Developing a molecular dynamics force field for both folded and disordered protein states

Paul Robustelli[a], Stefano Piana[a,1], and David E. Shaw[a,b,1]

[a]D. E. Shaw Research, New York, NY 10036; and [b]Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032

Molecular dynamics (MD) simulation is a valuable tool for characterizing the structural dynamics of folded proteins and should be similarly applicable to disordered proteins and proteins with both folded and disordered regions. It has been unclear, however, whether any physical model (force field) used in MD simulations accurately describes both folded and disordered proteins. Here, we select a benchmark set of 21 systems, including folded and disordered proteins, simulate these systems with six state-of-the-art force fields, and compare the results to over 9,000 available experimental data points. We find that none of the tested force fields simultaneously provided accurate descriptions of folded proteins, of the dimensions of disordered proteins, and of the secondary structure propensities of disordered proteins. Guided by simulation results on a subset of our benchmark, however, we modified parameters of one force field, achieving excellent agreement with experiment for disordered proteins, while maintaining state-of-the-art accuracy for folded proteins. The resulting force field, a99SB-*disp*, should thus greatly expand the range of biological systems amenable to MD simulation. A similar approach could be taken to improve other force fields.

computer simulations | intrinsically disordered proteins | protein dynamics

Many biologically important functions are carried out by disordered proteins or proteins containing structurally disordered regions. Unlike folded proteins, disordered proteins have native states that lack a well-defined tertiary structure. To structurally characterize such proteins, with the aim of ultimately giving mechanistic insight into their function, it is necessary to determine the heterogeneous ensembles of conformations that they adopt. One potential approach is molecular dynamics (MD) simulation, which, in principle, provides a direct computational route to determining structurally disordered states in atomic detail. The quality of MD simulation results is, however, strongly dependent on the accuracy of the physical model (force field) used.

Significant progress has recently been made in the ability of MD force fields to accurately describe folded proteins (1–6). Despite these remarkable successes, however, initial comparisons of MD simulations of disordered proteins and peptides with experimental measurements from techniques including NMR spectroscopy, small angle X-ray scattering (SAXS), and FRET showed significant discrepancies (7–9). Our study of multiple popular force fields and water models (7), for example, showed that all tested combinations produced disordered states that were substantially more compact than estimated from experiments.

There have been a number of attempts to improve the ability of force fields to describe disordered states. Head-Gordon and coworkers (9) optimized solvent–water van der Waals (vdW) interactions to reproduce experimental solvation free energies for a number of model organic compounds. Best et al. (8) rescaled protein–water interactions in the a03w protein force field (10) by a constant factor to produce more realistic dimensions of unfolded states of proteins. Recently, we found that dispersion interactions in the water models used for MD simulation are severely underestimated; simulations performed with a water model that was designed to have a more balanced description of dispersion and electrostatic interactions produced

disordered states that more closely agree with experimental measurements (7). Unfortunately, in preliminary tests, this water model sometimes resulted in less accurate simulations of folded proteins (7). Initial studies of the other force-field improvements mentioned above with folded proteins are more encouraging (8, 9). In the absence of large-scale systematic tests of force-field accuracy, however, it has been unclear whether any force field currently in use can accurately describe both folded and disordered protein states. A force field that is capable of providing accurate descriptions of both ordered and disordered proteins is naturally highly desirable, as it would enable simulations of, for example, proteins containing both ordered and disordered regions and proteins that transition between ordered and disordered states.

In this investigation, we systematically and quantitatively assess the accuracy of a number of state-of-the-art force fields from the CHARMM and Amber families through MD simulations of a variety of ordered and disordered proteins. We assembled a large benchmark set of 21 experimentally well-characterized proteins and peptides, including folded proteins, fast-folding proteins, weakly structured peptides, disordered proteins with some residual secondary structure, and disordered proteins with almost no detectable secondary structure. This benchmark set contains over 9,000 previously reported experimental data points. The Amber force fields tested were a99SB*-ILDN (11, 12) with the TIP3P water model (13), a99SB-ILDN with the TIP4P-D water model (7), the a03ws force field containing empirically optimized solute–solvent dispersion interactions (8), and the a99SB force field with modified Lennard–Jones (LJ)

Significance

Many proteins that perform important biological functions are completely or partially disordered under physiological conditions. Molecular dynamics simulations could be a powerful tool for the structural characterization of such proteins, but it has been unclear whether the physical models (force fields) used in simulations are sufficiently accurate. Here, we systematically compare the accuracy of a number of different force fields in simulations of both ordered and disordered proteins, finding that each force field has strengths and limitations. We then describe a force field that substantially improves on the state-of-the-art accuracy for simulations of disordered proteins without sacrificing accuracy for folded proteins, thus broadening the range of biological systems amenable to molecular dynamics simulations.

parameters proposed by Head-Gordon and coworkers (9). The CHARMM force fields tested were C22* (14) and C36m (6). C36m is a recent update to the C36 force field that was shown to greatly improve the structural properties of small disordered peptides, but that does not solve the problem of overcompactness of disordered proteins. In general, we find that the tested force fields give results in good agreement with experiment for many benchmark systems but that none of these previously existing force fields produce accurate dimensions and residual secondary structure propensities for disordered proteins while simultaneously providing accurate descriptions of folded proteins.

We complete our investigation by attempting to improve the parameters of an existing force field. Using the a99SB-ILDN protein force field with the TIP4P-D water model as a starting point, we optimized torsion parameters and introduced small changes in the protein and water vdW interaction terms, resulting in a force field, a99SB-*disp*, that achieves unprecedented levels of accuracy in simulations of disordered protein states while maintaining state-of-the-art accuracy for folded proteins. The parameters were obtained by iteratively introducing parameter modifications to reduce the observed discrepancies between simulations and experimental measurements on a subset of our benchmark dataset. The final parameters were tested on the remainder of the benchmark to reduce the risk of overfitting. We expect that a99SB-*disp* will enable substantially more accurate simulations to be carried out for a range of important biological systems. The example of a99SB-*disp* also shows that the simplified functional forms currently used in MD force fields are not incompatible with the accurate simulation of both ordered and disordered protein states with a single set of parameters; it is likely that the parameters of other force fields can similarly be improved by using the benchmark presented here.

Results

Composition of the Benchmark Set. To determine the accuracy of force fields for both ordered and disordered protein states, we assembled a benchmark set of 21 proteins and peptides, including folded and disordered systems (Fig. 1). Over 9,000 experimental data points are available for this set of proteins.

This benchmark set includes four folded proteins [ubiquitin, GB3, hen egg white lysozyme (HEWL), and bovine pancreatic trypsin inhibitor (BPTI)] that have been characterized by experimental NMR J couplings, residual dipolar couplings (RDCs), and order parameters that describe their conformational fluctuations. Calmodulin, a multidomain protein with two folded domains connected by a flexible linker (15), has been characterized by experimental NMR chemical shifts and RDCs that report on the structure and dynamics of the folded domains and flexible

linker and SAXS scattering curves that report on the overall dimensions of the solution ensemble. Simulations of calmodulin can simultaneously probe the ability of a force field to describe flexibility in the linker region, to avoid overly compact structures, and to maintain the structures of globular folded domains. To probe the ability of the force fields to accurately describe the equilibrium between ordered and disordered conformations, we examined the temperature-dependent native-state stability of the fast-folding variant of the villin head piece (referred to here as villin) (16), Trp-cage (17), the GTT variant of the WW domain FiP35 (referred to here as GTT) (18), the helical $(AAQAA)_3$ 15-mer peptide (19), and the small β-hairpin–forming peptide CLN025 (20).

We also selected for inclusion in the benchmark a set of proteins that are disordered under physiological conditions and for which extensive sets of NMR and SAXS data are available. Disordered proteins can vary widely in terms of local order, residual secondary structure propensities, and compactness, and our selections reflect this diversity. The benchmark includes the disordered proteins ACTR (21), drkN SH3 (22), α-synuclein (23), the N_{TAIL} domain of the measles virus nucleoprotein (24), Aβ40 (25), the ParE2-associated antitoxin PaaA2 (26), the proliferating cell nuclear antigen-associated factor $p15^{PAF}$ (27), the cyclin-dependent kinase inhibitor Sic1 (28), and an intrinsically disordered region from the *Saccharomyces cerevisiae* transcription factor Ash1 (29) (a region that we will refer to simply as Ash1). Simulations of disordered proteins were compared with experimental NMR J couplings, chemical shifts, and RDCs to assess the accuracy of local conformational distributions; experimental paramagnetic relaxation enhancements (PREs) to assess the accuracy of transient tertiary contacts; and experimental SAXS scattering data to determine the accuracy of simulated radii of gyration (Rg). We also included the bZip domain of the GCN4 transcription factor (which we refer to as GCN4) (30), a partially disordered dimer with an ordered helical coiled coil dimerization domain. Simulations of GCN4 were compared with experimental NMR chemical shifts and order parameters to assess local conformational distributions and fluctuations. To study the ability of force fields to describe an unstructured peptide, we examined the disordered polyalanine peptide Ala_5 (31), for which NMR J couplings are available. A full list of experimental measurements used to evaluate the accuracy of simulations is contained in *SI Appendix*, Table S1.

Assessment of the Ability of Current Force Fields to Reproduce the Experimental Data. We examined the ability of a number of force fields to accurately reproduce experimental data for the benchmark set. The force fields examined were the Amber force fields a99SB*-ILDN (11, 12) with TIP3P (13), a03ws (8), a99SB-ILDN with TIP4P-D (7), and a99SB with TIP4P-Ew (32) and the
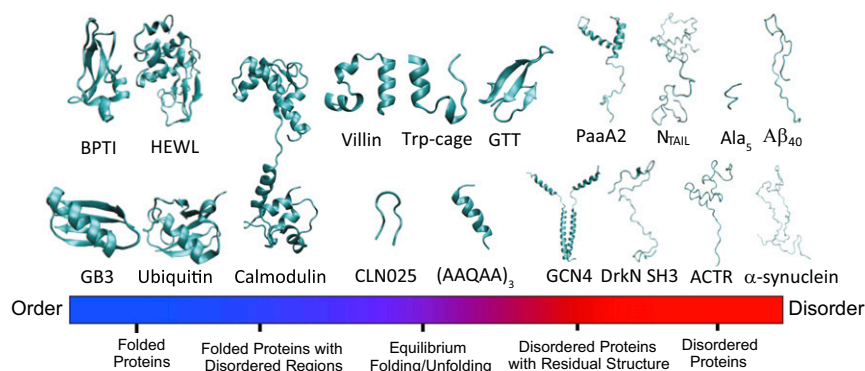
Fig. 1. Schematic illustration of systems contained in the benchmark set of proteins simulated in this work to assess and refine the accuracy of protein force fields for both ordered and disordered protein states.

Head-Gordon LJ (9) and dihedral (33) modifications (we refer to this force field as a99SB-UCB) as well as the CHARMM force fields C22* (14) with TIP3P-CHARMM (34) and C36m (6) with TIP3P-CHARMM. This set of force fields allowed for the comparison of the accuracy of two "helix-coil balanced" force fields optimized for use with three-point water models (a99SB-ILDN and C22*), the recently optimized C36m force field, and three force fields that use four-point water models. Comparisons of simulations with experimental measurements are represented by a normalized force-field score (*Methods*), where a normalized force-field score of one indicates that a simulation with a given force field produces the closest agreement with experiment among all of the force fields tested for all classes of the experimental observables considered. We discuss the most salient results from simulations with these six force fields in the text; a detailed comparison between calculated and experimental properties for each force field for each member of the benchmark set of proteins is reported in *SI Appendix*, Tables S2–S17. Results for simulations run using a99SB-UCB are reported in *SI Appendix*.

Folded proteins. To assess the performance of the force fields on folded proteins, we first simulated four proteins for which extensive NMR data are available; 10-μs simulations of the folded proteins ubiquitin, GB3, HEWL, and BPTI performed with a99SB*-ILDN, C36m, C22*, and a99SB-ILDN/TIP4P-D rather accurately reproduced the experimental NMR measurements (Fig. 2 and *SI Appendix*, Fig. S1 and Tables S3–S6 and S18). Simulations of folded proteins run with the a03ws and a99SB-UCB force fields gave substantially larger deviations from the experimental results. Analysis of the trajectories reveals that, in many cases, partial or complete unfolding of the proteins was responsible for these deviations. We also examined the stability of 11 additional folded proteins from the set by Huang et al. (6) in 20-μs simulations in each force field (*SI Appendix*, Table S20) and the agreement with NMR chemical shift and NOE measurements for 41 additional folded proteins taken from the set by Mao et al. (35) in 10-μs simulations in each force field (*SI Appendix*, Fig. S15 and Tables S21 and S26). We found that a99SB*-ILDN/TIP3P yielded both the most stable simulations of the proteins in the set by Huang et al. (6) and the closest agreement with experimental chemical shifts and NOEs in the dataset by Mao et al. (35). Simulations run with C36m and C22* were somewhat less stable and had a higher average fraction of NOE violations. Simulations run with a99SB-ILDN/TIP4P-D and a03ws destabilized a larger fraction of proteins and produced the largest average fraction of NOE violations. We found that the trends in stability and agreement with NMR measurements of these additional 52 proteins were generally consistent with conclusions drawn from the comparison of simulations of HEWL, BPTI, GB3, and ubiquitin with more extensive sets of experimental NMR data, with the exception that a99SB-ILDN/TIP4P-D partially unfolded some of these 52 additional proteins.

Disordered and partially disordered proteins. We next examined the accuracy of the force fields for simulating disordered proteins. In simulations of the disordered proteins drkN SH3, ACTR, N_{TAIL}, α-synuclein, Aβ40, PaaA2, p15^{PAF}, Sic1, and Ash1 run with a99SB*-ILDN, we observed a systematic underestimation of the average Rg values compared with experimental values, with proteins adopting compact molten globule-like structures. In the 30-μs timescale examined here, sampling in simulations run with a99SB*-ILDN tended to be restricted to conformations with very similar topologies and contacts to the first structures sampled after an initial hydrophobic collapse. As a result, the secondary structure propensities observed in these simulations were largely dictated by the conformation of the initial collapsed structure and were less dependent on the intrinsic secondary structure propensity of the local sequence. This interpretation is supported by the observation that an a99SB*-ILDN simulation of the
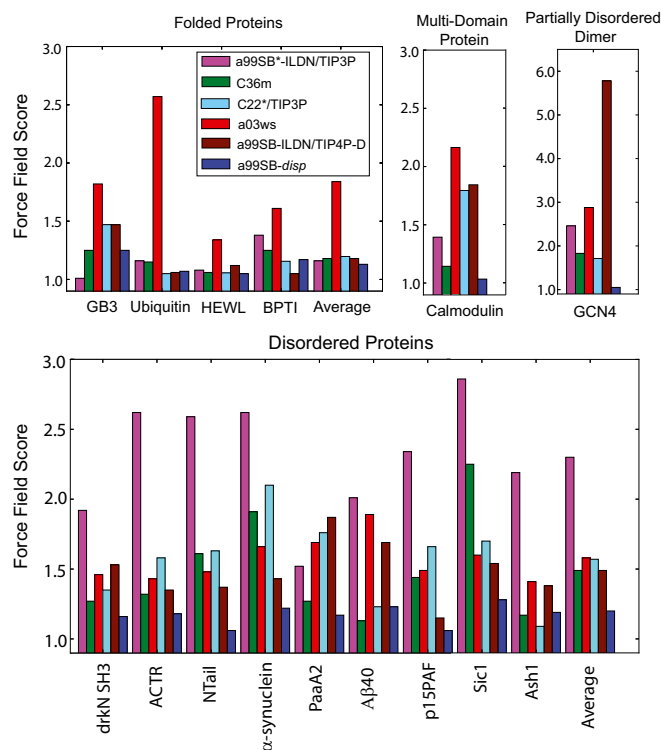

Fig. 2. Normalized force-field scores for simulations of folded and disordered proteins from the benchmark examined in this work. Average scores are also shown for calmodulin, which contains two folded globular domains connected by a flexible linker, and GCN4, a partially disordered dimer that contains an ordered coiled coil dimer interface. We note that simulations of GB3, ubiquitin, drkN SH3, ACTR, N_{TAIL}, and α-synuclein were used in the training of the parameters of a99SB-*disp*. Statistical uncertainties in the force-field score were estimated to be ~0.1 for individual proteins (*SI Appendix*, Fig. S14).

N_{TAIL} molecular recognition (MoRE) element, a truncated 31-residue N_{TAIL} construct that is too small to experience a restrictive hydrophobic collapse, showed helical propensities in excellent agreement with experimental measurements, whereas in the simulation of the entire N_{TAIL} domain, the MoRE element had restricted conformational flexibility because of the overly compact structures sampled and did not sample any helical conformations (*SI Appendix*, Fig. S2). Overcollapse generally resulted in persistent secondary structure forming in simulations where none is experimentally observed and overall poor agreement with experimental secondary structure propensities. One exception was the simulation of PaaA2, which was initiated from a conformation containing two helices in the correct locations. The helices remained intact in the initial collapsed structure and were stable throughout the simulation.

Disordered protein simulations run with C22* and C36m showed less restricted sampling and featured larger Rg fluctuations, larger average Rg values, and more frequent rearrangements of the chain topology than simulations run with a99SB*-ILDN, but simulations in both CHARMM force fields still substantially underestimated the Rg of larger (>60 residues) disordered proteins, producing ensembles that are not consistent with experimental data (*SI Appendix*, Fig. S3). Consistent with the less restricted sampling, C22* and C36m showed better agreement with experimental secondary structure propensities and NMR chemical shifts than a99SB*-ILDN for the majority of the proteins examined here (Figs. 2 and 3 and *SI Appendix*, Tables S7–S16) but did not capture residual helical propensities in drkN SH3, N_{TAIL}, and GCN4 (*SI Appendix*, Figs. S5, S11, and

S12); both force fields also contained some elevated β propensities in drkN SH3 and α-synuclein that were not in agreement with experimental chemical shifts and RDCs (*SI Appendix*, *Figs. S5 and S7 and Tables S7 and S10*). [We note that, in the C36m simulation of PaaA2, the stable helices seem to be the result of stabilization of initial helical conformations from an unphysical hydrophobic collapse (*SI Appendix*, *Fig. S3*).] Notably, of all of the force fields examined in this study, C36m produced the best agreement with the NMR measurements of Aβ40, a relatively compact disordered protein that is too short to experience a restrictive hydrophobic collapse in simulation. C36m resulted in slightly more expanded disordered-state ensembles compared with C22* (*SI Appendix*, *Fig. S3 and Tables S7–S16*), which uses the same water model, possibly as a result of the changes introduced to the vdW terms of alkanes in C36m. This result suggests that it may be interesting in future studies to further explore changes to alkane vdW parameters and water model parameters and their effect on disordered protein compactness and the hydrophobic effect.

Simulations of disordered proteins run with a99SB-ILDN/TIP4P-D, a99SB-UCB, and a03ws had Rg values much closer to experimental measurements. The percentage deviations of the average Rg from the experimental estimate for the disordered proteins in the benchmark set were 10, 13, and 9% for a99SB-ILDN/TIP4P-D, a99SB-UCB, and a03ws, respectively, compared with 22, 26, and 36% for C36m, C22*, and a99SB*-ILDN, respectively.
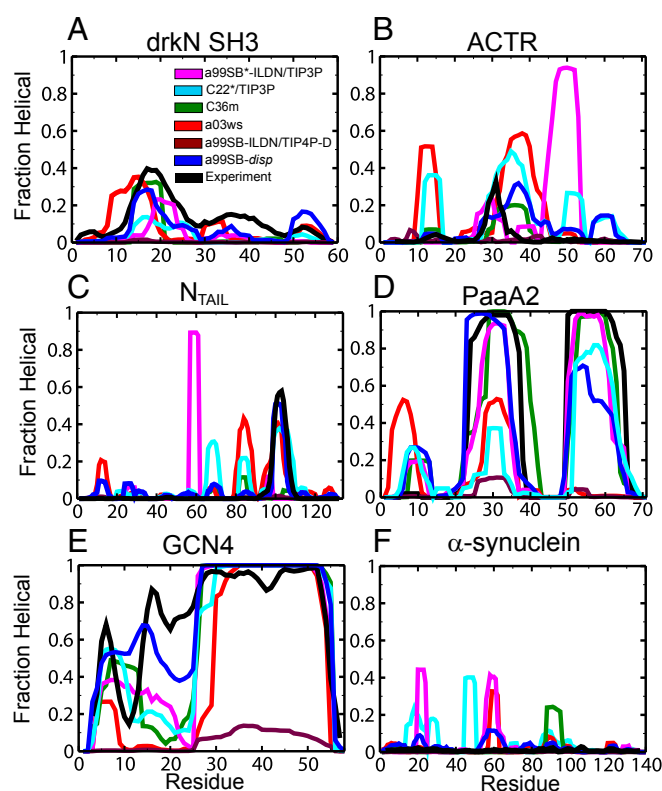

Fig. 3. Helical propensities observed in simulations of the disordered proteins drkN SH3 (*A*), ACTR (*B*), N$_{TAIL}$ (*C*), PaaA2 (*D*), GCN4 (*E*), and α-synuclein (*F*). Black lines are experimental estimates from restrained ensemble models calculated from NMR data [drkN SH3 (65), N$_{TAIL}$ (66), PaaA2 (26)] or predicted from experimental NMR chemical shifts using the program δ2d (67) (ACTR, α-synuclein, GCN4). We note that simulations of drkN SH3, ACTR, N$_{TAIL}$, and α-synuclein were used in the training of the parameters of a99SB-*disp*. For clarity of presentation, error bars have been omitted but are displayed in *SI Appendix*, *Figs. S5–S7 and S10–S12*.

Simulations run with a99SB-ILDN/TIP4P-D showed no helical propensity for any regions of the proteins in the benchmark set, and the helical coiled coil interface of the dimeric protein GCN4 was unstable and dissociated into unstructured monomers. These results suggest that the TIP4P-D water model in combination with the a99SB-ILDN force field strongly destabilizes helical conformations. In contrast, the simulated ensembles of Aβ40 and α-synuclein, which contain little or no secondary structure, showed good agreement with experimental NMR measurements (*SI Appendix*, *Tables S10 and S12*).

Simulations run with a99SB-UCB also substantially underestimated residual helicity in all of the proteins tested, although regions of drkN SH3 and N$_{TAIL}$ with stable experimental helices showed small amounts of helical propensity in simulation (*SI Appendix*, *Figs. S5–S12*). The GCN4 dimer also dissociated into unstructured monomers when simulated with a99SB-UCB. These results suggest the a99SB-UCB vdW overrides also strongly destabilize helical conformations. Simulations of Aβ40 and α-synuclein using a99SB-UCB produced excellent agreement with experimental NMR measurements, surpassing the agreement of simulations run with a99SB-ILDN/TIP4P-D.

Simulations run with a03ws had substantially more residual helicity than a99SB-ILDN/TIP4P-D and a99SB-UCB. In the a03ws simulations, several experimentally observed helices were populated, although several other regions showed helical propensity where it was not observed experimentally or lacked helical propensity where stable helices were detected in experiment (Fig. 3 and *SI Appendix*, *Figs. S5–S12*). We note that, even in the more expanded ensembles of a03ws, we still observed some contact-based secondary structure stabilization, although it tended to be with closely neighboring regions, as long-range contacts were much more transient in these ensembles. These results suggest that, although the relative stabilities of helix, sheet, and coil in a03ws are in more reasonable agreement with experiment than in a99SB-ILDN/TIP4P-D and a99SB-UCB, the relative stabilities of the secondary structure elements for different amino acids may require further tuning (36).

Ala5. For testing force-field performance on small peptides, we performed 500-ns simulations of Ala$_5$ with each force field and computed scalar couplings. In *SI Appendix*, *Table S2*, we report χ^2 values (*SI Appendix*, *Eq. S1*) for each force field, taking into account estimates of the errors produced by uncertainty in the Karplus equation coefficients (8). All force fields investigated here were parameterized against this NMR dataset or similar data and, thus, reproduced the experimental scalar couplings reasonably well. Simulations run with C36m, a03ws, a99SB-UCB, and a99SB-ILDN/TIP4P-D had χ^2 values < 1, which suggest agreement with experiment within the error of the Karplus equation predictions.

Fast-folding proteins and peptides. We performed simulated tempering runs of the two short peptides (AAQAA)$_3$ and CLN025 (Fig. 4), which have been widely used as force-field benchmarks due to their ability to form helical or β structure. Consistent with previous studies (2, 8), all force fields considered here considerably underestimated the cooperativity of both hairpin and helix formation. C22* best captured the helical propensity of (AAQAA)$_3$ at 300 K; a99SB*-ILDN, C36m, and a03ws performed similarly on (AAQAA)$_3$, showing helical propensities of 5–12% with relatively little temperature dependence. We observed no helicity in (AAQAA)$_3$ simulations run with a99SB-ILDN/TIP4P-D and a99SB-UCB. a99SB*-ILDN and C22* showed the closest agreement with the melting curve of CLN025. All other force fields underestimated the stability of the native CLN025 hairpin at 300 K to different extents.

In simulated tempering simulations of the fast-folding proteins Trp-cage, GTT, and villin, we found that simulations run using a99SB*-ILDN showed the closest agreement with the experimental melting curves, while overestimating the melting temperatures by
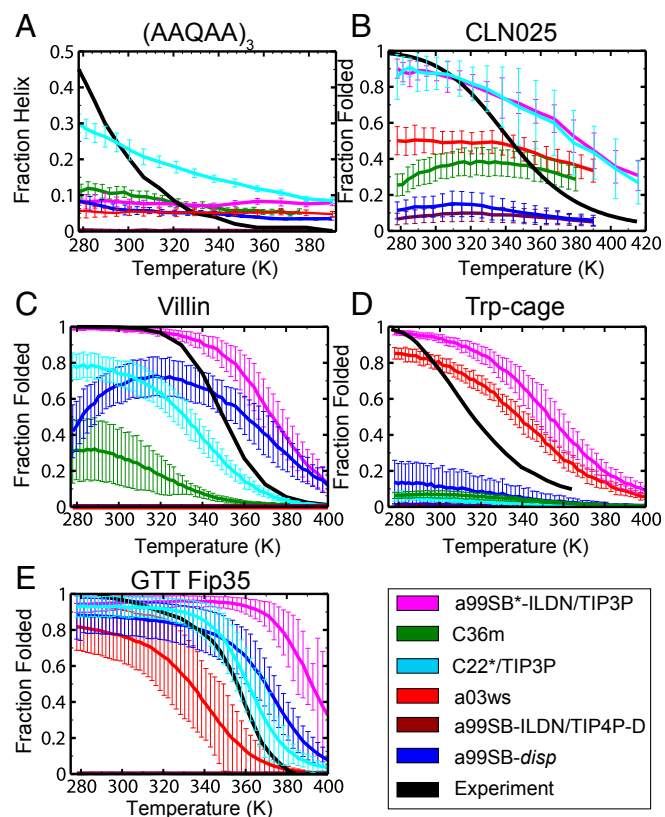
Fig. 4. Stability of the weakly structured peptides (AAQAA)$_3$ (*A*) and CLN025 (*B*) and the fast-folding proteins villin (*C*), Trp-cage (*D*), and GTT Fip35 (*E*) from simulated tempering simulations. Experimental melting curves are shown in black. We note that simulations of (AAQAA)$_3$ were used in the training of the parameters of a99SB-*disp*. No folded structures were observed in simulations of (AAQAA)$_3$, villin, Trp-cage, or GTT Fip35 run with a99SB-ILDN/TIP4P-D, and there were no folded structures observed in simulations of villin run with a03ws.

10–50 K (Fig. 4). In simulations run with C36m, the stabilities of villin and Trp-cage were underestimated, and no folded structures of GTT were observed. In simulations run with a03ws, we observed reasonable agreement with the experimental melting curve of Trp-cage and a moderate underestimation of the stability and melting temperature of GTT. No stable folded structures were observed in the a03ws simulation of villin, and a simulation of the native state of villin unfolded after 0.4 μs in a 300-K simulation. In simulated tempering simulations run with a99SB-ILDN/TIP4P-D and a99SB-UCB, we did not observe any folded species for any of the fast-folding proteins examined here. Simulations of the native folded structures of Trp-cage and villin run at 300 K confirmed that these structures are not stable on the microsecond timescale in these force fields. Simulations of the native state of GTT were stable at 300 K for 10 μs in C36m and a99SB-UCB, and we previously observed reversible folding of GTT in a99SB-ILDN/TIP4P-D at 395 K (7), suggesting that the absence of folded states in the 300-μs simulated tempering runs is likely the result of unconverged sampling.

Calmodulin. Experimental measurements for Ca^{2+}-bound calmodulin indicate that it consists of two stable globular domains connected by a flexible linker (15, 37–39). Ca^{2+}-bound calmodulin simulations were initiated from the "dumbbell"-shaped crystal structure (40), in which the dynamic linker is in an entirely helical conformation.

In simulations run with a99SB*-ILDN in TIP3P, the linker quickly frayed and formed interactions with the C-terminal tail of

the protein that were stable on the 30-μs timescale observed here. These interactions fortuitously stabilized a static domain orientation with an *Rg* in reasonable agreement with experiment. In the simulation run with C22*, the two domains collapsed together and then progressively unfolded throughout the remainder of the simulation. In the a03ws simulation, the N-terminal domain became destabilized and largely unfolded after 2 μs, while the C-terminal domain remained structured. In the a99SB-ILDN simulations with TIP4P-D, the linker was highly flexible and dynamic, and the two domains sampled a large number of orientations with an average *Rg* in excellent agreement with experiment, but the helical interfaces within the globular domains became somewhat destabilized. Calmodulin simulations run with a99SB-UCB were the least stable, with the N-terminal domain unfolding after 0.5 μs and the C-terminal domain unfolding after 5 μs, resulting in the poorest agreement with the experimental measurements. Calmodulin simulations in C36m showed flexibility in the linker domain, sampled several orientations of the two domains, and were in excellent agreement with experimental measurements.

Summary of force-field benchmark testing. In our benchmark testing, several of the force fields performed well in simulations of folded proteins, but none of the force fields produced accurate dimensions and residual secondary structure propensities across the set of disordered proteins while attaining state-of-the-art performance for folded proteins. a99SB*-ILDN, for example, performed well for simulations of folded proteins, small disordered peptides, and fast-folding proteins but produced unrealistic dimensions and poor agreement with residual secondary structure propensities for disordered proteins. Simulations run with C22* and C36m performed well for folded proteins and showed decent agreement with experimental measurements for small disordered proteins (<60 residues). Small peptides and fast-folding proteins, however, were understabilized in C36m. Simulations run with C22* and C36m also produced overly collapsed ensembles of longer disordered proteins and showed discrepancies in residual secondary structure propensities of some disordered proteins.

Simulations run with force fields optimized to prevent the overcollapse of disordered states produced more realistic dimensions for disordered proteins but often at the expense of the accuracy of descriptions of residual secondary structure propensity and/or the stability of folded proteins. Simulations run with a03ws, for example, accurately described the residual secondary populations of small peptides and the stability of some of the fast-folding proteins, but they often resulted in lower stability and degraded performance for folded proteins and inaccurate residual secondary structure content in disordered proteins.

Simulations of disordered proteins without residual secondary structure performed with a99SB-UCB were in good agreement with experimental measurements, but simulations of folded proteins were unstable, and the stability of residual secondary structure propensities was substantially underestimated in disordered proteins and small peptides. Simulations run with a99SB-ILDN/TIP4P-D performed well for folded proteins and also provided accurate descriptions of disordered protein regions with no residual secondary structure. The stability of secondary structure elements in small peptides and disordered proteins was severely underestimated, however, and the fast-folding proteins and small peptides were unstable in this force field.

Optimization of a99SB-*disp*. We next asked if the difficulty in consistently obtaining accurate results for both ordered and disordered proteins reflects an intrinsic limitation in the force-field functional forms or whether substantial improvements are possible through parameter optimization alone. As a starting point, we chose the a99SB-ILDN/TIP4P-D force field and

attempted to modify its parameters to improve its performance for both ordered and disordered proteins.

Inspired by previous successful efforts to reparameterize force-field torsion angles to obtain a more accurate balance between helix and coil states (10, 11, 14), we performed a similar torsion optimization targeting $(AAQAA)_3$ fraction helicity and polyalanine scalar couplings as described previously (14). At improved levels of helicity, we observed previously described (36) discrepancies in the helical propensities of charged residues, and we thus incorporated the corrections of the a99SB*-ILDN-Q force field (36). We found that, through torsion optimizations, it was possible to produce good agreement with the temperature-dependent helicity of $(AAQAA)_3$ and polyalanine scalar couplings ($\chi^2 = 0.94$). Simulations of disordered proteins performed with this torsion-optimized force field, however, produced ensembles that were too helical compared with experimental measurements (*SI Appendix*, Fig. S4) and, in the case of ACTR, induced a hydrophobic collapse. There seemed to be some cooperativity between helix formation and collapse in disordered proteins, as we observed that torsion parameters that accurately described helical propensities in small peptides, such as $(AAQAA)_3$, and small disordered peptides, such as N_{TAIL} MoRE, did not produce accurate simulations of partially helical disordered proteins that were large enough to experience hydrophobic collapse. This force field also substantially degraded the accuracy of simulations of GB3 and ubiquitin by destabilizing the packing of β sheets.

These results suggest that it may be difficult to accurately describe helical propensities, the dimensions of disordered proteins, and the stability of native states in a99SB*-ILDN-Q/TIP4PD using torsion optimization alone. To overcome this difficulty, we thus also tested modifications in the strength of the C6 dispersion term in our water model. In an attempt to alleviate the overcollapse of helical disordered proteins, we optimized a water model with a slightly stronger C6 dispersion term than that of TIP4P-D (960 kcal mol^{-1} Å^{-6} in our water model as opposed to 900 kcal mol^{-1} Å^{-6} in TIP4P-D) as described previously (7). We compared the liquid water properties of this water model, which we refer to as a99SB-*disp* water, with those of TIP4P-D in *SI Appendix*, Table S22; we found most properties to be very similar, although for some, such as the diffusion coefficient, slightly worse agreement with experiment was observed with a99SB-*disp* water. We also compared the solvation free energies of protein side-chain analogs in TIP3P, TIP4P-D, and a99SB-*disp* water (*SI Appendix*, Fig. S13) and found them to be very similar. We found that a99SB-*disp* water successfully reduced the occurrence of hydrophobic collapse of disordered helical states but also destabilized folded proteins and helical conformations in $(AAQAA)_3$, drkN SH3, ACTR, and N_{TAIL}.

Inspired by previous work (9), we then attempted to increase the stability of helical states and folded proteins by introducing modifications to the O-H LJ pair between backbone carbonyl oxygens and backbone amide hydrogens, which strengthened protein backbone hydrogen bonds. To find reasonable combinations of the carbonyl-oxygen and backbone amide hydrogen O-H LJ pair and backbone torsion adjustments while reducing the risk of overfitting, we optimized these parameters against a "training" subset of the benchmark set introduced above: ubiquitin, GB3, $(AAQAA)_3$, Ala$_5$, N_{TAIL} MoRE, N_{TAIL}, drkN SH3, ACTR, and α-synuclein. We also ran constant temperature folding simulations of villin and GTT near their melting temperatures to ensure that they could reversibly fold and unfold. In our search of parameter space, we found that we were unable to simultaneously reproduce the helicity of both $(AAQAA)_3$ and helical disordered proteins with high accuracy and ultimately accepted worse agreement with $(AAQAA)_3$ helicity in favor of more accurate descriptions of N_{TAIL} MoRE, N_{TAIL}, drkN SH3, ACTR, and α-synuclein.

Through iterative adjustments of the backbone torsion potential and of the strength of the LJ modification for carbonyl oxygen and amide hydrogen pairs, we were able to produce a force field, which we term a99SB-*disp*, that performed reasonably well across our training set of proteins. In addition to these modifications, a99SB-*disp* includes a series of side-chain torsion modifications targeting Protein Data Bank (PDB) rotamer distributions and quantum mechanical (QM) energy scans; a reparameterization of the side-chain charges of aspartate, glutamate, and arginine residues to match the guanidinium acetate association constant (14); and a reparameterization of glycine backbone torsion angles targeting a PDB coil library distribution (41, 42). The final parameters and further information regarding the parameterization of a99SB-*disp* are contained in *SI Appendix* (*SI Appendix*, Tables S22–S25).

In the training set, a99SB-*disp* performed comparably with the best-performing force fields for the folded proteins GB3 and ubiquitin (Fig. 2 and *SI Appendix*, Tables S3, S4, and S15), and $(AAQAA)_3$ helicity (Fig. 4) was comparable with a99SB*-ILDN, a03ws, and C36m. Simulations of N_{TAIL}, drkN SH3, ACTR, and α-synuclein were in good agreement with experiment (Figs. 2 and 3 and *SI Appendix*, Tables S7–S10 and S15), containing a reasonable amount of residual helicity in the correct regions of the protein sequences (Fig. 3). A similar level of accuracy was obtained for the simulations of the test set of proteins not used in the parameter optimization: HEWL, BPTI, Aβ40, PaaA2, p15^{PAF}, Sic1, Ash1, GCN4, and calmodulin (Figs. 2 and 3 and *SI Appendix*, Tables S5, S6, and S11–S17), suggesting that the parameters obtained are reasonably transferable and that the level of accuracy obtained was not the result of overfitting on the training set of proteins. Simulations run with a99SB-*disp* also produced the best agreement with experiment among all force fields tested on an additional 52 folded proteins from the test sets by Huang et al. (6) and Mao et al. (35) (*SI Appendix*, Fig. S15 and Tables S20, S21, and S26), none of which were used in parameterization, providing further evidence for the transferability of the parameters.

Importantly, in our benchmark set, which includes nine disordered proteins with lengths ranging from 40 to 140 residues and experimental *Rg* values ranging from 12 to 32 Å, simulations of disordered proteins run with a99SB-*disp* had the closest agreement with experimental *Rg* measurements of all force fields tested, with an average deviation of only 6% from experimental values. In particular, for all disordered proteins with more than 60 residues, simulations run with a99SB-*disp* produced ensembles that were substantially more expanded, in much closer agreement with experiment, than those in simulations run with the next best force field, C36m (27% deviation from experimental *Rg* values) (*SI Appendix*, Fig. S3). This difference is most pronounced in simulations of larger proteins with more hydrophobic sequences (α-synuclein, N_{TAIL}, and Sic1). In simulations of shorter, more compact disordered proteins, such as Aβ40 and drkN SH3, and in simulations of the highly charged disordered protein Ash1 (net charge of −15 at pH 7), both force fields produced *Rg* values in good agreement with experiment. On average, simulations of disordered proteins run with a99SB-*disp* were also in substantially better agreement with experimental NMR measurements than were simulations run with C36m, with similar levels of improvement observed in simulations of disordered proteins in both the training and test sets.

The a99SB-*disp* melting curves for villin, Trp-cage, and GTT, proteins that were not part of the training set, were also in much better agreement with experiment than the a99SB-ILDN/TIP4P-D melting curves, which showed no folded populations at any temperature for these proteins. There was no noticeable improvement in the folded population of CLN025 compared with simulations run with a99SB-ILDN/TIP4P-D (Fig. 4). We see some evidence of cold denaturation in the a99SB-*disp* melting

curve of villin, which is likely attributable to a subtle shift in the folding enthalpy and heat capacity induced by the water model (*SI Appendix*, Fig. S16).

Discussion

We have assessed six protein force fields commonly used in MD simulation and found that, although the force fields tested produced results in good agreement with experiment in many cases, simulations of the disordered proteins in our benchmark revealed limitations in each of the force fields. We have proposed a force field, a99SB-*disp*, with improved parameters that were trained on and tested against separate subsets of the benchmark; this force field advances the state of the art for accuracy for simulations of disordered proteins, while achieving accuracy comparable with the best force fields for folded proteins.

The transferability of a99SB-*disp* across the benchmark examined here suggests that it should be suitable for studying a number of systems that are not well-described by existing force fields. The ability of a99SB-*disp* to describe both ordered and disordered states should enable accurate simulations of proteins with both ordered and disordered domains as well as simulations of transitions between disordered and ordered states, such as those observed in the coupled folding-upon-binding of intrinsically disordered proteins with their binding partners. Further studies that examine the performance of a99SB-*disp* in simulations of a wider variety of disordered proteins and explore its performance from the perspective of polymer physics (43) should also be of considerable interest.

It is notable that the transferability of a99SB-*disp* was achieved within the constraints of the approximate functional forms used in current fixed charge force fields. The parameters of a99SB-*disp* are the result of introducing modest changes to an existing force field to enable the accurate description of both ordered and disordered proteins. We found that we were able to achieve this goal by modifying the water model and iteratively testing small changes in backbone torsion corrections and the strength of a backbone O-H LJ pair. We believe that the demonstration that the simplified functional form of a nonpolarizable force field is sufficient to describe folded proteins and a wide range of disordered and partially disordered systems provides a noteworthy proof of principle and that the accuracy achieved in a99SB-*disp* simulations of folded proteins, disordered proteins, fast-folding proteins, and multidomain proteins suggests that this force field could be useful for the accurate simulation of a wide variety of systems that present difficulties for existing force fields.

Due to the computational cost of obtaining sufficiently converged simulations of the proteins in our training set, our search for a set of optimal parameters was not exhaustive. It is thus possible that the performance of a99SB-*disp* could benefit from further optimization. In particular, it is likely that modifications not explored in this study, such as more extensive changes to the nonbonded parameters, could further improve a99SB-*disp* (and other fixed charge force fields). Joint optimizations of alkane vdW terms (such as those introduced in C36m) and water model parameters, for example, may better capture the physics of the hydrophobic effect and further improve the ability of fixed charge models to balance the stability of small peptides and the dimensions of disordered proteins. The benchmark set of proteins described here should provide a valuable tool in future efforts to develop force fields that accurately describe a broad range of disordered systems.

Methods

MD Simulations. Details of the MD simulation setup for each of the systems studied in this work can be found in *SI Appendix*, Table S16. All systems were simulated using the following force fields: a99SB*-ILDN (11, 12) with TIP3P (13), C22* (14) with TIP3P-CHARMM (34), C36m (6), a03ws (containing

modified TIP4P/2005 interactions) (8), a99SB and TIP4P-Ew (32) with the Head-Gordon vdW (9) and dihedral (33) modifications (termed a99SB-UCB), a99SB-ILDN (12) with TIP4P-D (7), and a99SB-*disp*. (The parameters for the a99SB-*disp* force field are listed in *SI Appendix*.) Systems were initially equilibrated at 300 K and 1 bar for 1 ns using the Desmond software (44). Production runs at 300 K were performed in the NPT ensemble (45–47) with Anton specialized hardware (48) using a 2.5-fs time step and a 1:2 RESPA scheme (49). Bonds involving hydrogen atoms were restrained to their equilibrium lengths using the M-SHAKE algorithm (50). Non-bonded interactions were truncated at 12 Å, and the Gaussian split Ewald method (51) with a $32 \times 32 \times 32$ mesh was used for the electrostatic interactions. All simulations were run at 300 K, with the exception of $(AAQAA)_3$, CLN025, and the fast-folding proteins Trp-cage, villin, and GTT, which used simulated tempering (52) to improve sampling. In simulated tempering simulations of $(AAQAA)_3$ and CLN025, 20 rungs were spaced geometrically spanning 278–390 K. In simulated tempering simulations of Trp-cage, villin, and GTT, 60 rungs were spaced geometrically spanning 278–400 K.

Calculation of Experimental Observables. Backbone scalar coupling constants were calculated using published Karplus relationships (53) for $^3J_{HNH\alpha}$, $^3J_{HNC'}$, $^3J_{HNC\beta}$ (54), $^3J_{H\alpha C'}$ (55), and $^3J_{C'C'}$ (56). Side-chain scalar coupling constants were calculated using published Karplus relationships for $^3J_{H\alpha H\beta}$, $^3J_{C'H\beta}$, $^3J_{C'C\gamma}$, and $^3J_{NC\gamma}$ (57), with the exception of $^3J_{C'C\gamma}$ and $^3J_{NC\gamma}$ values for Ile, Thr, and Val, which were computed using Karplus parameters from the work by Chou et al. (58). Through-hydrogen bond $^{3H}J_{NC'}$ scalar coupling constants were calculated according to the work by Barfield (59). RDCs of folded proteins were calculated as reported previously (60). RDCs of disordered proteins were calculated using PALES (61) using a local alignment window of 15 residues. Backbone amide and methyl S^2 order parameters were calculated from the value of the internal autocorrelation functions of the relevant bond vectors at lag times corresponding to the experimentally determined rotational correlation times as described previously (62). Internal autocorrelation functions were calculated after aligning trajectories to the backbone atoms of the simulation starting structures for ubiquitin, GB3, and HEWL and to backbone atoms of the stable leucine zipper coiled coil dimer interface for GCN4 (63). NMR chemical shifts were calculated using Sparta+ (64). PREs were calculated as described previously (7).

Calculation of Normalized Force-Field Scores. To compare the relative accuracy of each force field, we report normalized force-field scores. For folded proteins, the rmsd from each class of experimental data, such as side-chain scalar couplings, is normalized by the smallest observed rmsd among the seven force fields examined here. The normalized force-field score is determined by taking the average of the normalized rmsds over all classes of experimental measurements (the classes used for a specific protein are given in the first columns of *SI Appendix*, Tables S3–S6; note that a class may include multiple datasets listed in *SI Appendix*, Table S1):

$$\text{Folded Protein } FF_{Score} = \frac{1}{N} \sum_{i=1}^{N} \frac{FF_{rmsd}}{rmsd_{Norm}},$$

where N is the number of classes of experimental data considered, FF_{rmsd} is the rmsd of the simulated values from the corresponding experimental values for class i, and $rmsd_{Norm}$ is the smallest observed rmsd of all of the seven force fields examined in this study for class i. In this metric, a normalized FF_{Score} of one indicates that a force field produces the closest agreement with experiment among all of the force fields tested for all of the classes of experimental observables considered.

For disordered proteins, GCN4, and calmodulin, force-field scores are computed as a combination of a backbone NMR chemical shift score (CS_{Score}), a score based on additional NMR measurements (NMR_{Score}), and an Rg deviation penalty ($Rg_{Penalty}$). The CS_{Score} is determined analogously to the folded protein score by normalizing the rmsd for each class of chemical shift type (the classes are listed in *SI Appendix*, Tables S7–S17) (for drkN SH3, for example, the classes are C_α, H_α, H_N, C', and C_β) by the smallest rmsd observed for the seven force fields and taking an average of the normalized rmsds over all sets of experimental chemical shifts. The NMR score is computed analogously for all additional classes of NMR measurements. The $Rg_{Penalty}$ is zero if the average simulated Rg is within the experimentally estimated error ($Rg_{Exp\ error}$). For deviations larger than the estimated experimental error, the $Rg_{Penalty}$ is calculated as

$$Rg_{Penalty} = \frac{|Rg_{Exp} - Rg_{Sim}| - Rg_{Exp\ error}}{Rg_{Exp}}.$$

The combined disordered protein force-field score is computed as

$$Disordered\ Protein\ FF_{Score} = \frac{CS_{Score} + NMR_{Score}}{2} + Rg_{Penalty}.$$

We find it helpful to summarize the accuracy of each force field in this way as a single number. Clearly, however, the details of the definition of the score are, to some extent, arbitrary. To facilitate examination of alternative scores, we have included in *SI Appendix* the deviation of the simulated values from

experiment for each of the experimental measurements considered here for each force field. To provide a measure of the sensitivity of the calculated force-field scores to the initial simulation conditions on the timescales examined in this study, we repeated simulations of the folded and disordered proteins examined in this study using the a99SB-*disp* force field with a different set of randomized initial velocities. We compare the resulting force-field scores with those obtained from the previous set of simulations in *SI Appendix*, Fig. S14.

1. Lange OF, van der Spoel D, de Groot BL (2010) Scrutinizing molecular mechanics force fields on the submicrosecond timescale with NMR data. *Biophys J* 99:647–655.
2. Lindorff-Larsen K, et al. (2012) Systematic validation of protein force fields against experimental data. *PLoS One* 7:e32131.
3. Beauchamp KA, Lin YS, Das R, Pande VS (2012) Are protein force fields getting better? A systematic benchmark on 524 diverse NMR measurements. *J Chem Theory Comput* 8:1409–1414.
4. Lindorff-Larsen K, Piana S, Dror RO, Shaw DE (2011) How fast-folding proteins fold. *Science* 334:517–520.
5. Mittal J, Best RB (2010) Tackling force-field bias in protein folding simulations: Folding of Villin HP35 and Pin WW domains in explicit water. *Biophys J* 99:L26–L28.
6. Huang J, et al. (2017) CHARMM36m: An improved force field for folded and intrinsically disordered proteins. *Nat Methods* 14:71–73.
7. Piana S, Donchev AG, Robustelli P, Shaw DE (2015) Water dispersion interactions strongly influence simulated structural properties of disordered protein states. *J Phys Chem B* 119:5113–5123.
8. Best RB, Zheng W, Mittal J (2014) Balanced protein–water interactions improve properties of disordered proteins and non-specific protein association. *J Chem Theory Comput* 10:5113–5124.
9. Nerenberg PS, Jo B, So C, Tripathy A, Head-Gordon T (2012) Optimizing solute-water van der Waals interactions to reproduce solvation free energies. *J Phys Chem B* 116:4524–4534.
10. Best RB, Mittal J (2010) Protein simulations with an optimized water model: Cooperative helix formation and temperature-induced unfolded state collapse. *J Phys Chem B* 114:14916–14923.
11. Best RB, Hummer G (2009) Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. *J Phys Chem B* 113:9004–9015.
12. Lindorff-Larsen K, et al. (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* 78:1950–1958.
13. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935.
14. Piana S, Lindorff-Larsen K, Shaw DE (2011) How robust are protein folding simulations with respect to force field parameterization? *Biophys J* 100:L47–L49.
15. Chou JJ, Li S, Klee CB, Bax A (2001) Solution structure of Ca(2+)-calmodulin reveals flexible hand-like properties of its domains. *Nat Struct Biol* 8:990–997.
16. Kubelka J, Chiu TK, Davies DR, Eaton WA, Hofrichter J (2006) Sub-microsecond protein folding. *J Mol Biol* 359:546–553.
17. Neidigh JW, Fesinmeyer RM, Andersen NH (2002) Designing a 20-residue protein. *Nat Struct Biol* 9:425–430.
18. Piana S, et al. (2011) Computational design and experimental testing of the fastest-folding β-sheet protein. *J Mol Biol* 405:43–48.
19. Shalongo W, Dugad L, Stellwagen E (1994) Distribution of helicity within the model peptide acetyl(AAQAA)3amide. *J Am Chem Soc* 116:8288–8293.
20. Honda S, et al. (2008) Crystal structure of a ten-amino acid protein. *J Am Chem Soc* 130:15327–15331.
21. Iešmantavičius V, et al. (2013) Modulation of the intrinsic helix propensity of an intrinsically disordered protein reveals long-range helix-helix interactions. *J Am Chem Soc* 135:10155–10163.
22. Zhang O, Forman-Kay JD (1995) Structural characterization of folded and unfolded states of an SH3 domain in equilibrium in aqueous buffer. *Biochemistry* 34:6784–6794.
23. Bertoncini CW, et al. (2005) Release of long-range tertiary interactions potentiates aggregation of natively unstructured α-synuclein. *Proc Natl Acad Sci USA* 102:1430–1435.
24. Jensen MR, et al. (2011) Intrinsic disorder in measles virus nucleocapsids. *Proc Natl Acad Sci USA* 108:9839–9844.
25. Sgourakis NG, Yan Y, McCallum SA, Wang C, Garcia AE (2007) The Alzheimer's peptides Abeta40 and 42 adopt distinct conformations in water: A combined MD/NMR study. *J Mol Biol* 368:1448–1457.
26. Sterckx YG, et al. (2014) Small-angle X-ray scattering- and nuclear magnetic resonance-derived conformational ensemble of the highly flexible antitoxin PaaA2. *Structure* 22:854–865.
27. De Biasio A, et al. (2014) p15PAF is an intrinsically disordered protein with non-random structural preferences at sites of interaction with other proteins. *Biophys J* 106:865–874.
28. Mittag T, et al. (2010) Structure/function implications in a dynamic complex of the intrinsically disordered Sic1 with the Cdc4 subunit of an SCF ubiquitin ligase. *Structure* 18:494–506.
29. Martin EW, et al. (2016) Sequence determinants of the conformational properties of an intrinsically disordered protein prior to and upon multisite phosphorylation. *J Am Chem Soc* 138:15323–15335.
30. Bracken C, Carr PA, Cavanagh J, Palmer AG, 3rd (1999) Temperature dependence of intramolecular dynamics of the basic leucine zipper of GCN4: Implications for the entropy of association with DNA. *J Mol Biol* 285:2133–2146.
31. Graf J, Nguyen PH, Stock G, Schwalbe H (2007) Structure and dynamics of the homologous series of alanine peptides: A joint molecular dynamics/NMR study. *J Am Chem Soc* 129:1179–1189.
32. Horn HW, et al. (2004) Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J Chem Phys* 120:9665–9678.
33. Nerenberg PS, Head-Gordon T (2011) Optimizing protein–solvent force fields to reproduce intrinsic conformational preferences of model peptides. *J Chem Theory Comput* 7:1220–1230.
34. MacKerell AD, et al. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 102:3586–3616.
35. Mao B, Tejero R, Baker D, Montelione GT (2014) Protein NMR structures refined with Rosetta have higher accuracy relative to corresponding X-ray crystal structures. *J Am Chem Soc* 136:1893–1906.
36. Best RB, de Sancho D, Mittal J (2012) Residue-specific α-helix propensities from molecular simulation. *Biophys J* 102:1462–1467.
37. Bertini I, et al. (2004) Experimentally exploring the conformational space sampled by domain reorientation in calmodulin. *Proc Natl Acad Sci USA* 101:6841–6846.
38. Bertini I, et al. (2010) Conformational space of flexible biological macromolecules from average data. *J Am Chem Soc* 132:13553–13558.
39. Kukic P, Camilloni C, Cavalli A, Vendruscolo M (2014) Determination of the individual roles of the linker residues in the interdomain motions of calmodulin using NMR chemical shifts. *J Mol Biol* 426:1826–1838.
40. Chattopadhyaya R, Meador WE, Means AR, Quiocho FA (1992) Calmodulin structure refined at 1.7 A resolution. *J Mol Biol* 228:1177–1192.
41. Jiang F, Han W, Wu YD (2013) The intrinsic conformational features of amino acids from a protein coil library and their applications in force field development. *Phys Chem Chem Phys* 15:3413–3428.
42. Jiang F, Zhou C-Y, Wu Y-D (2014) Residue-specific force field based on the protein coil library. RSFF1: Modification of OPLS-AA/L. *J Phys Chem B* 118:6983–6998.
43. Mao AH, Crick SL, Vitalis A, Chicoine CL, Pappu RV (2010) Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. *Proc Natl Acad Sci USA* 107:8183–8188.
44. Bowers KJ, et al. (2006) Scalable algorithms for molecular dynamics simulations on commodity clusters. *Proceedings of the ACMIIEEE Conference on Supercomputing (SC06)* (IEEE, New York).
45. Nosé S (1984) A unified formulation of the constant temperature molecular dynamics methods. *J Chem Phys* 81:511–519.
46. Hoover WG (1985) Canonical dynamics: Equilibrium phase-space distributions. *Phys Rev A Gen Phys* 31:1695–1697.
47. Martyna GJ, Tobias DJ, Klein ML (1994) Constant pressure molecular dynamics algorithms. *J Chem Phys* 101:4177–4189.
48. Shaw DE, et al. (2009) Millisecond-scale molecular dynamics simulations on Anton. *Proceedings of the Conference on High Performance Computing, Networking, Storage and Analysis (SC09)* (ACM, New York).
49. Tuckerman M, Berne BJ, Martyna GJ (1992) Reversible multiple time scale molecular dynamics. *J Chem Phys* 97:1990–2001.
50. Kräutler V, Van Gunsteren WF, Hünenberger PH (2001) A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *J Comput Chem* 22:501–508.
51. Shan Y, Klepeis JL, Eastwood MP, Dror RO, Shaw DE (2005) Gaussian split Ewald: A fast Ewald mesh method for molecular simulation. *J Chem Phys* 122:54101.
52. Marinari E, Parisi G (1992) Simulated tempering: A new Monte Carlo scheme. *Europhys Lett* 19:451–458.
53. Karplus M (1959) Contact electron-spin coupling of nuclear magnetic moments. *J Chem Phys* 30:11–15.
54. Vögeli B, Ying J, Grishaev A, Bax A (2007) Limits on variations in protein backbone dynamics from precise measurements of scalar couplings. *J Am Chem Soc* 129:9377–9385.
55. Lindorff-Larsen K, Best RB, Vendruscolo M (2005) Interpreting dynamically-averaged scalar couplings in proteins. *J Biomol NMR* 32:273–280.
56. Li F, Lee JH, Grishaev A, Ying J, Bax A (2015) High accuracy of Karplus equations for relating three-bond J couplings to protein backbone torsion angles. *ChemPhysChem* 16:572–578.

57. Pérez C, Löhr F, Rüterjans H, Schmidt JM (2001) Self-consistent Karplus parametrization of 3J couplings depending on the polypeptide side-chain torsion chi1. *J Am Chem Soc* 123:7081–7093.

58. Chou JJ, Case DA, Bax A (2003) Insights into the mobility of methyl-bearing side chains in proteins from (3)J(CC) and (3)J(CN) couplings. *J Am Chem Soc* 125:8959–8966.

59. Barfield M (2002) Structural dependencies of interresidue scalar coupling (h3)J(NC′) and donor (1)H chemical shifts in the hydrogen bonding regions of proteins. *J Am Chem Soc* 124:4158–4168.

60. Lindorff-Larsen K, Best RB, Depristo MA, Dobson CM, Vendruscolo M (2005) Simultaneous determination of protein structure and dynamics. *Nature* 433:128–132.

61. Zweckstetter M, Bax A (2000) Prediction of sterically induced alignment in a dilute liquid crystalline phase: Aid to protein structure determination by NMR. *J Am Chem Soc* 122:3791–3792.

62. Trbovic N, Kim B, Friesner RA, Palmer AG, 3rd (2008) Structural analysis of protein dynamics by MD simulations and NMR spin-relaxation. *Proteins* 71:684–694.

63. Robustelli P, Trbovic N, Friesner RA, Palmer AG, 3rd (2013) Conformational dynamics of the partially disordered yeast transcription factor GCN4. *J Chem Theory Comput* 9:5190–5200.

64. Shen Y, Bax A (2010) SPARTA+: A modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network. *J Biomol NMR* 48:13–22.

65. Marsh JA, Forman-Kay JD (2009) Structure and disorder in an unfolded state under nondenaturing conditions from ensemble models consistent with a large number of experimental restraints. *J Mol Biol* 391:359–374.

66. Ozenne V, et al. (2012) Mapping the potential energy landscape of intrinsically disordered proteins at amino acid resolution. *J Am Chem Soc* 134:15138–15148.

67. Camilloni C, De Simone A, Vranken WF, Vendruscolo M (2012) Determination of secondary structure populations in disordered states of proteins using nuclear magnetic resonance chemical shifts. *Biochemistry* 51:2224–2231.

Robustelli et al.