and aβ conformations that better reproduces the spectroscopic data. The findings of this study pose a challenge in the development of the MD force fields because DFT calculations with explicit water are computationally demanding. Moreover, it is not clear whether the force field improvement can be achieved solely by implementing residue specificity into di-

hedral potentials. Indeed, we tested residue-specific force field RSFF1 on a subset of 6 guest residues x (= A, I, F, D^P, R, S) in GxG peptides and found that including residue specificity derived from protein coil libraries does not improve OPLS-AA/L beyond its latest version OPLS-AA/M. It could be argued that a better water model would account for improved peptide hydration and water-peptide hydrogen bonding, thereby mediating interactions between the side-chain and backbone groups which contribute to residue specificity. Indeed, replacing the simpler water model TIP3P by TIP4P/2005 in Amber ff14SB resulted in an overall reduction of χ^2 values for the six guest residues x (= A, I, F, D^P, R, S). However, the improvement in the χ^2 values is only modest and does not bring Amber ff14SB combined with TIP4P/2005 significantly closer to the Gaussian model predictions, indicating that multiple aspects of the MD force field potential energy function require modifications. This can be illustrated by the example of the protonated aspartic acid residue, whose conformational ensemble is particularly poorly captured by all MD force fields used in this study. As reported by Hagarman *et al.* and Rybka *et al.*, turn-forming conformations are prominent in conformational ensembles of D^P, S, and T, which exhibit hydrogen bonding capacity that has been associated with a relatively high turn content[31,32]. These observations combined suggest that the current MD force fields do not properly account for the hydrogen bonding capacity of the side chains.

In summary, our study elucidates the inability of the current MD force fields to capture the intrinsic conformational ensembles of guest amino acids x in GxG in water. This represents a major weakness of the MD methodology for a number of reasons. First, correctly reproduced Ramachandran distributions of amino acid residues in short unfolded peptides are required to accurately model disordered regions of IDPs in the absence of non-local interactions. This is particularly important when determining the entropic and enthalpic contributions to ligand-induced folding or binding to IDPs. Second, accurate MD predictions of global properties of IDPs, such as a radius of gyration and its scaling with the number of amino acids, require that turn-forming preferences of amino acid residues are adequately accounted for. Third, the errors in Ramachandran distributions of guest residues x in GxG peptides reported in this work are expected to accumulate in MD simulations of longer unfolded peptides, negatively affecting MD simulations of IDPs and resulting predictions, impeding the use of MD as a tool in unraveling the mysteries of IDPs.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The authors thank Carlos Simmerling for his help with Amber ff19SB simulations and insightful comments as well as Fan Jiang and Yun-Dong Wu for sharing the GROMACS files needed for RSFF1 implementation. This work has been supported by the National Science Foundation through the grant numbers MCB-1817650 (B.A., R.S.S., B.U.) and OAC-1919691 (B.U.). MD simulations were performed on hardware supported by Drexel's University Research Computing Facility.

Notes and references

1 P. E. Wright and H. J. Dyson, *Nat. Rev. Mol. Cell Biol.*, 2015, **16**, 18–29.
2 E. Nwanochie and V. N. Uversky, *Int. J. Mol. Sci.*, 2019, **20**, 4186.
3 S. Piana, A. G. Donchev, P. Robustelli and D. E. Shaw, *J. Phys. Chem. B*, 2015, **119**, 5113–5123.
4 J. Henriques, C. Cragnell and M. Skepo, *J. Chem. Theory Comput.*, 2015, **11**, 3420–3431.
5 R. B. Best, W. Zheng and J. Mittal, *J. Chem. Theory Comput.*, 2014, **10**, 5113–5124.
6 J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. de Groot, H. Grubmüller and A. D. MacKerell, Jr., *Nature Methods*, 2017, **14**, 71–73.
7 P. S. Shabane, S. Izadi and A. V. Onufriev, *J. Chem. Theory Comput.*, 2019, **15**, 2620–2634.
8 D. Petrov and B. Zagrovic, *PLoS Comput. Biol.*, 2014, **10**, e1003638.
9 M. Carballo-Pacheco, A. E. Ismail and B. Strödel, *J. Chem. Theory Comput.*, 2018, **14**, 6063–6075.
10 B. Andrews, K. Long and B. Urbanc, *J. Phys. Chem. B*, 2021, **125**, 6897–6911.
11 R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig and A. D. MacKerell, Jr., *J. Chem. Theory Comput.*, 2012, **8**, 3257–3273.
12 M. J. Robertson, J. Tirado-Rives and W. L. Jorgensen, *J. Chem. Theory Comput.*, 2015, **11**, 3499–3509.
13 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, 2015, **11**, 3696–3713.
14 J. Huang and A. D. MacKerell, Jr., *J. Comput. Chem.*, 2013, **34**, 2135–2145.
15 L. Wickstrom, A. Okur and C. Simmerling, *Biophys. J.*, 2009, **97**, 853–856.
16 P. S. Georgoulia and N. M. Glykos, *J. Phys. Chem. B*, 2011, **115**, 15221–15227.
17 A. T. Bogetti, H. E. Piston, J. M. G. Leung, C. C. Cabalteja, D. T. Yang, A. J. DeGrave, K. T. Debiec, D. S. Cerutti, D. A. Case, W. S. Horne and L. T. Chong, *J. Chem. Phys.*, 2020, **153**, year.
18 C. Tian, K. Kasavajhala, K. A. A. Belfon, L. Raguette, H. Huang, A. N. Migues, J. Bickel, Y. Wang, J. Pincay, Q. Wu and C. Simmerling, *J. Chem. Theory Comput.*, 2020, **16**, 528–552.
19 S. E. Toal and R. Schweitzer-Stenner, *Biomolecules*, 2014, **4**, 725–773.
20 W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225–11236.
21 G. A. Kaminski, R. A. Friesner, J. Tirado-Rives and W. L. Jorgensen, *J. Phys. Chem. B*, 2001, **105**, 6474–6487.
22 D. Meral, S. Toal, R. Schweitzer-Stenner and B. Urbanc, *J. Phys. Chem. B*, 2015, **119**, 13237–13251.

23 S. E. Toal, D. J. Verbaro and R. Schweitzer-Stenner, *J. Phys. Chem. B*, 2014, **118**, 1309–1318.

24 B. Milorey, H. Schwalbe, N. O'Neill and R. Schweitzer-Stenner, *J. Phys. Chem. B*, 2021, **125**, 11392–11407.

25 R. Schweitzer-Stenner, *Mol. BioSyst.*, 2012, **8**, 122–133.

26 S. Zhang, R. Schweitzer-Stenner and B. Urbanc, *J. Chem. Theory Comput.*, 2020, **16**, 510–527.

27 R. Schweitzer-Stenner, *J. Phys. Chem. B*, 2009, **113**, 2922–2932.

28 B. Andrews, S. Zhang, R. Schweitzer-Stenner and B. Urbanc, *Biomolecules*, 2020, **10**, 1121.

29 S. Zhang, B. Andrews, R. Schweitzer-Stenner and B. Urbanc, *J. Phys. Chem. B*, 2020, **124**, 11600–11616.

30 A. Hagarman, T. J. Measey, D. Mathieu, H. Schwalbe and R. Schweitzer-Stenner, *J. Am. Chem. Soc.*, 2010, **132**, 540–551.

31 A. Hagarman, D. Mathieu, S. Toal, T. J. Measey, H. Schwalbe and R. Schweitzer-Stenner, *Chem. Eur. J.*, 2011, **17**, 6789–6797.

32 K. Rybka, S. E. Toal, D. J. Verbaro, D. Mathieu, H. Schwalbe and R. Schweitzer-Stenner, *Proteins: Struct., Funct., Bioinf.*, 2013, **81**, 968–983.

33 R. Schweitzer-Stenner, A. Hagarman, S. Toal, D. Mathieu and H. Schwalbe, *Proteins: Struct., Funct., Bioinf.*, 2013, **81**, 955–967.

34 D. DiGuiseppi, J. Kraus, S. E. Toal, N. Alvarez and R. Schweitzer-Stenner, *J. Phys. Chem. B*, 2016, **120**, 10079–10090.

35 S. Izadi, R. Anandakrishnan and A. V. Onufriev, *J. Phys. Chem. Lett.*, 2014, **5**, 3863–3871.

36 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33–38.

37 H. J. C. Berendsen, D. van der Spoel and R. van Drunen, *Comput. Phys. Commun.*, 1995, **91**, 43–56.

38 E. Lindahl, B. Hess and D. van der Spoel, *J. Mol. Model.*, 2001, **7**, 306–317.

39 D. V. D. Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berendsen, *J. Comput. Chem.*, 2005, **26**, 1701–1718.

40 B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *J. Chem. Theory Comput.*, 2008, **4**, 435–447.

41 S. Pronk, S. Pall, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess and E. Lindahl, *Bioinformatics*, 2013, **29**, 845–854.

42 S. Pall, M. J. Abraham, C. Kutzner, B. Hess and E. Lindahl, Solving software challenges for exascale, 2015, pp. 3–27.

43 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1-2**, 19 – 25.

44 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.

45 A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, **102**, 3586–3616.

46 A. D. MacKerell, M. Feig and C. L. Brooks, *J. Comput. Chem.*, 2004, **25**, 1400–1415.

47 J. Abascal and C. Vega, *J. Chem. Phys.*, 2005, **123**, year.

48 F. Jiang, C.-Y. Zhou and Y.-D. Wu, *J. Phys. Chem. B*, 2014, **118**, 6983–6998.

49 H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura and T. Head-Gordon, *J. Chem. Phys.*, 2004, **120**, 9665–9678.

50 S. Pall and B. Hess, *Comput. Phys. Commun.*, 2013, **184**, 2641–2650.

51 G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**, 014101.

52 H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **8**, 3684–3690.

53 R. Schweitzer-Stenner, F. Eker, Q. Huang and K. Griebenow, *J. Am. Chem. Soc.*, 2001, **123**, 9628–9633.

54 A. C. Wang and A. Bax, *J. Am. Chem. Soc.*, 1996, **118**, 2483–2494.

55 J.-S. Hu and A. Bax, *J. Am. Chem. Soc.*, 1997, **119**, 6360–6368.

56 J. Wirmer and H. Schwalbe, *J. Biomol. NMR*, 2002, **23**, 47–55.

57 K. Y. Ding and A. M. Gronenborn, *J. Am. Chem. Soc.*, 2004, **126**, 6232–6233.

58 B. Milorey, R. Schweitzer-Stenner, B. Andrews, H. Schwalbe and B. Urbanc, *Biophys. J.*, 2021, **120**, 662–676.

59 J. Graf, P. H. Nguyen, G. Stock and H. Schwalbe, *J. Am. Chem. Soc.*, 2007, **129**, 1179–1189.

60 D. DiGuiseppi and R. Schweitzer-Stenner, *J. Raman Spectrosc.*, 2016, **47**, 1063–1072.

61 L. A. Nafie, M. R. Oboodi and T. B. Freedman, *J. Am. Chem. Soc.*, 1983, **105**, 7449–7450.

62 R. Schweitzer-Stenner and S. E. Toal, *Mol. BioSyst.*, 2016, **12**, 3294–3306.

63 N. V. Ilawe, A. E. Raeber, R. Schweitzer-Stenner, S. E. Toal and B. M. Wong, *Phys. Chem. Chem. Phys.*, 2015, **17**, 24917–24924.

64 G. Lanza and M. A. Chiacchio, *J. Phys. Chem. B*, 2016, **120**, 11705–11719.