Title
The combined force field-sampling problem in simulations of disordered amyloid-β peptides

Journal
The Journal of Chemical Physics, 150(10)

ISSN
0021-9606

Authors
Lincoff, James
Sasmal, Sukanya
Head-Gordon, Teresa

Publication Date
2019-03-14

DOI
10.1063/1.5078615

Peer reviewed

The Journal of Chemical Physics ARTICLE scitation.org/journal/jcp

The combined force field-sampling problem in simulations of disordered amyloid-β peptides

Cite as: J. Chem. Phys. 150, 104108 (2019); doi: 10.1063/1.5078615
Submitted: 26 October 2018 • Accepted: 20 February 2019 •
Published Online: 14 March 2019

James Lincoff,[1,a] Sukanya Sasmal,[1,a),b] and Teresa Head-Gordon[1,2,3,4,c]

AFFILIATIONS

[1] Department of Chemical and Biomolecular Engineering, University of California, Berkeley, California 94720, USA
[2] Department of Chemistry, University of California, Berkeley, California 94720, USA
[3] Department of Bioengineering, University of California, Berkeley, California 94720, USA
[4] Pitzer Theory Center, University of California, Berkeley, California 94720, USA

[a)] **Contributions:** J. Lincoff and S. Sasmal contributed equally to this work.
[b)] **Current address:** Department of Pharmaceutical Sciences, University of California, Irvine, California 92697, USA.
[c)] **Author to whom correspondence should be addressed:** thg@berkeley.edu

ABSTRACT

Molecular dynamics simulations of intrinsically disordered proteins (IDPs) can provide high resolution structural ensembles if the force field is accurate enough and if the simulation sufficiently samples the conformational space of the IDP with the correct weighting of sub-populations. Here, we investigate the combined force field–sampling problem by testing a standard force field as well as newer fixed charge force fields, the latter specifically motivated for better description of unfolded states and IDPs, and comparing them with a standard temperature replica exchange (TREx) protocol and a non-equilibrium Temperature Cool Walking (TCW) sampling algorithm. The force field and sampling combinations are used to characterize the structural ensembles of the amyloid-beta peptides Aβ42 and Aβ43, which both should be random coils as shown recently by experimental nuclear magnetic resonance (NMR) and 2D Förster resonance energy transfer (FRET) experiments. The results illustrate the key importance of the sampling algorithm: while the standard force field using TREx is in poor agreement with the NMR J-coupling and nuclear Overhauser effect and 2D FRET data, when using the TCW method, the standard and optimized protein-water force field combinations are in very good agreement with the same experimental data since the TCW sampling method produces qualitatively different ensembles than TREx. We also discuss the relative merit of the 2D FRET data when validating structural ensembles using the different force fields and sampling protocols investigated in this work for small IDPs such as the Aβ42 and Aβ43 peptides.

INTRODUCTION

Intrinsically disordered proteins (IDPs) are a class of biomolecules that do not adopt a well-defined equilibrium structure in solution, instead sampling an ensemble comprised of sub-populations of fully and/or partially disordered structures.[1,2] A classic IDP example is the amyloid-β (Aβ) peptide associated with Alzheimer's disease,[3] for which recent state-of-the-art solution-based nuclear magnetic resonance (NMR) and Förster resonance energy transfer (FRET) experiments have shown that the monomeric forms of the Aβ40 and Aβ42 peptides are largely random coils.[4–6] Roche and co-workers performed multiple types of J-coupling measurements on Aβ40 and Aβ42 to show there were no overt differences from random coil signatures for both peptides.[5] They further supported this result with very high resolution Nuclear Overhauser effect (NOE) spectra that showed that both Aβ40 and Aβ42 are dominated by short (i, i + 1 contacts) and to a lesser extent medium range (i, i + 2 to i, i + 4 contacts) NOEs, and thus, any longer-range helical or β–hairpin formation would only be expected to appear at levels below the detection limit.[5] Conicella and Fawzi employed $^{1}H_N$ and ^{15}N chemical shift data and H_N-H_α J-couplings, as well as ^{15}N R_2, ^{15}N R_1, and heteronuclear ^{15}N–^{1}H NOE measurements, to show that there are no structural differences in the N-terminus and central hydrophobic core (CHC) between the

Aβ42 and Aβ43 peptides, with only small structural differences in the C-terminus.[4] Since the two peptides differ only by a threonine residue at the C-terminus, it implies that the Aβ43 monomer is also largely random coil-like in structure, albeit with a greater propensity for aggregation.[4] Finally, using a single molecule and 2D FRET, Meng *et al.*[6] showed that both Aβ40 and Aβ42 have ensembles that are dominated by expanded conformations with no persistent sub-populations of secondary or tertiary structure (i.e., no long helices or β sheets above ~5%-10% population), in general agreement with the NOE spectra of Roche *et al.*

Computational techniques are often combined with such experimental information to create the structural ensemble and to characterize the sub-populations of an IDP of interest.[7–10] A complementary approach is to generate IDP ensembles using molecular dynamics (MD) simulation without experimental information as input, which therefore requires an accurate force field and a sampling method that can describe the conformational substates of the IDP ensemble. In early simulation studies of IDPs, research groups relied on off-the-shelf and pairwise-additive protein and water force fields,[11] such as the Amber,[12] GROMOS,[13] OPLS-AA,[14,15] and CHARMM[16] protein force fields in combination with TIP3P,[17] TIP4P,[18] and TIP4P-Ew[19] water models. But since standard pairwise additive force fields are parameterized using mostly folded protein crystallographic data,[20] they have been thought to be insufficient for modeling of IDPs because they exhibit a bias toward overly collapsed and ordered structural ensembles or poorly reproduce the equilibrium between unfolded and native conditions for globular proteins.[21–23]

Furthermore, multiple groups have shown that unfolded and IDP structural ensembles generated using different standard force fields vary considerably in terms of secondary structure content.[24–27] For example, in studies of the $A\beta_{16-22}$ peptide, Nguyen *et al.* have demonstrated that Amber99 predicts more helical structures and GROMOS96[28] favors more β-strand structures, whereas OPLS-AA demonstrates no particular secondary structure preference.[24] As we have discussed before, the quality of the water model is also critical for accurate molecular simulations of peptides and proteins by balancing the relative strengths of water-water and water-solute interactions.[29] Several groups including Song and co-workers,[30] Amini and co-workers,[31] and Viet and co-workers[32] have conducted straight MD simulations on Aβ peptides on the order of microseconds with standard protein force fields and using a variety of three-site and four-site water models and multiple starting structures. Using brute force MD, Robustelli and co-workers simulated Aβ40 using a single 30 μs trajectory of a standard protein force field Amber99ffsb-ildn* and TIP3P and found close to 90% β−sheet structure in different regions of the sequence.[33] In general, all of these simulations produce overstructured ensembles, in disagreement with the experimental results of Roche *et al.*[5] and Meng *et al.*[6] for Aβ peptides. One might surmise from the accumulation of evidence that the standard force fields for the protein and water are failing to describe disordered or unfolded protein ensembles.

To make better and more uniform predictions across different IDPs, as well as more accurate models for protein folding equilibrium, a number of research labs have modified the parameters used in standard force fields.[21–23,27,34,35] These force field modifications include adjusting the water-protein London dispersion interactions to be more favorable,[21,22] refining the peptide backbone parameters to produce more expanded structures[22,23,34] or reduce propensity to certain ordered conformations,[21] and/or changing the salt-bridge interactions.[21] Skepö and co-workers, using long MD simulations, have shown that some of these modified force fields produce better agreement with experimental small-angle x-ray scattering (SAXS) data in terms of radius of gyration, R_g, for the disordered histatin 5 peptide.[36,37] Huang *et al.*[21] also found better agreement with the SAXS profile for the RS peptide using the modified CHARMM force field. However, Robustelli *et al.* simulated Aβ40 with a single 30 μs trajectory for the recently modified A03ws force field by Best and co-workers[38] and found close to 25%-60% α−helix in large regions of the sequence that indicate that the newer IDP force fields may also be experiencing problems.[33]

But the other important and often not enough emphasized aspect of generating IDP structural ensembles is the sampling technique itself. Because IDPs have a relatively flat energy landscape with many local minima, it takes substantial sampling efficiency to determine all relevant configurations with the correct weighting of multiple small sub-populations. Enhanced sampling methods are therefore generally applied for the simulation of IDPs, as they accelerate the rate of convergence to time scales that are significantly less than possible with a brute force MD simulation.[39,40] The most common enhanced sampling technique used in the IDP field at present is the temperature replica exchange (TREx) method.[39–43] In fact, TREx simulations were used to identify perceived errors in standard force fields, which led to some of the modified force fields developed to improve modeling of unfolded proteins and IDPs.[26,27,38,44]

However, a noted deficiency of TREx for large systems is the diffusiveness of barrier crossing due to many closely spaced intermediate replicas when energy landscapes are dominated by entropic barriers.[39,45,46] Several alternative enhanced sampling methods with better performance than TREx have been developed and applied to the study of IDPs, including replica exchange with solute tempering (REST),[47] metadynamics,[48] and MD combined with the Markov State Model (MSM) analysis.[49] More specifically, Lin *et al.*[49] performed ~200 μs MD simulations over many initial conditions, and combined it with MSM to reach even longer time scales, to characterize the Aβ42 structural ensemble using the Amberff99SB protein force field and the TIP3P water model. Over this much longer sampling time scale, they obtained far higher quantities of extended, largely unstructured conformations, with the only noteworthy structural component being ~10%–20% observed helical content between residues 12-18. While there may be concerns that the clustering protocols may have introduced error in secondary structure populations, and improved clustering methods are now available,[88] it is evident that the full weighted ensemble based on the MSM is equivalent to the raw α-helix and β propensities from their production MD runs (as seen in their supplementary material). Thus, this more extensive sampling produced a significant improvement in generating random coil ensembles using a standard peptide-water force field, in good agreement with the experimental results of Roche *et al.*[5] and Meng *et al.*[6] for Aβ peptides and in significant disagreement with previous shorter MD runs[33] and TREx studies.[26,27,38,44]

We have developed the temperature cool walking (TCW) technique,[46,50] a non-equilibrium alternative to TREx, which uses only one high temperature replica to generate trial moves for the target

temperature replica. In previous studies, we have shown that TCW converges more quickly to the proper equilibrium distribution than TREx, and at much lower computational expense, for a 1D rough surface[46] and for alanine dipeptide and met-enkephalin—sufficiently small systems where well-defined and quantitative metrics of convergence are available.[50] More recently, we have been able to apply TCW to larger systems such as Aβ through its implementation in OpenMM.[50,51] In this work, we address both dimensions of the IDP problem by comparing different combinations of enhanced sampling techniques, TREx and TCW, and protein-water force fields, unmodified and newly optimized, testing combinations of each on Aβ42 and Aβ43. Again summarizing the recent experiments,[4–6] the structural ensembles of Aβ42 and Aβ43 peptides should be largely the same and exhibit no persistent structural ordering or long-range contacts. Thus, one would expect the computationally generated structural ensembles to be highly similar for Aβ42 and Aβ43 in terms of back-calculations to experimental observables such as chemical shifts or J-couplings[4,5] and that both peptides would display largely random coil configurations, lacking stable populations of organized structures such as β−hairpins and extended helices, to agree with reported NOE[4,5] and 2D FRET[6] data for Aβ.

Among our set of results, we attain the biggest improvement in IDP ensemble generation by switching the sampling method from TREx to TCW, since the latter sampling algorithm provides much better agreement with the full range of NMR J-coupling and NOE data[4,5] as well as with the 2D FRET data,[4–6] compared to the same force fields simulated with TREx. More specifically, the unmodified force field when sampled with TCW yields Aβ ensembles that are largely unstructured, in qualitative agreement with the robust MD/MSM simulation[49] of the Aβ42 peptide using a similar unmodified force field, with only small populations of structures containing longer-range contacts at levels (~5%-10%) that would be undetectable by the NOE and 2D FRET experiments. We believe that the presented evidence supports the conclusion that TCW is more capable than TREx of sampling the disordered protein energy landscape and is in support of recent work by Granata and co-workers that have shown that disordered conformations are lower in free energy than ordered structures.[52] Finally, using the TCW sampling protocol, we find that the newly modified force fields do produce more extended ensembles that are in better quantitative agreement with the 2D FRET data, whereas the standard force field is in better quantitative agreement with the NMR J-coupling data. We conclude that more work is needed in regard to interpreting FRET data[6,53–57] and more extensive testing in general is needed before standard force fields are abandoned or more force field changes are pursued for IDPs.

METHODS

Peptide simulations

The starting Aβ42 and Aβ43 configurations were created using the *tleap* module in Amber,[58] and the peptides were subsequently minimized and equilibrated in the NPT ensemble at 1 bar to obtain the correct density. Structures at the density of maximum probability were selected as initial structures for production, producing boxes that were approximately 60 Å on each side. The Ewald summation was used for the long-range with a cutoff of 9.0-9.5 Å for the real space electrostatics and Lennard-Jones forces. Trajectories were analyzed for results at 287 K.

TREx simulation protocol

The Amber14 molecular dynamics package[58] was used to perform 100 ns TREx simulations with 58 temperature replicas in the temperature range 287-450 K, testing both peptides with (1) the Amber ff99SB[12] force field with TIP4P-Ew[19] water and (2) the CHARMM36m[21] force field with CHARMM-TIP3P water.[16] The temperature schedule was chosen such that the exchange probability between each pair of replicas was in the range 18%-22%, which has been shown to be optimal for TREx.[59] Exchanges between neighboring replicas were attempted every 0.5 ps. The TREx simulations were performed in the NVT ensemble with a time step of 1 fs and with constraints on the heavy atom hydrogen bonds. A Langevin thermostat was used to maintain a constant temperature. Each set of conditions—peptide + force field + water model—was run in duplicate. The first 50 ns of data was discarded as equilibration, with the last 50 ns being analyzed and presented, averaged across the two independent simulations. TREx simulations with this setup were also attempted for the Amber ff99SB-ILDN[22,60] force field with TIP4P-D[22] water, but using the same temperature ladder and simulation package, we were unable to obtain similar exchange probabilities across the entire temperature span. As we could not run an optimal TREx simulation with this force field combination, no data are presented for this simulation setup.

Over the last several years, several groups have run much longer TREx simulations for IDPs, on the order of 750–1000 ns, noting that even when optimally run, structural properties such as the radius of gyration and secondary structure propensities can take several hundred ns to reach apparent equilibration in TREx.[6,44,61,62] To evaluate this difference, we additionally performed one 800 ns TREx simulation each for Aβ42 and Aβ43 using Amber ff99SB + TIP4P-Ew water using the Amber16 molecular dynamics package.[63] All parameters were kept the same as in the original 100 ns TREx simulations with the exception of using a 2 fs time step and attempting exchanges between adjacent pairs of replicas every 1 ps. For these simulations, the first 300 ns of data was discarded as equilibration, with the last 500 ns of data at 287 K analyzed. Averaging over two 250 ns blocks was done to mimic the averaging over two independent simulations performed for the 100 ns TREx simulations and the TCW simulations.

TCW protocol

The TCW enhanced sampling method uses only two temperature replicas to generate an equilibrium ensemble at the target temperature.[46,50] Starting with expanded structures at the high temperature, sequential cooling is performed to obtain structures at the low temperature such that detailed balance is satisfied. The same set of maximum and minimum temperatures and cooling schedule was used as for all TREx simulations. The temperature was regulated using an Andersen thermostat. Trial moves were initiated every 8 ps. The cooling rate was set such that the peptides spent 40 fs on average at each intermediate temperature, with the first trial exchange per cooling run occurring after having annealed

through 40 of the 56 intermediate temperature steps on average. Additional high temperature replica propagation was performed to further decorrelate subsequent configurations at a ratio of 8:1 to each fs of annealing performed. TCW simulations were performed using modifications to the OpenMM software package,[64] and a code is available upon request from the authors. TCW simulations were performed with the (1) Amber ff99SB force field + TIP4P-Ew water model, (2) CHARMM36m force field + CHARMM-TIP3P water, and (3) Amber ff99SB-ILDN force field + TIP4P-D water. Simulations were run in duplicate for 200 ns, with the first 50 ns of each discarded as equilibration and results from the remaining 150 ns of each simulation averaged together.

Trajectory analysis

The structural ensembles were analyzed using both the *cpptraj*[65] module of Amber and in-house codes. Contact maps were generated by calculating the fraction of structures where pairs of residues had at least one pair of heavy atoms within 5 Å of each other. The DSSP criterion was used to assign secondary structures.[66] Details about the back calculation of NMR observables have been reported in previous publications by our group,[10,29,67,68] including chemical shifts from ShiftX2[69] and J-coupling constants.[70] In this work, we focus primarily on J-couplings using the Karplus equation

$$\langle J \rangle = \langle A\cos^2\phi + B\cos\phi + C \rangle, \tag{1}$$

where $\langle \ldots \rangle$ denote ensemble averages. We calculate the χ^2 parameter from the simulated J-coupling constants ($J_{HN\text{-}H\alpha}$) for each ensemble as compared to the experiment,

$$\chi^2 = \frac{1}{N}\sum_{i=1}^{N}\frac{(\langle J_i \rangle_{sim} - J_{i,expt})^2}{\sigma^2}, \tag{2}$$

where J_i is the scalar coupling constant for the ith residue, N is the total number of experimental $J_{HN\text{-}H\alpha}$ observables, subscripts *sim* and *expt* refer to the simulated and the experimental values, respectively, and σ^2 is the RMSD error when using the Karplus parameters introduced by Vögeli *et al.*[71] We additionally perform a Bayesian analysis on the scalar couplings, Experimental Inferential Structure Determination (EISD),[72] that accounts for uncertainties in the values of the Karplus parameters as well as the individual per-coupling experimental uncertainties. The relative magnitudes of ensemble scores for a peptide represent the relative likelihood of that structural

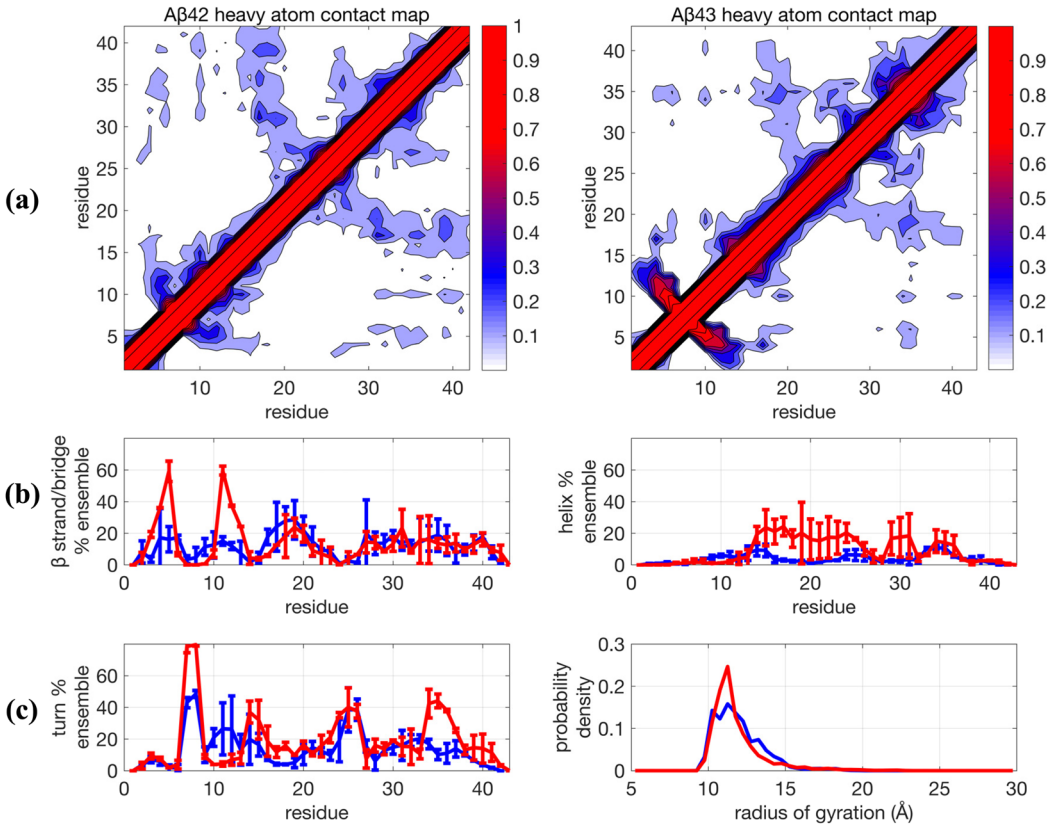

FIG. 1. Comparison of structural properties for Aβ42 and Aβ43 using the force field Amber99SB + TIP4P-Ew and using TREx simulated at 0.1 μs/replica. (a) Contact maps, (b) β−strand and helix propensities, and (c) turn propensity and radius of gyration distribution. For (b) and (c), blue lines are for Aβ42 and red lines are Aβ43 and represent the average for two independent trajectories. Error bars are plus and minus one standard deviation of the calculated propensity at the given residue for the two trajectories and generally represent the degree of agreement between the two trajectories.

ensemble matching the experimental data against which they are compared with a larger score corresponding to a higher probability.

The observable directly obtained in MD simulations of the FRET experiments, the efficiency distribution $E(t)$,

$$E_{FRET} = \frac{1}{1 + (R_{ee}/R_0)^6}, \quad (3)$$

and its average E_{FRET}, is back-calculated from the simulated end-to-end distance, R_{ee} of the untagged peptide across the ensemble, with a Förster radius $R_0 = 5.2$ nm for the dye pair of Alexa 488 and 647.[6] As in the work of Meng *et al.*, the R_{ee} is first calculated as the distance between the C_α atoms of the first and last residue of each peptide and then scaled up to approximate the additional distance between the two fluorophores, relative to the distance between the first and last residues.

RESULTS

We tested five different force field-sampling combinations for each of the two peptides: (1) duplicate TREx simulations using the standard Amber ff99SB[12] protein force field and the TIP4P-Ew[19] water model, (2) duplicate TREx simulations using the modified CHARMM36m[21] protein force field and the CHARMM-TIP3P[16] water model, (3) duplicate TCW simulations using the standard Amber ff99SB protein force field and the TIP4P-Ew water model (4) duplicate TCW simulations using the modified CHARMM36m protein force field and the CHARMM-TIP3P water model, and (5) duplicate TCW simulations using the Amber ff99SB-ILDN[60] protein force field and the TIP4P-D[22] water model.

Figure 1 shows that the 0.1 μs TREx + Amber ff99SB + TIP4P-Ew simulations predict that both peptides are very collapsed, as evidenced by their contact maps and R_g distributions, which originate from the abundance of organized backbone structure that underlies the secondary structure propensities (Fig. 1 and Table I). In addition, the Aβ42 and Aβ43 ensembles show significant differences, with the Aβ43 ensemble exhibiting an increased α−helical structure in the central hydrophobic cluster (CHC) of residues 16-30, as well as increases in the turn populations at residues 36-38 and the 6-9 region with a simultaneous increase in the β strand formed by residues 3-5 and 10-12, structural sub-populations that are greatly diminished (but still present) in the Aβ42 ensemble. This result is in direct contradiction with results reported by Conicella

TABLE I. Simulated properties for the Aβ42 and Aβ43 peptides for each sampling method-force field combination. Mean and standard deviation averaged over two independent trajectories, except for the 0.8 μs TREx simulation, where block averaging over two 250 ns blocks was used. We noted a systematic shift in J-couplings between the two experimental datasets from Conicella and Fawzi[4] and Roche *et al.*; we have applied a −0.4 Hz shift to the Aβ43 J_i values from Conicella and Fawzi before comparison which would bring the Aβ42 results in line with each other to the simulated J-couplings.

	Sampling method and force field combination					
Peptide	TREx (0.1 μs) + TIP4P-Ew + Amberff99SB	TREx (0.1 μs) + CHARMM36m + CHARMM-TIP3P	TREx (0.8 μs) + Amberff99SB + TIP4P-Ew	TCW (0.2 μs) + Amberff99SB + TIP4P-Ew	TCW (0.2 μs) + CHARMM36m + CHARMM-TIP3P	TCW (0.2 μs) + Amber99SB-ILDN + TIP4P-D
	χ^2 between simulated and experimental[4,5] J-coupling constants					
Aβ42	3.70	4.65	3.70	2.70	3.01	2.89
Aβ43	4.75	3.76	3.65	2.47	2.96	2.71
	EISD score					
Aβ42	39.479	3.858	40.014	57.833	47.741	54.937
Aβ43	23.234	25.090	36.530	47.488	39.328	45.338
	Mean and standard deviation of the end-to-end-distance, $\langle R_{ee} \rangle$ (in Å)					
Aβ42	24.3 ± 0.6	39.3 ± 3.0	20.5 ± 1.7	28.4 ± 0.9	36.8 ± 1.3	33.5 ± 7.9
Aβ43	26.5 ± 3.5	44.4 ± 1.1	20.3 ± 0.7	29.1 ± 2.0	38.2 ± 1.3	31.9 ± 0.9
	Mean and standard deviation of FRET efficiencies, $\langle E_{FRET} \rangle$					
Aβ42	0.93 ± 0.002	0.64 ± 0.084	0.96 ± 0.007	0.88 ± 0.009	0.71 ± 0.031	0.77 ± 0.150
Aβ43	0.92 ± 0.049	0.56 ± 0.001	0.97 ± 0.007	0.87 ± 0.034	0.69 ± 0.020	0.82 ± 0.004
	Mean and standard deviation of the radius of gyration, $\langle R_g \rangle$ (in Å)					
Aβ42	12.0 ± 0.3	17.2 ± 2.2	11.3 ± 0.0	12.9 ± 0.1	15.9 ± 0.5	14.9 ± 1.6
Aβ43	11.8 ± 0.6	17.8 ± 0.4	11.8 ± 0.1	13.2 ± 0.4	16.4 ± 0.2	15.7 ± 2.0

and Fawzi,[4] which showed that there are no major structural differences in the N-terminus and CHC region for the two peptide monomers.

When using the same TREx sampling method for 0.1 μs, but changing to the CHARMM36m + CHARMM-TIP3P force field, there is significant reduction in long-range structure and secondary structure propensities for the two peptides, and hence, the ensembles are less collapsed compared to the standard force field combination for the two peptides (Table I and Fig. S1). However, the TREx simulations using the newer protein-water force field produce ensembles that still predict differences in the sub-populations of structure for both Aβ42 and Aβ43 that are not observed in the experiment,[4] such as the relative enrichment in β content for Aβ42. The long-range structure evident in the contact maps shown in Fig. S1 directly contradicts the NOE data of Roche *et al.*, which found no evidence for long-range contacts.[5] The large error bars indicate that the two independent trajectories did not converge to the same result, a sign that one or both of the independent trajectories were stuck in local minima on the 0.1 μs/replica time scale of the TREx simulation. We show this data to emphasize the point that TREx simulations on these non-converged time scales were used to identify errors in standard force fields and thus informed the development of modified force fields for unfolded proteins and IDPs.[27,38]

Since more recently it has become standard to perform TREx simulations using on the order of 1 μs/replica in recent IDP studies,[6,21,44,61,62] we conducted one additional trajectory using a TREx simulation out to 800 ns/replica for each peptide using Amberff99SB + TIP4P-Ew, with the resulting structural data in Fig. 2. Our TREx simulation of 0.8 μs is generally consistent with the 1.0 μs/replica TREx simulations performed by Rosenman *et al.*[44] With longer simulation time scales, the previously significant differences between Aβ42 and Aβ43 are reduced to being within statistical error for all secondary structure categories. While there is a clear improvement in some structural properties at this longer time scale using TREx, such as reduction in the population of α−helices, the β−sheet propensities still show a large variability of ±20% at residues 5-6, 19, and 31. As a result, their structural ensembles are in stark disagreement with the available NMR and 2D FRET data by being too collapsed and highly structured. It is therefore very understandable why one would continue to conclude that there is a deficiency in the standard force fields based on evidence such as Fig. 2.

Furthermore, the TREx + Amberff99SB + TIP4P-Ew Aβ ensembles strongly disagree with the Aβ42 ensemble generated by

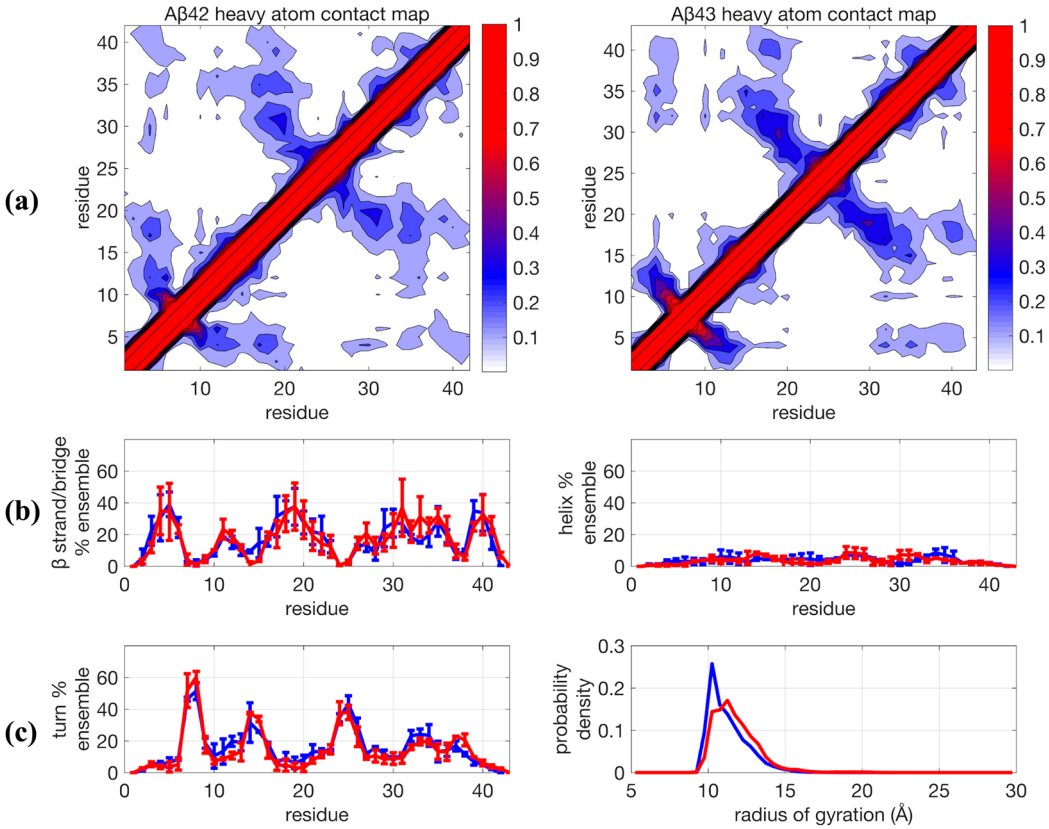

FIG. 2. Comparison of structural properties for Aβ42 and Aβ43 using the force field Amberff99sb + TIP4P-Ew and using Temperature Replica Exchange (TREx) simulated at 0.8 μs/replica. (a) Contact maps, (b) β−strand and helix propensities, and (c) turn propensity and radius of gyration distribution. For (b) and (c), blue lines are Aβ42 and red lines are Aβ43. Further details are given in the caption of Fig. 1.

Lin *et al.*[49] In their more extensive MD/MSM simulation, they found negligible β-sheet content, with the only persistent secondary structure being ~10%–20% α-helical content for residues 12-18. While the MD/MSM study used the same peptide force field as in this work (Amber ff99SB), they performed their simulations with the TIP3P water model[17] instead of TIP4P-Ew, which likely accounts for some quantitative differences in the Aβ ensembles. A study of the folding of Trp-cage using Amber ff99SB found that TIP3P enriched the sampling of helical content by ~10% and reduced the sampling of β-sheet content by a similar amount when compared to TIP4P-Ew.[73] These relatively minor population shifts are too small to explain the large discrepancies between the TREx and MD/MSM results; so much of the differences between the ensembles must result from the sampling efficiency. We believe this demonstrates that even on the μs time scale, TREx simulations are not able to fully capture the structural ensemble of Aβ peptides.

We next consider, whether an alternative sampling, namely, the TCW method, would converge faster on the ~0.1-0.2 μs/replica time scale to evaluate the different force field combinations (Fig. 3 and Table I). We have established in previous work that TCW is superior to the TREx approach using test systems where quantitative measures of convergence are available.[46,50] In addition, Figs. S2 and S3 show that the TCW method on the ~0.2 μs time scale reaches comparable convergence to the ~0.8 μs TREx simulation for all secondary structure categories. But as seen when comparing Figs. 2 and 3, the TREx and TCW enhanced sampling methods yield very different Aβ peptide ensembles using the same unmodified protein-water force field combination. More specifically, the TCW + ff99SB + TIP4P-Ew result is far less structured than that found with TREx and is much more similar to the MD/MSM results by Lin *et al.*[49] Although there are regions of the sequence that exhibit larger uncertainties in the turn population, this variation in turn content suggests the transient sampling of still overall unstructured conformations since the sampling of β and helical conformations is consistently low rather than the more significant sampling of highly structured conformations seen in the TREx simulations. For the unmodified force field combination using TCW, there is a very small amount (~5%-10%) for the contact region formed by residues 16-20 and 30-37 for both Aβ42 and Aβ43. It is unlikely that the experimental NOEs could absolutely rule out the presence of such a small population of long-ranged structure, and in fact the simulated NOEs would support this conclusion since the average NOE distance for these residues would be dominated by the ~90% of the unstructured populations.

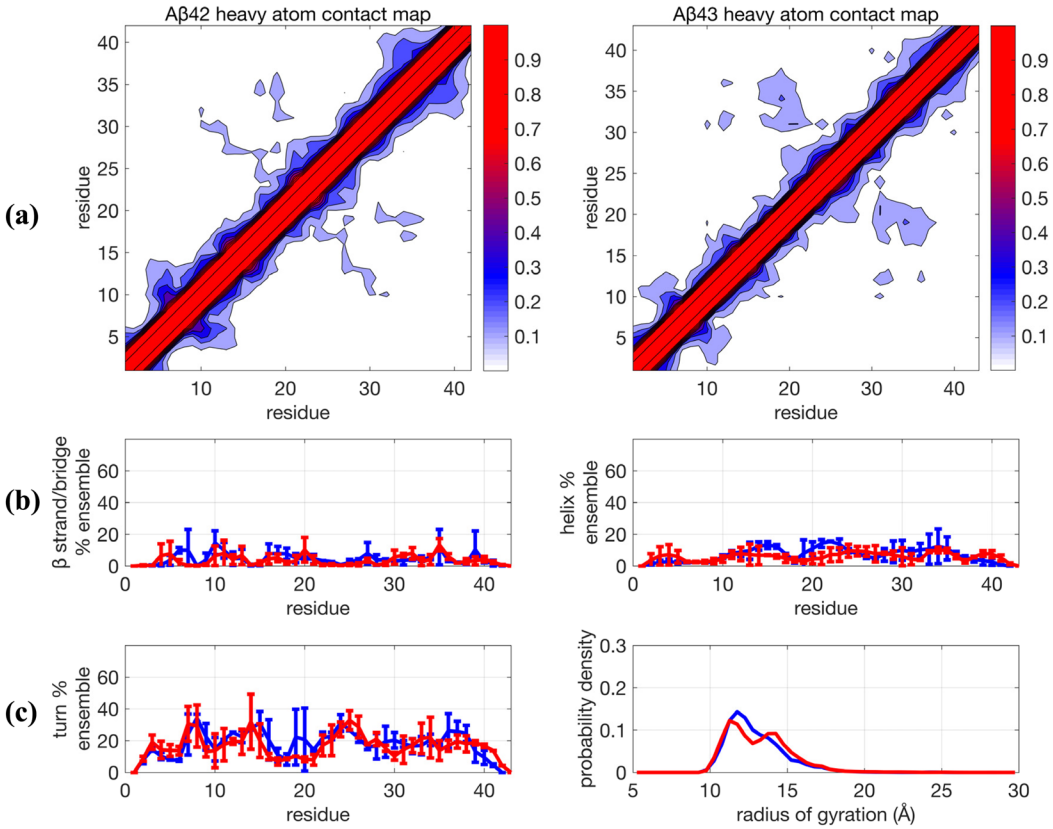

FIG. 3. Comparison of structural properties generated for Aβ42 and Aβ43 using the force field combination Amber99SB + TIP4P-Ew and using Temperature Cool Walking (TCW). (a) Contact maps, (b) β−strand and helix propensities, and (c) turn propensity and radius of gyration distribution. For (b) and (c), blue lines are Aβ42 and red lines are Aβ43. Further details are given in the caption of Fig. 1.

For completeness, we consider the newer force fields that have been shown to improve descriptions of IDP ensembles using TREx but now simulated with TCW. We first consider the Amber ff99SB-ILDN + TIP4P-D combination (Fig. 4), where the force field optimization is centered on increasing the strength of attractive London dispersion forces on the water oxygen atoms by ~50% relative to that of other four-site water models. Using TREx simulations, this has been reported to produce less structured IDP ensembles by increasing the strength of protein-water attractions relative to protein-protein attractions using standard Lennard-Jones mixing rules.[22] The results for TCW shown in Fig. 4 also confirm that this recently introduced water model generally reduces structural order and maintains structural similarity between the ensembles of the two peptides, with $\langle R_g \rangle$ ~ 14.9 Å and 15.7 Å for Aβ42 and Aβ43, respectively (Table I). There is again, as in the TCW + ff99SB + TIP4P-Ew result, some variability in the turn population, suggesting the sampling of some different unstructured substates in the different trajectories, but the β−sheet and α−helical populations are consistently small.

Similar structural results are obtained for the CHARMM36m + CHARMM-TIP3P force field combination using the TCW sampling method (Fig. 5), predicting more expanded structures as evidenced by a shifted and broader R_g distribution, with $\langle R_g \rangle$ ~ 15.9-16.4 Å for the two peptides (Table I), with no long-range contacts found using TREx as demonstrated through the absence of secondary structure signatures as well as the contact maps, in agreement with the NOEs for Aβ42.[5,6] Furthermore, the structural ensemble of the Aβ42 and Aβ43 free monomers is seen to be nearly identical, in agreement with the experimental results of Conicella and Fawzi.[4]

Experimental validation of simulated ensembles

In order to better validate these results for the different force fields, and to better understand the differences found between TREx and TCW, we consider back-calculations of the NMR and FRET data as a metric for comparing Aβ42 and Aβ43 ensembles to the experiment (Table I). Unfortunately, current state-of-the-art chemical shift calculators developed specifically for protein applications have large associated intrinsic back-calculation errors, making quantitative comparisons problematic for IDPs. For example, the RMS error for ^1H_N chemical shifts in the SHIFTX2 calculator is 0.17 ppm,[69] which is much larger than the experimental difference (<0.05 ppm) between the Aβ42 and Aβ43 ensembles.[4] Thus, chemical shifts are not an ideal metric to distinguish between the different simulated

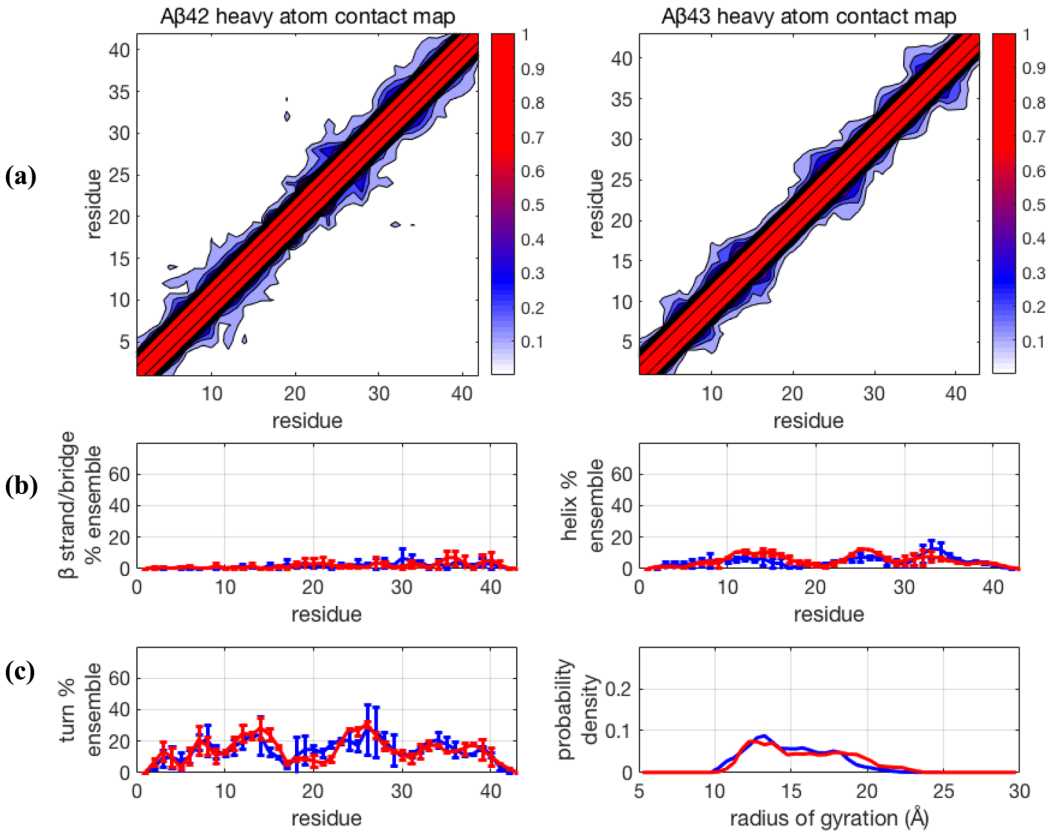

FIG. 4. Comparison of structural properties generated for Aβ42 and Aβ43 using the force field Amber99SB-ILDN + TIP4P-D and using Temperature Cool Walking (TCW). (a) Contact maps, (b) β−strand and helix propensities, and (c) turn propensity and radius of gyration distribution. For (b) and (c), blue lines are Aβ42 and red lines are Aβ43. Further details are given in the caption of Fig. 1.

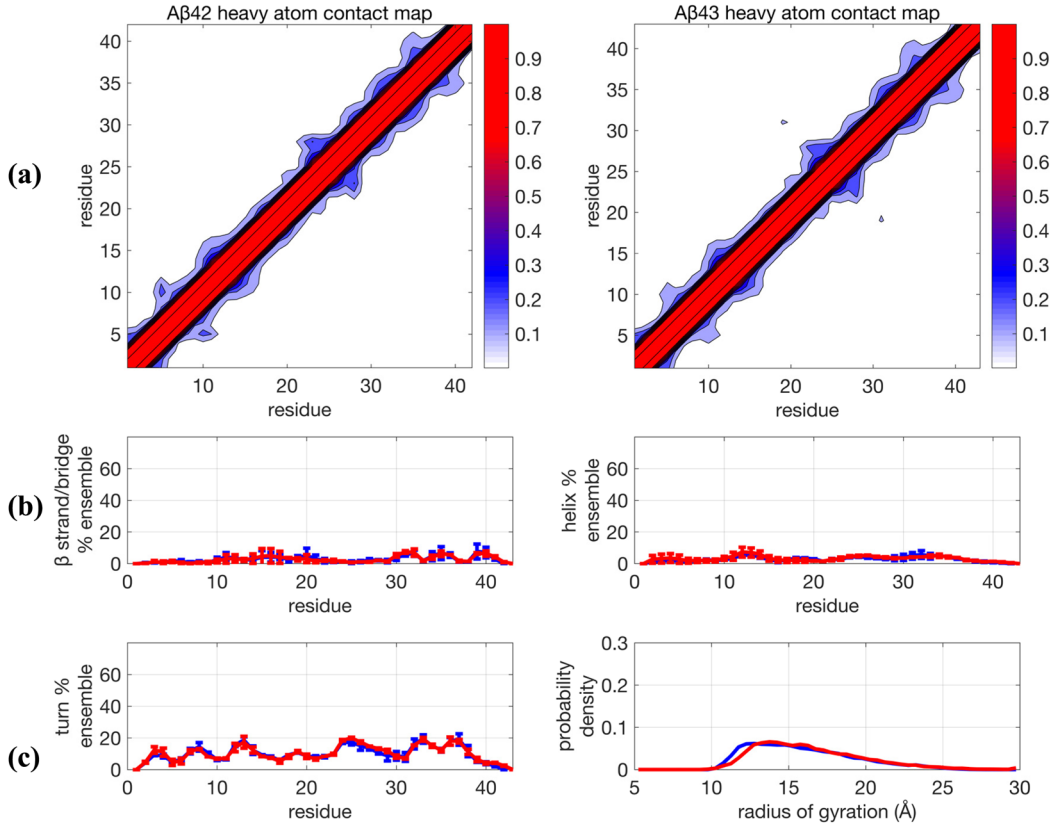

FIG. 5. Comparison of structural properties generated for Aβ42 and Aβ43 using the force field CHARMM36m + CHARMM-TIP3P and using Temperature Cool Walking (TCW). (a) Contact maps, (b) β−strand and turn propensities, and (c) α−helix propensity and radius of gyration distribution. For (b) and (c), blue lines are Aβ42 and red lines are Aβ43. Further details are given in the caption of Fig. 1.

ensembles, and as expected, the simulations are unable to capture the subtle differences in the $^{1}H_{N}$ chemical shifts in the C-terminus of the two peptides (examples given in Fig. S4).[33]

Table I shows the calculated χ^2 parameter from the simulated J-coupling constants ($J_{HN\text{-}H\alpha}$) for each ensemble as compared to the experiment, in which lower values of the χ^2 metric indicate better agreement with experimental J-coupling constants, with values near one indistinguishable from the scalar coupling back-calculation error. We compare to the Roche *et al.* experimental dataset[5] consisting of 38 J_i values for Aβ42 and the Conicella and Fawzi[4] experimental dataset consisting of 22 J_i values for Aβ43. The simulated J-couplings from the computational ensembles support the general conclusion that the TCW sampling method is in better agreement with the NMR J-couplings than found for TREx, regardless of force field for both Aβ42 and Aβ43 peptides, with the standard force field in best quantitative agreement.[4] We note that the χ^2 values appear to be commensurate with the results of Meng *et al.* for Aβ42 using 750 ns TREx simulations with Amber ff99SBws (χ^2 = 2.89) and Amber ff03ws (χ^2 = 4.58), two IDP-optimized force fields, using the same experimental data and set of Karplus parameters. However, their χ^2 values include a block-averaging error due to variance within the trajectories that depresses the value of the χ^2 (up to ~10%). We did not apply this block averaging error since this hides the intrinsic sampling problem we are investigating.

The conclusion that the TCW method yields better agreement with J-couplings is bolstered when using a Bayesian analysis we have developed, the Experimental Inferential Structure Determination (EISD) method.[72] The EISD method is designed to assess agreement given the available experimental J-coupling data (as well as chemical shifts, which we ignore here given their low predictive value) that take into account the intrinsic experimental and back-calculation uncertainties through optimization. In this case, the optimization occurs within the variance of the Gaussian distributed model for the back-calculation error for Karplus parameters A, B, and C in Eq. (1). We analyze each ensemble with EISD against the same set of experimental data used for the χ^2 analysis and find that the ranking of the ensembles does not change from the χ^2 analysis (Table I), i.e., the available J-coupling data are sufficient for concluding that the TCW ensembles are in better agreement with the NMR data, with significant better agreement for the standard force field.

Although we do not invoke a full scale simulation of NOE data as we have done in previous studies,[10,29,67,68] we can make some qualitative comparisons to the NOE data for Aβ42[5] using the

contact maps. Clearly, the newer force fields are in excellent agreement with the NOE data. Furthermore, the standard force field using the TCW simulations is also in good agreement with the NOE data, unlike the TREx simulations that contain a very high percentage of long-range contacts. For the unmodified force field combination using TCW, there is a very small amount (~5%-10%) for the contact region formed by residues 16-20 and 30-37 for both Aβ42 and Aβ43, but as we have already stated above, it is unlikely that the experimental NOEs could absolutely rule out the presence of such a small population of transient long-ranged structure.

Next, we consider the comparison of the different sampling and force field combinations to 2D FRET that have been reported recently for Aβ42;[6] we again make the reasonable assumption that FRET efficiencies, E_{FRET}, for Aβ43 will be nearly identical based on the results of Conicella and Fawzi.[4] In principle, E_{FRET} should be calculated for all conformations using the end-to-end distance between the dyes or tags, $R_{ee}^{tag}(t)$, i.e., the simulations should use the same sequence construct as the experiment that includes additional residues and the chemical specifics of the covalently bound dye molecules. The possibility that the IDP ensemble will be perturbed to some degree by these tags, as we have seen previously for the MTSL tag used in EPR studies,[51] obscures the means to compare R_{ee}^{untag}, from the simulation data of untagged peptides, with experimentally derived values of R_{ee}^{tag}. Hence, a model for the missing tags must be developed to make contact with the 2D FRET data.

It is useful to consider the FRET model used by Meng *et al.*[6] in which the distance between the dyes $R_{ee}^{tag}(t)$ is implicitly accounted for by scaling the R_{ee}^{untag} with an approximation that the effect of the dyes is equivalent to adding N_{tag} additional residues to the sequence length,[74]

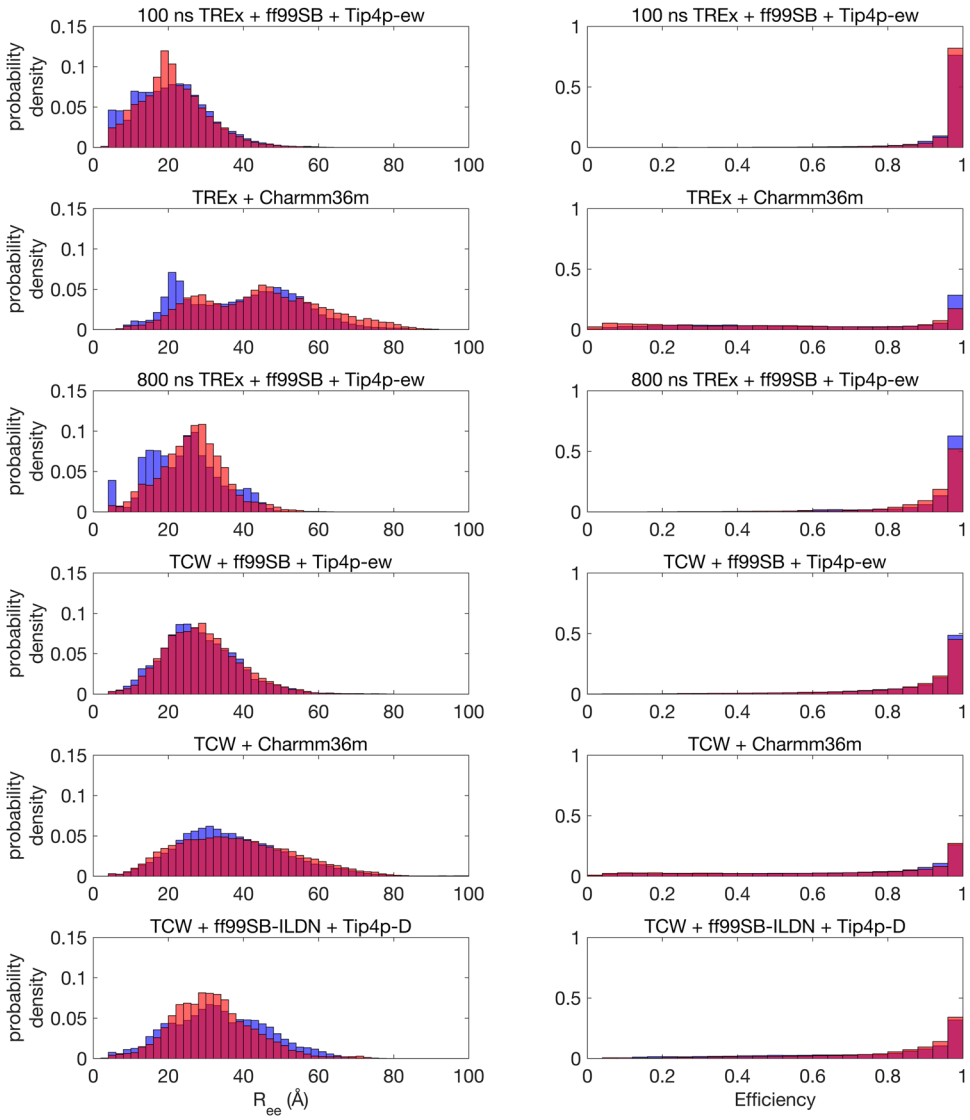

FIG. 6. Simulated end-to-end distance histograms (left) and corresponding FRET efficiency histograms using Eq. (3) (right) for all force field- sampling combinations. For all plots, blue bars are for Aβ42 and red bars are for Aβ43. All simulations are for the untagged peptides, with the results shown having incorporated a shift to model the additional residues and tags.

$$R_{ee}^{tag}(t) = R_{ee}^{untag}(t)\left(\frac{N + N_{tag}}{N}\right)^{0.5}. \quad (4)$$

The origin of the scaling model used for the missing tags [Eq. (4)] and value for N_{tag} has precedent in the literature from a study by McCarney and co-workers, who conducted 1 ns standard MD simulations on a model of Alexa 488 attached to a single cysteine residue; they obtained an ensemble average that does provide support for $N_{tag} = 12$ for a pair of dyes.[75] However, they also noted that this average conceals the chemical nature of the bimodal distribution, which results from an extended conformation corresponding to $N_{tag} \sim 16$, and a collapsed conformation due to hydrophobic interactions between the dye and its linker, with $N_{tag} \sim 4.4$. While McCarney *et al.* noted that the collapsed conformation would not likely persist in experiments with sufficient denaturants present, as is the case in many FRET studies including their own,[75] that would not be and is not the case for native, denaturant-free studies, as in the experiments of Meng *et al.*[6] for Aβ42 that study the peptide in more native-like conditions.

The resulting untagged peptide simulations that have been scaled using Eq. (4) with $N_{tag} = 12$ show a highly skewed distribution with a dominant peak at $E_{FRET} = 1$ (Fig. 6) that is in disagreement with the experiment which is peaked around the average FRET efficiency of ~0.63 for Aβ42. This same difference between the experimental and simulated result is also evident in the supplementary material in the work of Meng *et al.*[6] using the new IDP force fields developed by Best and co-workers.[23,27,38] The extremely high FRET efficiency is as expected given the relatively short length of the Aβ peptides and the large Förster radius for the dye pair of Alexa 488 and 647, i.e., as per Eq. (3), all conformations with the scaled values of R_{ee}^{untag} less than ~40 Å will yield $E_{FRET} \sim 1$.

Even so, the TCW simulations using the different force fields yield FRET efficiencies for Aβ42 of $\langle E_{FRET} \rangle = 0.71$–$0.88$ (Table I); to compare to the work of Meng *et al.* for Aβ42, the range of simulated FRET efficiencies they found using two different force fields designed for IDPs with the same value of $N_{tag} = 12$ gave values of $E_{FRET} = 0.68$–0.83 which they state is in good agreement with the experimental results.[6] By contrast, the 0.8 μs TREx simulation using a standard force field yields $E_{FRET} = 0.96$–0.97. Increasing N_{tag} to 16 or even 20 would only reduce E_{FRET} by a few percent; hence, this general difference between sampling methods would remain the same. Among the TCW simulations, those using modified force fields produce values of E_{FRET} that are lower and closer to the experimental values of Meng *et al.*, consistent with those force fields producing more expanded ensembles than standard ones.

Given the complete summary of the experimental validation, we also must conclude that the secondary structure propensities for the standard force field for the Aβ peptides are not as egregiously incorrect as ascertained from the TREx sampling method using a standard 0.8 μs/replica simulation. There is no question that the newer force fields agree well with the experimental data, but the TCW + Amber ff99SB + TIP4P-Ew also agrees as well and in fact even slightly better for the NMR J-coupling χ^2 and EISD evaluation. Given the uncertainties in the 2D FRET models, we can only conclude that the TCW simulations for the standard force fields are acceptable based on the upper bound value reported by Meng *et al.*[6] ($E_{FRET} = 0.83$) for one of their modified force fields. Using TCW, the standard force field yields negligible β-sheet and α-helical content, and the only observed long-range contacts are at very low population, and hence qualitatively consistent with the 2D FRET and NOEs taken on the Aβ42 peptide, where the 5%-10% population of long-range contacts will likely not be captured in the experiment.[5,6] As a result, the structural ensemble of both peptides generated using the TCW sampling method is more extended than what was found under the TREx protocol for the standard force field, with $\langle R_g \rangle \sim 12.9$–$13.2$ Å for the two peptides (Table I).

DISCUSSION

Given the better agreement with all of the experimental data using the TCW protocol over the TREx simulations, we use the TCW results to next address the question as to whether the new modified force fields introduce a genuine improvement over the standard force field. One assessment is whether the force fields are yielding structural ensembles consistent with a random coil ensemble modeled as a Gaussian random chain or a self-avoiding random walk (SARW). Although the Aβ ensemble using the standard force field does contain a small amount of residual structure, a rapidly interconverting unfolded or IDP ensemble will sample both extended and compact conformations containing regions of secondary structure seen in folded proteins that are still consistent with a random coil ensemble.[76]

The chain length scaling exponents for polymer models used to interpret SAXS, NMR, and FRET measurements for unfolded proteins and IDPs have been shown to be dependent on sequence characteristics such as charge and hydrophobicity.[77,78] Given that Aβ has both a net positive and relatively high hydrophobicity in the CHC region of its sequence, the scaling exponent might be expected to reduce to the Θ-limit for this IDP,[6]

$$R_g = R_0 N^{\nu}, \quad (5)$$

where $R_0 = 2.0$ Å as given by Fitzkee and Rose[79] and $\nu = 0.5$ for the Θ-limit which would yield $\langle R_g \rangle \sim 13$ Å that is consistent with the value calculated from the unmodified force field results simulated with TCW given in Table I. Under the Gaussian random coil model, we would conclude that the modified force fields have resulted in an over-correction by producing much more expanded ensembles than is warranted. However, if we assume a SARW model, we determine a larger value for R_g, i.e., using $\nu = 0.54$–0.6 in a good solvent[56] would yield $\langle R_g \rangle \sim 15$–$19$ Å, in better agreement with the modified force fields, thereby suggesting that the standard force fields are in fact too collapsed. Furthermore, for a Gaussian random chain, we would expect the following relationship to hold:

$$R_g = R_{ee}/\sqrt{6}. \quad (6)$$

But this correlation between R_g and R_{ee} is poor given the simulated data in Table I, except for the TCW + CHARMM36m + CHARMM-TIP3P combination, which is inconsistent with the Gaussian model based on R_g.

Thus, we view polymer physics models and analyses as largely inconclusive for differentiating the quality of force fields for this IDP system for several reasons. First is that the differences between the Gaussian and SARW models are better differentiated for much longer polymers than the small Aβ peptides investigated here. Furthermore, Fuertes *et al.* have presented some novel analytical techniques on a set of IDPs and denatured proteins using a variety of

dyes and denaturant conditions to evaluate the relationship between FRET and SAXS.[57] They suggest that one should generally decouple R_g from R_{ee} and thus avoid using a simple scaling law such as Eq. (6) that is independent of peptide chemistry.[57] Finally, the field is in need of new heteropolymer-centric theoretical models that are able to capture sequence details to extend beyond simple scaling laws and empirical relations, models which are now starting to be developed in recent work for the IDP class of proteins.[80,81]

We next turn to a more quantitative assessment using the 2D FRET data to ascertain the differences in force fields. The simulation of FRET efficiencies, especially for IDPs, involves a series of assumptions that introduce uncertainty that must be acknowledged when comparing to the FRET observable given the presence of the fluorescent dyes. As of late 2018, there is active debate on the perturbations introduced by fluorophore tags.[53–57] Fuertes *et al.*[54] assert that there is no perturbation of the structural ensemble across a series of disordered peptides upon addition of dye labels. Part of the basis of their conclusion is that they did not find significant shifts in the SAXS profiles or $\langle R_g \rangle$ of tagged and untagged peptides at high denaturant conditions.[57] If this is the case, then our 2D FRET results and analysis support the view that the standard force field for the untagged Aβ ensembles and tags modeled using Eq. (4) is in adequate to good agreement with the experiment, with the modified force fields performing only slightly better.

By contrast, Riback *et al.* concluded that at native conditions with no denaturant, the addition of tags in FRET experiments leads to interactions with the IDP that will contribute to FRET signals that overemphasize its contraction and thus artificially increases its FRET efficiency under denaturant-free conditions.[56] Furthermore, the perturbative effect of the dyes was seen to be larger for the smaller peptides in their experiments, where the addition of labels and the residues to which they are bonded has a greater effect on the mass and resulting dynamics and structure of the peptide. For example, the addition of the tags produced shifts from −0.3 to +0.5 nm in the average R_g of the two smallest peptides studied, N49 and NLS, which are natively 36 and 44 residues.[56] Even the results of Meng *et al.* found an increase in E_{FRET} upon explicit representation of the Alexa fluorescent tags using a standard MD calculation for Aβ42,[6] which is consistent with greater compaction of the ensemble. Our own recent work demonstrated that addition of a hydrophobic MTSL-Cys tag to Aβ42 can perturb the structural ensemble through interactions between the dye and the peptide that in turn leads to more a collapsed structural ensemble compared to the original peptide.[51] The simulations of Fuertes *et al.*, however, modeled the unlabeled peptide using implicit solvent and then built up an ensemble of structures of the labeled peptide by stochastically adding the dyes to the unlabeled conformers.[57] While this produces a useful test for verifying that many different values of R_{ee} can be obtained from an ensemble with the same R_g, it does not address the potential phenomenon of the dyes affecting the chain dynamics and structure directly. If the tags do induce an artificial compaction of the ensemble, then the true experimental ensembles would be more expanded with even lower FRET efficiencies than reported for Aβ42. If that is the case, then the standard force fields yield IDP ensembles that are in fact too collapsed, and force field modifications are warranted, especially for small Aβ42 and Aβ43 peptides studied in this work.

CONCLUSION

We have simulated the disordered structural ensembles of the Aβ42 and Aβ43 peptides, which according to recent experiments[4–6] should be largely the same and exhibit no persistent structural ordering or long-range contacts. But two types of error can occur during computational studies of IDP structural ensembles that prevent connections to such experiments, namely, statistical sampling error and systematic error in energy and forces. Statistical sampling error occurs when the simulations have not been run sufficiently long to achieve convergence, while systematic error happens when the energy surface of the peptide-water system is not modeled with accurate molecular interactions. Not surprisingly, these two potential errors are intertwined, and hence, we have attempted to consider them both by comparing two sampling methods, TREx and TCW, as well as comparing a standard protein-water force field and those that have been recently modified to yield better modeling of disordered states. We have also attempted to validate the various simulated ensembles by comparing to recent state of the art NMR and 2D FRET experiments.[4–6]

While it is starting to become established that long MD trajectories of ~100 μs–1 ms are often necessary to reveal force field deficiencies,[11,49] simulation time scales that are largely routine only on specialized hardware such as Anton[82] or Folding@home,[83] the hope is that better enhanced or accelerated sampling methods might converge more quickly, i.e., with one to two orders of magnitude less effort. In this work, we have shown that even at μs time scales there appear to be limitations in the TREx sampling method, producing far more structured ensembles that are in disagreement with NMR and 2D FRET validation data on Aβ peptides. Changing the sampling method from TREx to TCW produces ensembles that are qualitatively in agreement with J-couplings, NOEs, and FRET efficiencies, regardless of the force field that is simulated. The TCW results for Aβ42 are also in very good agreement with the very long sub-millisecond MD/MSM results by Lin and co-workers.[49] Thus, our evidence has shown that what was thought to be primarily a force field problem was masked by what is at least in part a problem of poor sampling. This work establishes that the TCW method is more effective than TREx when entropic barriers dominate and, when applied to IDPs, supports the recent hypothesis that IDPs have an inverted free energy landscape in which disordered conformations are lower in free energy than ordered structures.[52]

Sampling is an important consideration for establishing the transferability of any force field, by demonstrating that appropriate conformational equilibria are reached across a range of thermodynamic conditions in order to describe folded and unfolded states of globular proteins as well as IDP sequences. For example, some standard protein and water force field combinations have proven robust for understanding mechanistic questions about protein folding,[84] which requires a good model of the structure and internal dynamics of the unfolded states in addition to the folded state of a globular protein.[85] But historically, it was the ability to sample multiple folding and unfolding events using these standard force fields which allowed them to gain validation through direct comparisons to robust folding experiments.[86,87] This work rescues some of these standard force fields in the sense that they require extensive

sampling to definitively show whether they are also capable of simulating accurate IDP ensembles.

The modified force fields may have had the effect of inherently lowering the entropic barriers and, in the best case, still maintaining the folded-unfolded equilibrium. However, unlike the case of protein folding, experimental validation is more limited and underdetermined for IDPs for a variety of reasons. At present, there is an impedance mismatch for chemical shifts and FRET data; for chemical shifts, the fault is on the theoretical side because of the reduced capability to back-calculate shifts from structure,[72] whereas for FRET measurements on very small IDPs, there remains the possibility that the presence of the fluorescent tags might perturb the IDP ensemble from its equilibrium state.[56] However, the impedance match between theory and experiment for scalar couplings and NOEs has provided strong support for the conclusion that, when simulated with TCW or using a very long MD/MSM simulation, the Amberff99SB + TIP4P-Ew, CHARMM36m + CHARMM-TIP3P, and Amberff99SB-ILDN + TIP4P-D are all appropriate force fields for IDPs. While clearly some of the new force field modifications can promote more expanded monomer ensembles to reproduce many experimental IDP properties more expediently, it is important to remember that the IDP-specific force fields[38] and other modified force fields may come with their own limitations, such as now manifesting native state instability,[33] thereby forgoing the ability to simulate disorder to order transitions in folding upon binding events that are part of the greater functional repertoire of proteins with intrinsic disorder.[2]

SUPPLEMENTARY MATERIAL

See supplementary material for comparison of other sampling methods and force fields, end-to-end distance and FRET efficiency histograms, simulated chemical shifts, and EISD results between the Aβ43 and Aβ42 structural ensembles.

ACKNOWLEDGMENTS

We would like to thank Ad Bax, Robert Best, Nick Fawzi, Julie Forman-Kay, Alex MacKerell, Stefano Piana-Agostinetti, and Kevin Plaxco for helpful discussions. We thank Reviewer 1 for very helpful comments and David H. Brookes for early calculations using EISD. S.S. and T.H.-G. were supported by the National Institutes of Health (NIH) Grant No. U01GM121667-02, and J.L. was partially supported on No. 5R01GM127627-01. J.L. also acknowledges partial support from an NIH Molecular Biophysics Training Grant No. T32-GM008295 during early stages of this study. This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. This research also used GPU resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

REFERENCES

[1] A. K. Dunker, I. Silman, V. N. Uversky, and J. L. Sussman, Curr. Opin. Struct. Biol. **18**(6), 756–764 (2008).
[2] H. J. Dyson and P. E. Wright, Nat. Rev. Mol. Cell Biol. **6**(3), 197–208 (2005).
[3] M. Goedert and M. G. Spillantini, Science **314**(5800), 777–781 (2006).
[4] A. E. Conicella and N. L. Fawzi, Biochemistry **53**(19), 3095–3105 (2014).
[5] J. Roche, Y. Shen, J. H. Lee, J. Ying, and A. Bax, Biochemistry **55**(5), 762–775 (2016).
[6] F. Meng, M. M. J. Bellaiche, J.-Y. Kim, G. H. Zerze, R. B. Best, and H. S. Chung, Biophys. J. **114**(4), 870–884 (2018).
[7] J. A. Marsh and J. D. Forman-Kay, Proteins: Struct., Funct., Bioinf. **80**(2), 556–572 (2012).
[8] M. Krzeminski, J. A. Marsh, C. Neale, W.-Y. Choy, and J. D. Forman-Kay, Bioinformatics **29**(3), 398–399 (2013).
[9] A. Bhowmick, D. H. Brookes, S. R. Yost, H. J. Dyson, J. D. Forman-Kay, D. Gunter, M. Head-Gordon, G. L. Hura, V. S. Pande, D. E. Wemmer, P. E. Wright, and T. Head-Gordon, J. Am. Chem. Soc. **138**(31), 9730–9742 (2016).
[10] K. A. Ball, D. E. Wemmer, and T. Head-Gordon, J. Phys. Chem. B **118**, 6405–6416 (2014).
[11] P. S. Nerenberg and T. Head-Gordon, Curr. Opin. Struct. Biol. **49**, 129–138 (2018).
[12] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling, Proteins: Struct., Funct., Bioinf. **65**(3), 712–725 (2006).
[13] M. Christen, P. H. Hünenberger, D. Bakowies, R. Baron, R. Bürgi, D. P. Geerke, T. N. Heinz, M. A. Kastenholz, V. Kräutler, C. Oostenbrink, C. Peter, D. Trzesniak, and W. F. van Gunsteren, J. Comput. Chem. **26**(16), 1719–1751 (2005).
[14] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, J. Am. Chem. Soc. **118**(45), 11225–11236 (1996).
[15] G. A. Kaminski, R. A. Friesner, J. Tirado-Rives, and W. L. Jorgensen, J. Phys. Chem. B **105**(28), 6474–6487 (2001).
[16] 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. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus, J. Phys. Chem. B **102**(18), 3586–3616 (1998).
[17] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**(2), 926–935 (1983).
[18] W. L. Jorgensen and J. D. Madura, Mol. Phys. **56**(6), 1381–1392 (1985).
[19] 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. **120**(20), 9665–9678 (2004).
[20] W. Wang, W. Ye, C. Jiang, R. Luo, and H.-F. Chen, Chem. Biol. Drug Des. **84**(3), 253–269 (2014).
[21] J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. de Groot, H. Grubmüller, and A. D. MacKerell, Jr., Nat. Methods **14**, 71 (2016).
[22] S. Piana, A. G. Donchev, P. Robustelli, and D. E. Shaw, J. Phys. Chem. B **119**(16), 5113–5123 (2015).
[23] R. B. Best and J. Mittal, J. Phys. Chem. B **114**(46), 14916–14923 (2010).
[24] P. H. Nguyen, M. S. Li, and P. Derreumaux, Phys. Chem. Chem. Phys. **13**(20), 9778–9788 (2011).
[25] C. M. Siwy, C. Lockhart, and D. K. Klimov, PLoS Comput. Biol. **13**(1), e1005314 (2017).
[26] S. Rauscher, V. Gapsys, M. J. Gajda, M. Zweckstetter, B. L. de Groot, and H. Grubmüller, J. Chem. Theory Comput. **11**(11), 5513–5524 (2015).
[27] R. B. Best and J. Mittal, Proteins: Struct., Funct., Bioinf. **79**(4), 1318–1328 (2011).
[28] W. R. P. Scott, P. H. Hünenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Krüger, and W. F. van Gunsteren, J. Phys. Chem. A **103**(19), 3596–3607 (1999).
[29] N. L. Fawzi, A. H. Phillips, J. Z. Ruscio, M. Doucleff, D. E. Wemmer, and T. Head-Gordon, J. Am. Chem. Soc. **130**(19), 6145–6158 (2008).
[30] W. L. Song, Y. Y. Wang, J. P. Colletier, H. Y. Yang, and Y. C. Xu, Sci. Rep. **5**, 11024 (2015).
[31] Z. Amini, M. H. Fatemi, and A. Rauk, Can. J. Chem. **94**(10), 833–841 (2016).
[32] M. H. Viet, P. H. Nguyen, P. Derreumaux, and M. S. Li, ACS Chem. Neurosci. **5**(8), 646–657 (2014).
[33] P. Robustelli, S. Piana, and D. E. Shaw, Proc. Natl. Acad. Sci. U. S. A. **115**(21), E4758–E4766 (2018).

[34]P. S. Nerenberg and T. Head-Gordon, J. Chem. Theory Comput. **7**(4), 1220–1230 (2011).
[35]P. S. Nerenberg, B. Jo, C. So, A. Tripathy, and T. Head-Gordon, J. Phys. Chem. B **116**(15), 4524–4534 (2012).
[36]J. Henriques, C. Cragnell, and M. Skepö, J. Chem. Theory Comput. **11**(7), 3420–3431 (2015).
[37]J. Henriques and M. Skepö, J. Chem. Theory Comput. **12**(7), 3407–3415 (2016).
[38]R. B. Best, W. Zheng, and J. Mittal, J. Chem. Theory Comput. **10**(11), 5113–5124 (2014).
[39]C. Abrams and G. Bussi, Entropy **16**(1), 163 (2014).
[40]R. C. Bernardi, M. C. R. Melo, and K. Schulten, Biochim. Biophys. Acta **1850**(5), 872–877 (2015).
[41]Y. Sugita and Y. Okamoto, Chem. Phys. Lett. **314**(1-2), 141–151 (1999).
[42]U. Hansmann, Chem. Phys. Lett. **281**, 140–150 (1997).
[43]D. M. Zuckerman and E. Lyman, J. Chem. Theory Comput. **2**, 1200–1202 (2006).
[44]D. J. Rosenman, C. Wang, and A. E. García, J. Phys. Chem. B **120**(2), 259–277 (2016).
[45]H. Nymeyer, J. Chem. Theory Comput. **4**(4), 626–636 (2008).
[46]S. Brown and T. Head-Gordon, J. Comput. Chem. **24**(1), 68–76 (2003).
[47]A. Hicks and H.-X. Zhou, J. Chem. Phys. **149**(7), 072313 (2018).
[48]T. N. Do, W.-Y. Choy, and M. Karttunen, J. Chem. Theory Comput. **12**(1), 395–404 (2016).
[49]Y. S. Lin, G. R. Bowman, K. A. Beauchamp, and V. S. Pande, Biophys. J. **102**(2), 315–324 (2012).
[50]J. Lincoff, S. Sasmal, and T. Head-Gordon, J. Chem. Phys. **145**(17), 174107 (2016).
[51]S. Sasmal, J. Lincoff, and T. Head-Gordon, Biophys. J. **113**(5), 1002–1011 (2017).
[52]D. Granata, F. Baftizadeh, J. Habchi, C. Galvagnion, A. De Simone, C. Camilloni, A. Laio, and M. Vendruscolo, Sci. Rep. **5**, 15449 (2015).
[53]J. A. Riback, M. A. Bowman, A. Zmyslowski, C. R. Knoverek, J. Jumper, E. B. Kaye, K. F. Freed, P. L. Clark, and T. R. Sosnick, Science **361**(6405), eaar7949 (2018).
[54]G. Fuertes, N. Banterle, K. M. Ruff, A. Chowdhury, R. V. Pappu, D. I. Svergun, and E. A. Lemke, Science **361**(6405), eaau8230 (2018).
[55]R. B. Best, W. Zheng, A. Borgia, K. Buholzer, M. B. Borgia, H. Hofmann, A. Soranno, D. Nettels, K. Gast, A. Grishaev, and B. Schuler, Science **361**(6405), eaar7101 (2018).
[56]J. A. Riback, M. A. Bowman, A. M. Zmyslowski, C. R. Knoverek, J. M. Jumper, J. R. Hinshaw, E. B. Kaye, K. F. Freed, P. L. Clark, and T. R. Sosnick, Science **358**(6360), 238 (2017).
[57]G. Fuertes, N. Banterle, K. M. Ruff, A. Chowdhury, D. Mercadante, C. Koehler, M. Kachala, G. Estrada Girona, S. Milles, A. Mishra, P. R. Onck, F. Gräter, S. Esteban-Martín, R. V. Pappu, D. I. Svergun, and E. A. Lemke, Proc. Natl. Acad. Sci. U. S. A. **114**(31), E6342 (2017).
[58]D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods, J. Comput. Chem. **26**(16), 1668–1688 (2005).
[59]N. Rathore, M. Chopra, and J. J. de Pablo, J. Chem. Phys. **122**(2), 024111 (2004).
[60]K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, and D. E. Shaw, Proteins: Struct., Funct., Bioinf. **78**(8), 1950–1958 (2010).
[61]N. G. Sgourakis, Y. L. Yan, S. A. McCallum, C. Y. Wang, and A. E. Garcia, J. Mol. Biol. **368**(5), 1448–1457 (2007).
[62]D. J. Rosenman, C. R. Connors, W. Chen, C. Y. Wang, and A. E. Garcia, J. Mol. Biol. **425**(18), 3338–3359 (2013).
[63]D. A. Case, R. M. Betz, D. S. Cerutti, T. E. Cheatham III, T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K. M. Merz, G. Monard, H. Nguyen, H. T. Nguyen, I. Omelyan, A. Onufriev, D. R. Roe, A. Roitberg, C. Sagui, C. L. Simmerling, W. M. Botello-Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, L. Xiao, and P. A. Kollman, AMBER 2016, University of California, San Francisco, 2016.
[64]P. Eastman, M. S. Friedrichs, J. D. Chodera, R. J. Radmer, C. M. Bruns, J. P. Ku, K. A. Beauchamp, T. J. Lane, L.-P. Wang, and D. Shukla, J. Chem. Theory Comput. **9**(1), 461 (2013).
[65]D. R. Roe and T. E. Cheatham, J. Chem. Theory Comput. **9**(7), 3084–3095 (2013).
[66]W. Kabsch and C. Sander, Biopolymers **22**(12), 2577–2637 (1983).
[67]K. A. Ball, A. H. Phillips, P. S. Nerenberg, N. L. Fawzi, D. E. Wemmer, and T. Head-Gordon, Biochemistry **50**(35), 7612–7628 (2011).
[68]K. A. Ball, A. H. Phillips, D. E. Wemmer, and T. Head-Gordon, Biophys. J. **104**(12), 2714–2724 (2013).
[69]B. Han, Y. Liu, S. W. Ginzinger, and D. S. Wishart, J. Biomol. NMR **50**, 43–57 (2011).
[70]G. W. Vuister and A. Bax, J. Am. Chem. Soc. **115**, 7772–7777 (1993).
[71]B. Vögeli, J. Ying, A. Grishaev, and A. Bax, J. Am. Chem. Soc. **129**(30), 9377–9385 (2007).
[72]D. H. Brookes and T. Head-Gordon, J. Am. Chem. Soc. **138**(13), 4530–4538 (2016).
[73]D. Paschek, R. Day, and A. E. García, Phys. Chem. Chem. Phys. **13**(44), 19840–19847 (2011).
[74]W. Zheng, A. Borgia, M. B. Borgia, B. Schuler, and R. B. Best, J. Chem. Theory Comput. **11**, 5543 (2015).
[75]E. R. McCarney, J. H. Werner, S. L. Bernstein, I. Ruczinski, D. E. Makarov, P. M. Goodwin, and K. W. Plaxco, J. Mol. Biol. **352**(3), 672–682 (2005).
[76]L. J. Smith, K. M. Fiebig, H. Schwalbe, and C. M. Dobson, Folding Des. **1**(5), R95–R106 (1996).
[77]M. Aznauryan, L. Delgado, A. Soranno, D. Nettels, J.-r. Huang, A. M. Labhardt, S. Grzesiek, and B. Schuler, Proc. Natl. Acad. Sci. U. S. A. **113**(37), E5389–E5398 (2016).
[78]H. Hofmann, A. Soranno, A. Borgia, K. Gast, D. Nettels, and B. Schuler, Proc. Natl. Acad. Sci. U. S. A. **109**(40), 16155–16160 (2012).
[79]N. C. Fitzkee and G. D. Rose, Proc. Natl. Acad. Sci. U. S. A. **101**(34), 12497–12502 (2004).
[80]Y.-H. Lin and H. S. Chan, Biophys. J. **112**(10), 2043–2046 (2017).
[81]T. Firman and K. Ghosh, J. Chem. Phys. **148**(12), 123305 (2017).
[82]K. Lindorff-Larsen, P. Maragakis, S. Piana, M. P. Eastwood, R. O. Dror, and D. E. Shaw, PLoS One **7**(2), e32131 (2012).
[83]V. S. Pande, I. Baker, J. Chapman, S. P. Elmer, S. Khaliq, S. M. Larson, Y. M. Rhee, M. R. Shirts, C. D. Snow, E. J. Sorin, and B. Zagrovic, Biopolymers **68**(1), 91–109 (2003).
[84]S. Piana, J. L. Klepeis, and D. E. Shaw, Curr. Opin. Struct. Biol. **24**, 98–105 (2014).
[85]I. Echeverria, D. E. Makarov, and G. A. Papoian, J. Am. Chem. Soc. **136**(24), 8708–8713 (2014).
[86]A. Soranno, B. Buchli, D. Nettels, R. R. Cheng, S. Müller-Späth, S. H. Pfeil, A. Hoffmann, E. A. Lipman, D. E. Makarov, and B. Schuler, Proc. Natl. Acad. Sci. U. S. A. **109**(44), 17800–17806 (2012).
[87]W. A. Eaton and P. G. Wolynes, Proc. Natl. Acad. Sci. U. S. A. **114**(46), E9759–E9760 (2017).
[88]B. E. Husic and V. S. Pande, J. Am. Chem. Soc. **140**(7), 2386–2396 (2018).