

Model-Dependent Solvation of the K-18 Domain of the Intrinsically Disordered Protein Tau

Sthitadhi Maiti and Matthias Heyden*



Cite This: *J. Phys. Chem. B* 2023, 127, 7220–7230



Read Online

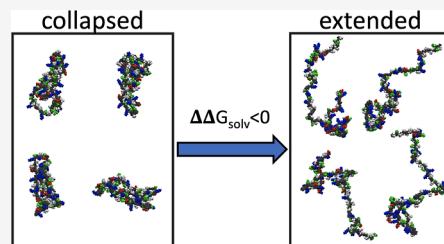
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: A known imbalance between intra-protein and protein–water interactions in many empirical force fields results in collapsed conformational ensembles of intrinsically disordered proteins in explicit solvent simulations that disagree with experiments. Multiple strategies have been introduced in the literature to modify protein–water interactions, which improve agreement between experiments and simulations. In this work, we combine simulations with standard and modified force fields with a spatially resolved analysis of solvation free energy contributions and compare the consequences of each strategy. We find that enhanced Lennard-Jones (LJ) interactions between protein atoms and water oxygens primarily improve the solvation of nonpolar functional groups of the protein. In contrast, modified electrostatics in the water model or strengthened LJ interactions between the protein and water hydrogens mainly affect the hydration of polar functional groups. Modified electrostatics further impact the average orientation of water molecules in the hydration shell. As a result, protein–water interactions with the first hydration layers are strengthened, while interactions with water molecules in higher hydration shells are weakened. Hence, distinct strategies to balance intra-protein and protein–water interactions in simulations have qualitatively different effects on protein solvation. These differences are not necessarily captured by comparisons to experiments that report on global parameters describing protein conformational ensembles, e.g., the radius of gyration, but will influence the tendency of a protein to form aggregates or phase-separated droplets.



INTRODUCTION

While the structural information for folded proteins continues to grow^{1–4} and de novo predictions are feasible,^{5–7} the discovery of intrinsically disordered proteins (IDPs) that lack folded domains^{8–11} challenges the structure–function paradigm in molecular and structural biology. In addition to IDPs, many proteins with folded domains contain extensive intrinsically disordered regions (IDRs), which are required for function.^{12,13} IDPs and IDRs constitute about 33% of the total eukaryotic proteome, while in humans, this fraction is found to be closer to 50%.^{14,15} The expression of IDPs is often tightly regulated in the cell,¹⁶ and they exhibit a wide range of biological functions such as cellular signaling¹⁷ and molecular recognition.^{18,19} The biological function of some IDPs and IDRs can be tied to folding or partial folding, e.g., upon binding to a specific target,^{20,21} but the formation of high-affinity complexes has been reported even for fully unfolded IDPs.²²

IDPs are involved in multiple diseases, such as cancer, amyloidosis, and neurodegenerative diseases.^{23–27} To shed light on the pathological behavior of IDPs and the relationship between their sequence, function, and characteristic features such as the tendency for liquid–liquid phase separation,^{28–31} we need to understand how conformational ensembles accessible to IDPs in solution depend on interactions with their molecular environment. In principle, structural and dynamical properties of conformational ensembles explored

by IDPs can be characterized in small-angle X-ray scattering (SAXS), nuclear magnetic resonance (NMR), and Förster resonance energy transfer (FRET) experiments,^{32–34} while atomistic molecular dynamics (MD) simulations provide a valuable tool to obtain the underlying microscopic picture.^{35–37} However, MD simulations with standard empirical force-field models typically fail to reproduce experimental observations and tend to predict overly compact conformations of solvated IDPs.^{35,36,38} This led to the development of IDP-specific protein force-field parameters, which re-weight dihedral angles to modify the conformational ensembles of IDPs in simulations.³⁹ However, the shortcomings of simulations with standard force fields to describe IDPs are more commonly ascribed to an imbalance between non-covalent interactions within the protein, which stabilize compact states, and protein–solvent interactions, which stabilize extended conformations with a higher solvent-accessible surface area (SASA).^{35,36,38}

Received: March 14, 2023

Revised: July 20, 2023

Published: August 9, 2023



Table 1. Force-Field Parameter Sets Used in MD Simulations of the K-18 Domain of the Tau Protein

system name	protein FF	water FF	modification
A03w	AMBER03w ^a	TIP4P/2005 ^b	none
A03ws	AMBER03ws ^c	TIP4P/2005 ^b	1.10 scaling of LJ- $\epsilon_{O_wX_p}$
A99w	AMBER99SBw ^c	TIP4P/2005 ^b	none
A99ws	AMBER99SBws ^c	TIP4P/2005 ^b	1.10 scaling of LJ- $\epsilon_{O_wX_p}$
A99*	AMBER99SB*-ILDN ^d	TIP4P/2005 ^b	none
A99*Ew	AMBER99SB*-ILDN ^d	TIP4P-Ew ^e	none
A99*D	AMBER99SB*-ILDN ^d	TIP4P-D ^f	re-parameterized water model
C36m	CHARMM36m ^g	TIPS3P ^h	none
C36m*	CHARMM36m ^g	TIPS3P ^h	2.17 scaling of LJ- $\epsilon_{H_wX_p}$

^aReference 44. ^bReference 43. ^cReference 36. ^dReference 59. ^eReference 58. ^fReference 35. ^gReference 46. ^hReference 47.

This motivated multiple strategies to modify standard force fields to improve their ability to reproduce experimental results.³⁷ In one approach that combines protein force fields in the AMBER family^{40–42} with the TIP4P/2005 water model,⁴³ Lennard-Jones (LJ) interaction energies between water oxygens and protein atoms have been increased by 10%.^{36,44,45} Using a similar strategy, scaled LJ interactions between water hydrogens and protein atoms were proposed for the CHARMM36m force field⁴⁶ in combination with the TIPS3P water model⁴⁷ (a modified version of the TIP3P⁴⁸ water model used in CHARMM that includes water hydrogen LJ sites). In a different approach, Shaw and co-workers re-parameterized the TIP4P water model to improve dispersion interactions between the protein and the solvent while also modifying electrostatics and parameters in AMBER-type protein force fields.^{35,49–51} Encouraging results have also been achieved with new types of water models that emphasize the role of electrostatics and charge distributions over molecular geometry.^{52,53}

In each of the cases cited above, modified protein–water interactions result in more favorable extended conformations for IDPs in solution and improve the agreement between atomistic simulations and experimental data. However, the thermodynamic consequences of modified protein–water interactions for the solvation of proteins have so far only been analyzed for side-chain analogues.^{36,37} A more detailed analysis of protein solvation free energies including a decomposition of local enthalpic and entropic contributions has so far been missing. The latter is key to understanding the microscopic consequences of modified protein–water interactions that remain hidden by comparisons to structural parameters obtained in experiments.

In this study, we simulated the intrinsically disordered K-18 domain of the Tau protein^{32,54} with varying combinations of standard and modified proteins and water force fields. In addition to an analysis of the conformational ensembles obtained with each simulation model, we analyzed the protein solvation free energy and its enthalpic and entropic contributions for a selected conformational state with spatial resolution using the previously developed 3D-2PT (3D-two-phase thermodynamics) approach.^{55,56} Our analysis reveals that distinct strategies for the modification of protein–water interactions have distinguishable thermodynamic consequences. Further, protein solvation is not just affected globally, but each modification of force-field parameters results in unique

changes in the solvation thermodynamics of polar and non-polar groups of the protein that affect the conformational ensembles sampled in simulations.

METHODS

Conformational Sampling. All simulations were performed using the GROMACS 2018.1⁵⁷ software package. We chose the K-18 domain of the Tau protein as an example to study the impact of varying force-field parameterizations on the solvation of IDPs and their thermally accessible conformational space.^{32,54} We used four pairs of force-field parameters in our simulations, each containing one set of standard parameters and one set with modified protein–water interactions intended to stabilize extended conformations and improve agreement with experiments. We provide a list of the pairs of force-field parameter sets and modifications applied to protein–water interactions in Table 1, where we also introduce shorthand notations used throughout this work. The A03w and A99w protein force fields are modified versions of AMBER03 and AMBER99SB for use in combination with the highly optimized TIP4P/2005 water model.^{43,44} In A03ws and A99ws, the ϵ parameter that describes the strength of LJ interactions has been scaled by a factor of 1.10 for interactions between protein atoms and water oxygens.^{36,44} In A99*D, the four-site water model has been re-parameterized with increased dispersion interactions (affecting protein–water and water–water interactions simultaneously) and a larger water dipole moment that modifies electrostatic interactions and screening.³⁵ The resulting TIP4P-D water model shares its geometry with the TIP4P/2005 water model that we use here as a reference in the A99* system without modifying the protein force field. To ensure that our results do not depend on this combination of protein and water force fields, we performed an alternative version of our analysis with the TIP4P-Ew water model⁵⁸ as a reference for comparisons to TIP4P-D (A99*Ew). Lastly, in the C36m* model, the ϵ parameter for LJ interactions between protein atoms and water hydrogens has been increased by a factor of ~2.17.⁴⁶ The existence of such interactions is a specific feature of the TIPS3P water model, which has been developed in conjunction with the CHARMM force-field family for proteins.⁴⁷

To generate an initial atomistic structure model for MD simulations of the protein, we selected the first backbone conformation for the K-18 domain of Tau in the PED00017 entry of the protein ensemble database (PED)⁶⁰ (6AAC in a

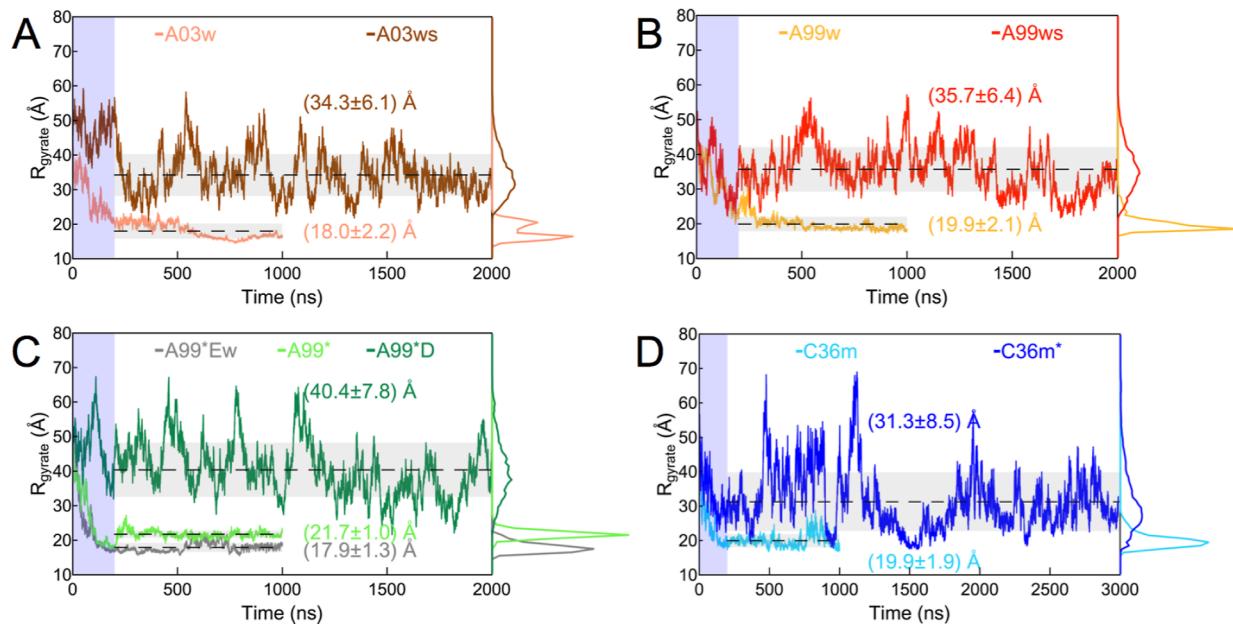


Figure 1. Time traces of the protein R_g in MD simulations of the K-18 domain of the Tau protein with distinct sets of force-field parameters (see Table 1). An equilibration time of 200 ns is highlighted and excluded from further analysis. Dashed horizontal lines indicate averages, and shaded gray backgrounds represent standard deviations due to conformational fluctuations. The numerical values are given as insets, and histograms of R_g are indicated on the alternative y -axis.

previous version of the PED³²). We then used the SCWRL4 package⁶¹ to add the missing side chains, which resulted in 1986 protein atoms for the 130 amino acid residues. We set the protonation state of each amino acid residue according to a pH of 7 based on standard side chain pK_a 's in solution. We then solvated the protein in a cubic box with a 164 Å side length by adding approximately 145,000 water molecules and 828 ions (409 of Na^+ and 419 of Cl^- ions), neutralizing the protein charge and resulting in a physiological salt concentration of 150 mM. Salt can influence the conformational ensemble and phase separation behavior of IDPs via electrostatic screening and salting-out effects, as discussed in recent work by Zheng et al.⁶²

We applied periodic boundary conditions (PBCs) in all dimensions and used the particle mesh Ewald algorithm⁶³ to describe long-ranged electrostatics using a 1.2 Å real space grid. We used a 10 Å real space cut-off for short-ranged electrostatic and LJ interactions, constrained all bond lengths in the protein with the LINCS⁶⁴ algorithm, and constrained the geometry of all water molecules in the system with SETTLE.⁶⁵ Simulations with all force fields were preceded by 1500 steps of energy minimization with the steepest descent algorithm.

We then equilibrated each system in the isothermal–isobaric ensemble (NPT) at 300 K and 1 bar for 1 ns using a time step of 1 fs, a velocity rescaling thermostat⁶⁶ with a 0.1 ps time constant, and a Berendsen weak-coupling barostat⁶⁷ with a time constant of 2.0 ps. For each system, we then performed 1–2 μ s simulations in the NPT ensemble at 300 K and 1 bar using a time step of 2 fs and a Nose–Hoover thermostat⁶⁸ and a Parrinello–Rahman barostat⁶⁹ with a time constant of 2.0 ps to sample the conformational space. We monitored the conformational dynamics of the simulated proteins via the root mean squared displacements (RMSDs) from the initial structure, the radius of gyration (R_g), and the SASA and considered the first 200 ns of each trajectory as additional equilibration time.

Solvation Thermodynamics. For an in-depth analysis of the solvation thermodynamics with 3D-2PT,^{55,56,70} we selected a moderately compact structure ($R_g = 24.6$ Å) with an extended SASA of 14,100 Å² from the A03ws (see Table 1) simulation. This specific structure was chosen due to the simultaneous presence of fully extended and partially collapsed portions of the protein (see Results).

For the 3D-2PT analysis, we followed procedures in previous work^{70–74} and performed an additional set of simulations during which the protein conformation was kept fixed, while hydration water configurations were sampled. For this purpose, position restraints with a force constant of 5000 kJ mol⁻¹ Å⁻² were applied to all protein atoms during all the following steps. First, we re-solvated the protein in its selected conformation with 30,000 water molecules in a 100 Å × 100 Å × 100 Å cubic simulation box and generated topologies for all force-field models described in Table 1. After performing a steepest descent energy minimization, we equilibrated each system for 1 ns in the NPT ensemble using a Berendsen weak-coupling thermostat and barostat⁶⁷ with a time constant of 1 ps, a reference temperature of 300 K, and a reference pressure of 1 bar. The applied position restraints maintained the RMSD for all protein atoms below 0.1 Å. Following the equilibration, we performed for each system 10 ns simulations in the NVT ensemble at 300 K using a Nose–Hoover thermostat with a time constant of 1 ps.⁷⁵ From this simulation, we extracted 100 independent solvent micro-states (coordinates and velocities) at regular 100 ps time intervals. These microstates were then used to initialize NVE simulations of 100 ps length, which were used for the 3D-2PT analysis. In these simulations, the integration time step was set to 1 fs, and coordinates and velocities were stored every 8 fs.

The 3D-2PT analysis, which is described in detail in ref 55 and in the Supporting Information, was carried out on a 98 × 98 × 98 analysis grid with a grid constant of 1 Å. The grid was used to compute local contributions to the solvation enthalpy,

ΔH_{solv} and solvation entropy, ΔS_{solv} , which are combined to describe contributions to the solvation free energy, ΔG_{solv} .

$$\Delta G_{\text{solv}} = \Delta H_{\text{solv}} - T\Delta S_{\text{solv}} \quad (1)$$

Ignoring negligible contributions from volume work, ΔH_{solv} is computed as a sum of pair-wise solute–solvent (uv) interactions, ΔU_{uv} , and changes in solvent–solvent (vv) interactions, ΔH_{vv} , caused by the presence of the solute. A similar separation is performed for ΔS_{solv} based on Ben-Naim's proof that solute-induced changes of the enthalpy and entropy cancel out exactly.^{76,77}

$$\Delta H_{\text{vv}} - T\Delta S_{\text{vv}} = 0 \quad (2)$$

Hence, ΔG_{solv} can alternatively be described as the sum of enthalpy and entropy changes caused solely by solute–water interactions.

$$\Delta G_{\text{solv}} = \Delta H_{\text{uv}} - T\Delta S_{\text{uv}} \quad (3)$$

As described in the *Supporting Information*, contributions to each term are computed per voxel of the analysis grid, either as an average local contribution per water molecule or as a spatial density. The latter can be summed over all voxels of the analysis grid or a fraction of it, e.g., to obtain the total solvation free energy or contributions from distinct components of the hydration shell, respectively.

RESULTS AND DISCUSSION

Comparison of Conformational Ensembles. We begin our analysis of the impact of force-field modifications on simulations of IDPs in solution with a comparison of conformational ensembles. For this purpose, we performed microsecond timescale simulations of the K-18 domain of the Tau protein.

As a reference, we used simulations with force-field parameters from the AMBER and CHARMM families that have been optimized for simulations of folded proteins (A03w, A99w, A99*/A99*-Ew, and C36m in Table 1).^{44,46,59} We then compared results from these reference simulations to simulations with modified force fields optimized to improve the description of IDPs in solution either via explicit modification of protein–water interactions or a re-parameterization of the water model (A03ws, A99ws, A99*D, and C36m*, see Table 1).^{35,36,44,46}

In Figure 1, we plot the time evolution of the radius of gyration R_G of the protein for all sets of simulations. The results show that the K-18 domain quickly collapses into a compact conformation in simulations with standard force-field parameters (A03w, A99w, A99*, and C36m), resulting in an average R_G between 20 and 22 Å with minor fluctuations. The sampling of these systems was stopped after 1 μ s due to the lack of conformational changes after the collapse. These observations are in agreement with previous simulation studies of other protein systems using standard force fields but are incompatible with experimental observations.^{36,38,46}

Experimentally, the average radius of gyration of the K-18 domain was determined as (33.7 ± 7.6) Å by Blackridge and co-workers using chemical shifts and residual dipolar couplings.³² All modified force fields optimized for simulations of IDPs tested in this study generate conformational ensembles that are in much better agreement with this experimental result. We extended simulations with these force fields up to 2–3 μ s to sample conformational fluctuations. Especially simulations with scaled LJ interactions between protein and

water atoms (A03ws, A99ws, and C36m*) accurately reproduce the experimental average, while simulations with the TIP4P-D water model (A99*D) slightly overshoot and predict an average radius of gyration that is too large.

A complementary comparison of the conformational ensembles generated in our simulations to single-molecule FRET data⁷⁸ is provided in Figure S1 of the *Supporting Information*. The experiments report an average C_α – C_α distance between the two cysteine residues of the K-18 domain of 38 Å. This distance is significantly smaller in simulations with standard force fields (20–24 Å), which matches our observations for R_G . For simulations with modified force fields, A03ws shows close agreement with the experiment, while we observed larger average distances of 45–50 Å for A99ws, A99*D, and C36m*, albeit with large fluctuations.

Both the NMR and single-molecule FRET experiments were performed in 50 mM sodium phosphate buffer at pH 6.8,^{32,78,79} which corresponds to an ionic strength of 78 mM (Debye length 10.9 Å). Compared to the salt concentration in our simulations (150 mM, Debye length 7.8 Å), we expect decreased electrostatic screening of repulsive interactions between the predominantly positive charges of the K-18 domain (more extended conformations) in these experiments. Compared to chloride, salting-out effects for anionic components in the experimental buffer are stronger,⁸⁰ but their concentration is lower.⁶² In summary, one may expect that conformations of the K-18 domain in experiments are slightly more extended than in our simulation. This is notable since our results in Figures 1 and S1 indicate that the average R_G and C_α – C_α distances observed in our simulations with modified force fields mostly exceed the experimental reference values. A03ws predicts average R_G and C_α – C_α distances that remain closest to the experiments. C36m* is an exception as it predicts an average R_G that is lower than observed experimentally but significantly overestimates the experimentally measured C_α – C_α distance.

In Figure S2 of the *Supporting Information*, we show a complementary analysis of the SASA versus simulation time, which reproduces the general trend observed in Figures 1 and S1: simulations with standard force-field parameters result in collapsed conformations, while extended conformations are observed in simulations with the modified force fields.

Joint probability distributions of R_G and SASA shown in Figures S3 and S4 of the *Supporting Information* further illustrate differences between the conformational ensembles observed in simulations with standard and modified force-field parameters. More importantly, the joint probability distributions highlight quantitative differences between the simulations with distinct sets of modified force fields. While all simulations with modified force-field parameters result in improved agreement with experimental observables, the conformational ensembles generated in each of these simulations remain noticeably distinct from each other.

Solvation Thermodynamics. To provide additional insights on how the distinct force-field modifications promote the formation of more extended and solvent-exposed protein conformations, we analyzed in the following the solvation thermodynamics of the K-18 domain using 3D-2PT.⁵⁶ To ensure comparability between the simulations, we perform these studies for a fixed conformation of the protein and sample only the protein–water and water–water interactions. The protein conformation was selected from the A03ws

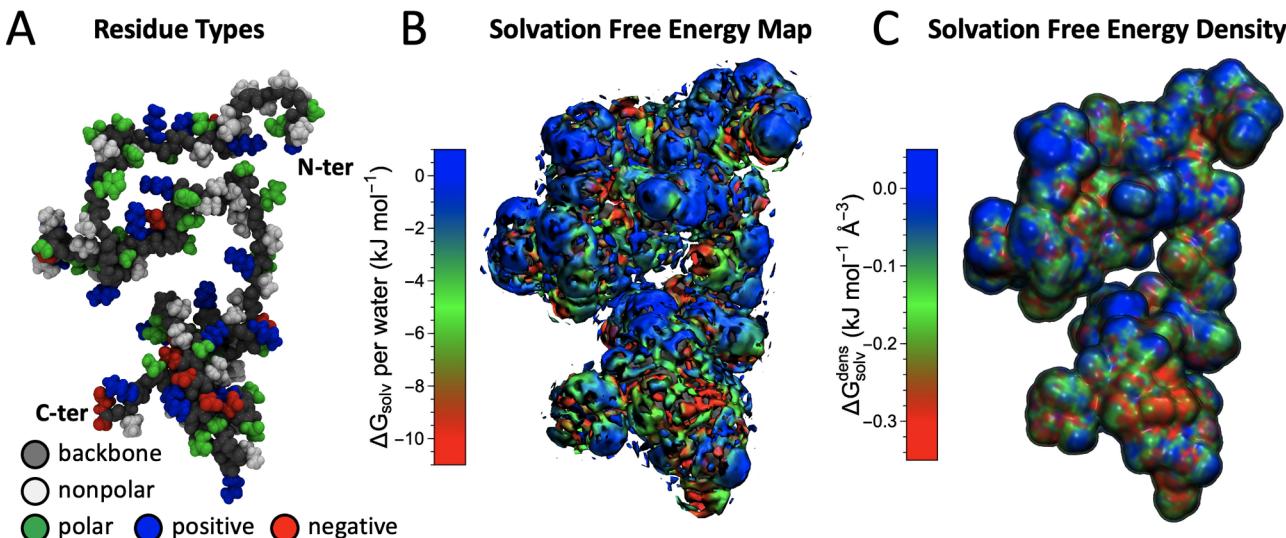


Figure 2. Solvation of the K-18 domain in a fixed conformation. (A) The selected conformation with $R_G = 24.6 \text{ \AA}$ and SASA = $13,260 \text{ \AA}^2$ was sampled as a comparably compact state in the A03w simulation, which features extended and collapsed domains. All atoms in the protein are shown as van der Waals spheres colored by the corresponding residue type, as indicated. (B) Exemplary solvation free energy map obtained with the 3D-2PT formalism using the A03w force field. Local contributions to ΔG_{solv} per water molecule (eq S10 in the Supporting Information) are projected on an iso-surface of the water density at 1.25 g/cm^3 , which effectively highlights the first hydration shell of the protein. (C) Solvation free energy density obtained from the dataset in (B) (eq S11 in the Supporting Information) color-coded on two surfaces with a constant distance of 3.0 and 3.5 \AA from the closest atom of the protein. The outer surface is semi-transparent to allow an average view on both surfaces.

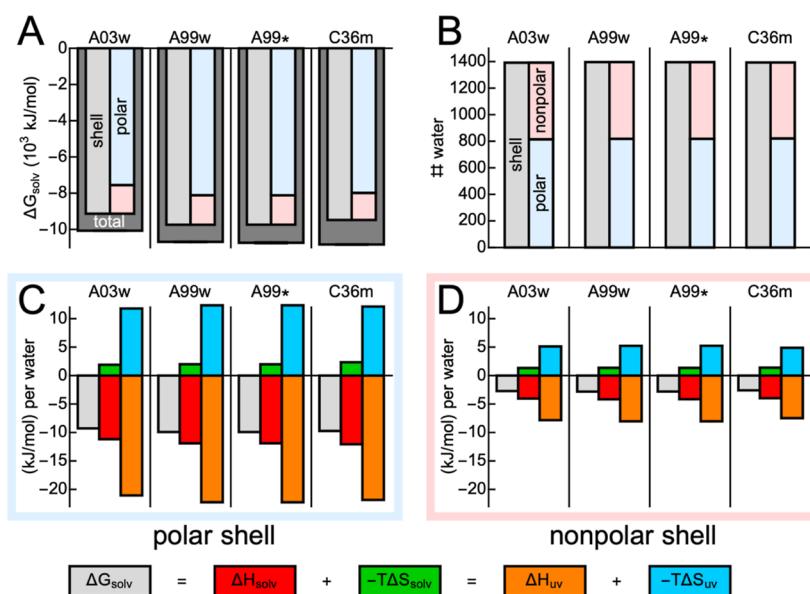


Figure 3. Solvation thermodynamics obtained with 3D-2PT for the fixed conformation of the K-18 domain in simulations with standard force-field models (see Table 1). (A) Total solvation free energy (gray), solvation free energy contributions of a 5 \AA shell around the protein defined by the distance to the closest non-hydrogen protein atom (light gray), and decomposition into contributions of hydration shells of polar (light blue) and nonpolar (light red) protein atoms. (B) Number of water molecules in the total 5 \AA hydration shell and its polar and nonpolar components (colors as in A). (C,D) Solvation free energy contributions per water molecule in the polar (C) and nonpolar (D) components of the hydration shell. The latter (light gray) is decomposed into enthalpic and entropic contributions, including (ΔH_{solv} in red and $-T\Delta S_{\text{solv}}$ in green) and excluding (ΔH_{uv} in orange and $-T\Delta S_{\text{uv}}$ in cyan) canceling contributions from solute-induced changes of water–water interactions (see Methods and Supporting Information). Error bars describing the error of the mean are included in all panels but not visible on the shown scale.

simulation and includes a mixture of compact and extended features. The spatially resolved analysis of the solvation thermodynamics for a single structure allows us to separate contributions to the solvation free energy, ΔG_{solv} , from the hydration shells of polar and nonpolar functional groups of the K-18 domain as well as longer-ranged contributions from water molecules with a minimum distance of 5 \AA from the closest

protein atom. While ΔG_{solv} will depend strongly on the protein conformation, our analysis aims to highlight overall differences in the protein solvation thermodynamics for distinct force fields that are not expected to depend strongly on the specific protein conformation.

IDP Solvation with Standard Force-Field Parameters. The 3D-2PT approach spatially resolves local contributions to

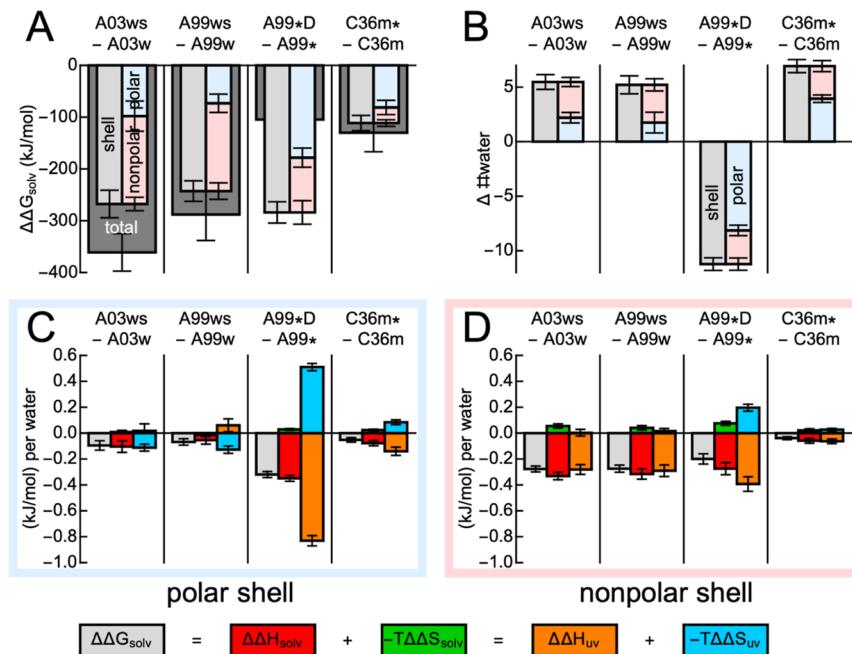


Figure 4. Change in solvation due to modified force-field parameters. (A) Change in the total solvation free energy (gray), solvation free energy contributions of a 5 Å shell around the protein defined by the distance to the closest non-hydrogen protein atom (light gray), and decomposition into contributions of hydration shells of polar (light blue) and nonpolar (light red) protein atoms. (B) Change in the number of water molecules in the total 5 Å hydration shell and its polar and nonpolar components. (C,D) Change of solvation free energy contributions per water molecule in the polar (C) and nonpolar (D) components of the hydration shell. The latter is decomposed into enthalpic and entropic contributions, including (ΔH_{solv} in red and $-\Delta S_{\text{solv}}$ in green) and excluding (ΔH_{uv} in orange and $-\Delta S_{\text{uv}}$ in cyan), canceling contributions from solute-induced changes of water–water interactions (see Methods and Supporting Information). Error bars indicate the statistical error of the mean.

the solvation free energy, enthalpy, and entropy on a 3D analysis grid. Each property can be expressed either per water molecule at a given site or as an energy density (per volume). The results are shown together with a 3D representation of the selected conformation of the K-18 domain in Figure 2. We visualize the selected conformation in panel A and highlight the location of charged, polar, and nonpolar side chains. In panel B, we illustrate a 3D map of local contributions to the solvation free energy. For this purpose, we used a color code to project spatially resolved solvation free energy contributions per water molecule on an iso-density surface of the water number density at 1.25 g/cm³. This iso-density surface encloses regions of increased water density characteristic for the first hydration shell of the protein where contributions to the solvation free energy are primarily located. In panel C, we show the closely related solvation free energy density projected on two surfaces with a constant distance to the closest protein atom, i.e., 3.0 and 3.5 Å, respectively (see caption). Both representations of the solvation free energy maps provide complementary information. Blue colors indicate sites with small or even positive (unfavorable) contributions to the protein solvation free energy, while green and red colors indicate local sites with significant favorable contributions to ΔG_{solv} . Differences between the results obtained from simulations with distinct force fields are not easily discernible at this scale, which is why we show here only the result for A03w.

Quantitative differences between distinct force fields become apparent upon integration over the entire analysis grid, the hydration shell (voxels within 5 Å of the closest non-hydrogen protein atom), and its polar and nonpolar components. Here, we distinguish polar and nonpolar components of the

hydration shell based on the partial charge (in the AMBER03 force field) of the closest non-hydrogen atom of the protein. If a voxel of the analysis grid lies within the protein hydration shell and the closest non-hydrogen atom of the protein carries a partial charge q with $|q| > 0.2$, we define it as part of the polar hydration shell. All other voxels of the protein hydration shell are defined as the nonpolar hydration shell.

We first analyzed distinct contributions to the solvation free energy of the K-18 domain in simulations with standard force fields described in Table 1. In panel A of Figure 3, we compare the integrated solvation free energies of the K-18 domain in simulations with the distinct standard force fields (prior to the introduction of any modifications). The total solvation free energy obtained upon integrating over the entire analysis grid is most favorable for the C36m system and least favorable for A03w. The solvation free energy is a critical component of the free energy surface sampled by the protein in unconstrained simulations. However, the total free energy surface of the protein also includes the intramolecular protein potential energy and its conformational entropy,⁷³ which are excluded from our present analysis.

Figure 3A shows that a dominant fraction of the total solvation free energy stems from contributions in the 5 Å hydration shell (~90%). Within the hydration shell, solvation free energy contributions are dominated by contributions from the polar hydration shell in all systems. The latter is partially caused by the respective numbers of water molecules in the polar and nonpolar hydration shells shown in panel B of Figure 3. The 5 Å hydration shell consists of just below 1400 water molecules in all systems. Approximately 60% of these water molecules are located in the polar and ~40% in the nonpolar hydration shell. However, more important than these different

populations are the distinct thermodynamic properties of water in polar and nonpolar hydration shells, which are quantified and decomposed into enthalpic and entropic components (per water molecule) in panels C and D of Figure 3.

Independent of the force-field model, water molecules in the polar hydration shell contribute ~ 9 kJ/mol to the solvation free energy, while water molecules in the nonpolar hydration shell only contribute ~ 2 kJ/mol. This difference is not surprising given the hydrophilic and hydrophobic properties of the polar and nonpolar functional groups of the protein. Decomposing local solvation free energy contributions into enthalpic and entropic contributions allows us to analyze the distinct protein–water interactions in more detail.

If protein-induced changes in the water–water interactions are included, the solvation entropy ΔS_{solv} appears only as a minor factor. The positive $-T\Delta S_{\text{solv}}$ term compensates 18–21% of the corresponding negative ΔH_{solv} term in the polar hydration shell, which seemingly dominates the favorable solvation free energy contribution of ~ 9 kJ/mol per water molecule. In the nonpolar hydration shell, both terms are smaller and the positive $-T\Delta S_{\text{solv}}$ term compensates 35–36% of the negative ΔH_{solv} term, which leads to the weakly favorable ~ 2 kJ/mol contribution to ΔG_{solv} per water molecule.

As discussed in the context of eqs 2 and 3, ΔH_{solv} and ΔS_{solv} include the exactly canceling protein-induced changes of water–water interactions, ΔH_{vv} and ΔS_{vv} , which have no impact on the free energy. To focus only on non-canceling contributions caused directly by solute–water interactions, we thus computed ΔH_{uv} and ΔS_{uv} . This reveals the actual thermodynamic driving forces that govern solvation and allows for direct comparisons to analytical theory. Specifically, linear response theory for the solvation of ions, dipoles, etc., in a polar solvent, i.e., solute–solvent interactions dominated by electrostatics,^{81,82} predicts the following ratio^{81,82}

$$\gamma = \Delta G_{\text{solv}} / \Delta H_{uv} = 1/2 \quad (4)$$

For both the polar and nonpolar hydration shells, ΔH_{uv} and $-T\Delta S_{uv}$ are substantially larger than ΔH_{solv} and $-T\Delta S_{\text{solv}}$ terms. Consequently, the degree of cancellation between ΔH_{uv} and $-T\Delta S_{uv}$ is also increased. In the polar hydration shell, 43% of the negative ΔH_{uv} term is compensated by positive $-T\Delta S_{uv}$ contributions. In other words, the ratio γ evaluates to 0.43, which is just below the linear response prediction for electrostatically dominated solvation. In the nonpolar hydration shell, this ratio drops to just ~ 0.34 with a weak dependence on the force field used.

We previously introduced deviations of γ from the linear response prediction as a measure of the relative importance of electrostatic solute–solvent interactions with $\gamma = 0.5$ as an upper limit.⁷³ We then used an empirical expression for γ to predict the solvation free energies of a small peptide as a function of its conformation based on ensemble averages of ΔH_{uv} .⁷³ In the present study, the observed decrease of γ from 0.43 in the polar to 0.34 in the nonpolar hydration shell follows the expected trend, i.e., the solvation of polar functional groups is dominated by electrostatics, while the solvation of nonpolar functional groups is not.

Changes in IDP Solvation with Modified Force Fields.

In Figure 4, we analyzed changes in the solvation of the K-18 domain induced by modified force-field parameters, i.e., we analyzed the difference between solvation free energy contributions obtained in simulations with modified and standard force-field parameters. An alternative version of this

analysis is shown in Figure S5 of the Supporting Information, where the A99*Ew system was used for comparison with A99*D. As shown in panel A, the solvation free energy changes, $\Delta\Delta G_{\text{solv}}$, are favorable for all four force-field pairs, i.e., the modified force fields result in an overall more negative ΔG_{solv} of the protein. The latter is expected given the more expanded protein conformations observed in simulations with the modified force fields in Figures 1 and S1–S4. However, the magnitude, localization, and decomposition of $\Delta\Delta G_{\text{solv}}$ differ between distinct pairs of standard and modified force fields. In Figure 4A, the magnitude of the total $\Delta\Delta G_{\text{solv}}$ is significantly larger for the comparisons of A03ws and A99ws with their reference systems than for C36m*. For A99*D, the magnitude of the total $\Delta\Delta G_{\text{solv}}$ depends on the choice of the reference system. If A99* (with the TIP4P/2005 water model) is used as a reference, the total $\Delta\Delta G_{\text{solv}}$ is comparable to C36m*. However, if we use A99*Ew (with the TIP4P-Ew water model) as a reference, the total $\Delta\Delta G_{\text{solv}}$ is comparable to A03ws and A99ws. The latter is interesting as it does not directly correlate with changes of the sampled conformational ensembles sampled with the modified force fields analyzed in Figures 1 and S1–S4. The lack of correlation can be attributed to different intra-protein and protein–water interactions in the corresponding reference systems.

It is further important to note that the different effects of force-field modifications on the solvation of polar and nonpolar groups discussed below imply that the total ΔG_{solv} and $\Delta\Delta G_{\text{solv}}$ for the compared pairs of modified and standard force fields are expected to depend on the protein conformation selected for the analysis, i.e., the relative exposure of polar and nonpolar functional groups. Hence, we focus our following analysis on changes of solvation parameters that we expect to be less dependent on the protein conformation, i.e., qualitative contributions to $\Delta\Delta G_{\text{solv}}$ from the first and higher hydration shells and per water molecule contributions in hydration shells of polar and nonpolar functional groups.

In the systems in which the protein–water interactions were modified directly, i.e., A03ws, A99ws, and C36m*, favorable contributions to $\Delta\Delta G_{\text{solv}}$ are found in the protein hydration shell as well as to a lesser extent in higher hydration shells for distances larger than 5 Å from the protein. The latter results in an increasing magnitude of $\Delta\Delta G_{\text{solv}}$ when the integration is performed over the full analysis grid instead of just the hydration shell. Notably, this situation is significantly different for A99*-D with the re-parameterized TIP4P-D water model (independent of the choice of the reference system, see Figure S5 in the Supporting Information). Here, favorable $\Delta\Delta G_{\text{solv}}$ contributions in the hydration shell are partially compensated by unfavorable contributions at larger distances from the protein. We can rationalize the latter by the nature of the force-field modifications.

In the A99*D system, the re-parameterized water model features a larger dipole moment, which affects the preferential orientation of water molecules within the electric field of the protein. These changes are largest at short distances and in the polar hydration shell, where local contributions to $\Delta\Delta G_{\text{solv}}$ are most pronounced. Water molecules at distances >5 Å from the protein interact more strongly with water molecules in the first hydration shell (with which they can form direct hydrogen bonds) than with the protein. Consequently, water molecules in higher hydration shells prioritize their interactions with water in the first hydration shell over weaker direct interactions with the protein. Thus, altered water orientations in the first

hydration shell of the A99*D system propagate into succinct hydration shells, where they result in unfavorable changes of protein–water interactions and thus unfavorable contributions to $\Delta\Delta G_{\text{solv}}$.

In addition to $\Delta\Delta G_{\text{solv}}$, we analyzed changes in the number of water molecules within the 5 Å hydration shell due to the modified force-field parameters in Figure 4B. As one might expect, the number of water molecules increased in all cases where protein–water interactions were modified to be more attractive, i.e., due to changes in protein–water LJ interactions in the A03ws, A99ws, and C36m* systems. However, for the A99*D system, where the re-parameterized water model affects both protein–water and water–water interactions, the number of water molecules in the hydration shell decreases despite the favorable $\Delta\Delta G_{\text{solv}}$ reported in panel A. While less compared to the overall number of hydration water molecules (see Figure 3B), these changes highlight that the impact of simultaneously modified protein–water and water–water interactions is non-trivial to predict. The same applies to changes in the solvation free energy: only changes in water–water interactions induced by the presence of the protein solute cancel out as described in eq 2, while modified water–water interactions in the force-field model do not.

In Figure 4C,D, we decomposed $\Delta\Delta G_{\text{solv}}$ contributions from the polar and nonpolar hydration shells into contributions per water molecule. This decomposition reveals additional details that inform how the distinct modifications of the force-field parameters impact protein solvation in each case.

For A03ws and A99ws, $\Delta\Delta G_{\text{solv}}$ contributions per water molecule are largest in the nonpolar hydration shell, while the opposite is the case for A99*D and C36m*. We begin the discussion of these differences with the nonpolar hydration shell in the A03ws and A99ws systems, where LJ interactions between protein atoms and water oxygens have been scaled.

In the nonpolar hydration shell, protein–water LJ interactions do not compete with electrostatic protein–water interactions such as hydrogen bonds. Existing attractive LJ interactions simply become stronger upon scaling, while changes in the water structure are minimal. The latter can be deduced by the minor difference between $\Delta\Delta H_{\text{solv}}$ and $\Delta\Delta H_{\text{uv}}$, which describes the change of water–water interactions $\Delta\Delta H_{\text{vv}}$. Further, the decomposition of $\Delta\Delta G_{\text{solv}}$ shows that the latter is essentially identical to $\Delta\Delta H_{\text{uv}}$, and no significant compensation by the $-T\Delta\Delta S_{\text{uv}}$ term is observed. For hydrophobic surfaces in water, ΔS_{uv} is dominated by repulsive protein–water interactions,⁸³ which are insensitive to scaling of the corresponding LJ potentials. The steepness of the repulsive term prevents the population of strongly repulsive configurations even prior to scaling.

This situation is different in the polar hydration shell of the A03ws and A99ws systems. Here, LJ interactions between protein atoms and water oxygens, which are isotropic and do not prefer any specific water orientation, compete with directional protein–water hydrogen bonds. Scaling the LJ interactions thus increases the weight of LJ interactions within this balance and results in overall less directional protein–water interactions. The consequences can be compared to previous studies of bulk water with modified water models, which showed how an increased weight of isotropic LJ interactions weakens the structure of the water hydrogen bond network.⁸⁴ In the polar hydration shell of the A03ws and A99ws systems, we thus observe a (small) favorable $-T\Delta\Delta S_{\text{uv}}$ term, i.e., an increase of the solute–solvent entropy. The

resulting change in the hydration water structure is particularly pronounced in the A99ws system, where it even leads to a net unfavorable change of the protein–water interaction energy, i.e., $\Delta\Delta H_{\text{uv}}$.

Compared to the observation for A03ws and A99ws, changes in solvation are qualitatively different for the A99*D and C36m* systems, where the introduced changes in the force field affect the preferential orientation of hydration water molecules.

In the A99*D system, the increased dipole moment of the water model results in an enhanced strength of electrostatic protein–water interactions such as protein–water hydrogen bonds, which is particularly evident in the $\Delta\Delta H_{\text{uv}}$ term for the polar hydration shell. This effect on electrostatic interactions seems to be more important than the enhanced dispersion interactions in the TIP4P-D water model. In both the polar and non-polar hydration shells, ~50% of the favorable $\Delta\Delta H_{\text{uv}}$ is compensated by an unfavorable $-T\Delta\Delta S_{\text{uv}}$ term, which is expected for primarily modified electrostatic interactions (following the same linear response approach discussed in the context of eq 4). Since electrostatic protein–water interactions are significantly more relevant in the polar hydration shell, the resulting contributions to $\Delta\Delta G_{\text{solv}}$ are larger than in the nonpolar hydration shell. The large difference between $\Delta\Delta H_{\text{solv}}$ and $\Delta\Delta H_{\text{uv}}$, especially in the polar hydration shell, indicates a significant change in the hydration shell structure and water–water interactions. The latter is a consequence of the previously discussed modified preferential orientation of hydration water molecules and the simultaneous change of water–water interaction potentials in the water model.

In the C36m* system, the scaling of short-ranged LJ interactions between protein atoms and water hydrogens, which are specific to the TIPSP water model, primarily enhance the strength of protein–water hydrogen bonds. Consequently, the impact on $\Delta\Delta G_{\text{solv}}$ (while small in comparison to the other modified force fields) is primarily localized in the polar hydration shell. Interestingly, the partial cancellation of ~50% of the favorable $\Delta\Delta H_{\text{solv}}$ term by an unfavorable $-T\Delta\Delta S_{\text{uv}}$ term closely resembles the expectations for modified electrostatic interactions despite the actual nature of the force-field modification.

CONCLUSIONS

We used all-atom MD simulations of the K-18 domain of the Tau protein to compare the impact of previously proposed force-field modifications to improve agreement between conformational ensembles of IDPs in simulations and experiments. We utilized four distinct pairs of standard and modified force fields that implement distinct strategies to increase favorable protein–water interactions and stabilize extended conformational states. We initially analyzed conformational ensembles sampled in simulations on the microsecond timescale before performing a detailed analysis of spatially resolved solvation free energy maps.

As expected, all force-field modifications that stabilize extended protein conformations result in favorable changes in the protein solvation free energy compared to standard parameters. However, our analysis based on spatially resolved solvation free energy maps of a single protein conformation identifies significant differences that arise from each specific strategy to increase attractive protein–water interactions. These differences are expected to affect the properties of

conformational ensembles sampled in MD simulations as well as water-mediated interactions between proteins that influence the ability to form complexes in solution, to aggregate, or even phase-separate.^{85,86}

For example, simulations with the TIP4P-D water model were found to enhance protein–water interactions in the first hydration layers relative to reference simulations with either TIP4P/2005 or TIP4P-Ew, while changes of protein–water interactions over longer distances were found to be unfavorable. We can associate this behavior with a change in the preferential orientation of water molecules in the first hydration shell in response to the modified protein–water interactions. Such restructuring of inner hydration shells propagates into the solution, where water molecules tend to optimize interactions with their immediate solvent environment and interactions with more distant protein atoms only play a secondary role.

Quantitative definitions of solvent-mediated interactions, e.g., between two proteins in solution, include the cost of depopulating each other's hydration shells upon complex formation. As a consequence, simulations with the TIP4P-D water model facilitate the approach of other proteins up to distances of up to 5 Å compared to simulations with the reference water models. Only the final desolvation of the protein–water interface, required to form direct protein–protein contacts, is expected to involve more work due to the increased thermodynamic cost of removing water molecules in the hydration shell compared to simulations with a standard force field.

In addition, the distinct force-field modifications exhibit distinct effects on the solvation of polar and nonpolar functional groups of the protein. A03ws and A99ws primarily enhance the solvation of nonpolar functional groups, while A99*D and C36m* primarily enhance the solvation of polar functional groups.

Distinct preferences to enhance the solvation of polar and nonpolar groups will have direct consequences on changes of conformational ensembles sampled with standard and modified force-field parameters. Simulations with A03ws and A99ws will tend to stabilize extended and solvent-exposed states of nonpolar residues, while simulations with A99*D and C36m* will tend to stabilize extended and solvent-exposed states of polar residues (relative to the corresponding standard force field).

For IDPs whose sequence includes both polar and nonpolar residues, such differences are not likely to be evident in comparisons of global parameters to characterize protein conformational states such as the radius of gyration or end-to-end distance. However, such differences will become more dominant if a protein primarily consists of polar or nonpolar amino acids.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpcb.3c01726>.

Theoretical background on 3D-2PT solvation free energy maps and their evaluation and analysis of C_{α} – C_{α} distances, SASA, and additional data for the A99*Ew reference system ([PDF](#))

AUTHOR INFORMATION

Corresponding Author

Matthias Heyden – School of Molecular Sciences, Arizona State University, Tempe, Arizona 85287, United States;
DOI: orcid.org/0000-0002-7956-5287; Email: mheyden1@asu.edu

Author

Sthitadhi Maiti – School of Molecular Sciences, Arizona State University, Tempe, Arizona 85287, United States

Complete contact information is available at:
<https://pubs.acs.org/10.1021/acs.jpcb.3c01726>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors are grateful to the Gordon and Betty Moore Foundation for financial support (Award #9162.10) and acknowledge Research Computing at Arizona State University for providing high-performance computing resources that have contributed to the research results reported within this work.

REFERENCES

- (1) Kendrew, J. C.; Bodo, G.; Dintzis, H. M.; Parrish, R.; Wyckoff, H.; Phillips, D. C. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* **1958**, *181*, 662–666.
- (2) Berman, H. M. The protein data bank: a historical perspective. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **2008**, *64*, 88–95.
- (3) Goodsell, D. S.; Zardecki, C.; Di Costanzo, L.; Duarte, J. M.; Hudson, B. P.; Persikova, I.; Segura, J.; Shao, C.; Voigt, M.; Westbrook, J. D.; et al. RCSB Protein Data Bank: Enabling biomedical research and drug discovery. *Protein Sci.* **2020**, *29*, 52–65.
- (4) Armstrong, D. R.; Berrisford, J. M.; Conroy, M. J.; Gutmanas, A.; Anyango, S.; Choudhary, P.; Clark, A. R.; Dana, J. M.; Deshpande, M.; Dunlop, R.; et al. PDBe: improved findability of macromolecular structure data in the PDB. *Nucleic Acids Res.* **2020**, *48*, D335–D343.
- (5) Senior, A. W.; Evans, R.; Jumper, J.; Kirkpatrick, J.; Sifre, L.; Green, T.; Qin, C.; Žídek, A.; Nelson, A. W.; Bridgland, A.; et al. Improved protein structure prediction using potentials from deep learning. *Nature* **2020**, *577*, 706–710.
- (6) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589.
- (7) Anishchenko, I.; Baek, M.; Park, H.; Hiranuma, N.; Kim, D. E.; Dauparas, J.; Mansoor, S.; Humphreys, I. R.; Baker, D. Protein tertiary structure prediction and refinement using deep learning and Rosetta in CASP14. *Proteins: Struct., Funct., Bioinf.* **2021**, *89*, 1722–1733.
- (8) Tompa, P. Intrinsically unstructured proteins. *Trends Biochem. Sci.* **2002**, *27*, 527–533.
- (9) Dunker, A. K.; Babu, M. M.; Barbar, E.; Blackledge, M.; Bondos, S. E.; Dosztányi, Z.; Dyson, H. J.; Forman-Kay, J.; Fuxreiter, M.; Gsponer, J.; et al. What's in a name? Why these proteins are intrinsically disordered: Why these proteins are intrinsically disordered. *Intrinsically Disord. Proteins* **2013**, *1*, No. e24157.
- (10) Tompa, P. Intrinsically disordered proteins: a 10-year recap. *Trends Biochem. Sci.* **2012**, *37*, 509–516.
- (11) Van Der Lee, R.; Buljan, M.; Lang, B.; Weatheritt, R. J.; Daughdrill, G. W.; Dunker, A. K.; Fuxreiter, M.; Gough, J.; Gsponer, J.; Jones, D. T.; et al. Classification of intrinsically disordered regions and proteins. *Chem. Rev.* **2014**, *114*, 6589–6631.
- (12) Bickers, S. C.; Sayewich, J. S.; Kanelis, V. Intrinsically disordered regions regulate the activities of ATP binding cassette transporters. *Biochim. Biophys. Acta Biomembr.* **2020**, *1862*, 183202.

- (13) Maiti, S.; Acharya, B.; Boorla, V. S.; Manna, B.; Ghosh, A.; De, S. Dynamic studies on intrinsically disordered regions of two paralogous transcription factors reveal rigid segments with important biological functions. *J. Mol. Biol.* **2019**, *431*, 1353–1369.
- (14) Deiana, A.; Forcelloni, S.; Porrello, A.; Giansanti, A. New classification of intrinsic disorder in the Human proteome. **2018**, bioRxiv, 446351.
- (15) Colak, R.; Kim, T.; Michaut, M.; Sun, M.; Irimia, M.; Bellay, J.; Myers, C. L.; Blencowe, B. J.; Kim, P. M. Distinct types of disorder in the human proteome: functional implications for alternative splicing. *PLoS Comput. Biol.* **2013**, *9*, No. e1003030.
- (16) Babu, M. M.; van der Lee, R.; de Groot, N. S.; Gsponer, J. Intrinsically disordered proteins: regulation and disease. *Curr. Opin. Struct. Biol.* **2011**, *21*, 432–440.
- (17) Wright, P. E.; Dyson, H. J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 18–29.
- (18) Krieger, J. M.; Fusco, G.; Lewitzky, M.; Simister, P. C.; Marchant, J.; Camilloni, C.; Feller, S. M.; De Simone, A. Conformational recognition of an intrinsically disordered protein. *Biophys. J.* **2014**, *106*, 1771–1779.
- (19) Hazy, E.; Tompa, P. Limitations of induced folding in molecular recognition by intrinsically disordered proteins. *Chem.-PhysChem* **2009**, *10*, 1415–1419.
- (20) Santofimia-Castaño, P.; Rizzuti, B.; Xia, Y.; Abian, O.; Peng, L.; Velázquez-Campoy, A.; Neira, J. L.; Iovanna, J. Targeting intrinsically disordered proteins involved in cancer. *Cell. Mol. Life Sci.* **2019**, *77*, 1695–1707.
- (21) Kim, D.-H.; Han, K.-H. Transient secondary structures as general target-binding motifs in intrinsically disordered proteins. *Int. J. Mol. Sci.* **2018**, *19*, 3614.
- (22) Borgia, A.; Borgia, M. B.; Bugge, K.; Kissling, V. M.; Heidarsson, P. O.; Fernandes, C. B.; Sottini, A.; Soranno, A.; Buhholzer, K. J.; Nettels, D.; et al. Extreme disorder in an ultrahigh-affinity protein complex. *Nature* **2018**, *555*, 61–66.
- (23) Uversky, V. N.; Oldfield, C. J.; Midic, U.; Xie, H.; Xue, B.; Vucetic, S.; Iakoucheva, L. M.; Obradovic, Z.; Dunker, A. K. Unfoldomics of human diseases: linking protein intrinsic disorder with diseases. *BMC Genom.* **2009**, *10*, S7.
- (24) Uversky, V. N.; Dave, V.; Iakoucheva, L. M.; Malaney, P.; Metallo, S. J.; Pathak, R. R.; Joerger, A. C. Pathological unfoldomics of uncontrolled chaos: intrinsically disordered proteins and human diseases. *Chem. Rev.* **2014**, *114*, 6844–6879.
- (25) Uversky, V. N.; Oldfield, C. J.; Dunker, A. K. Intrinsically disordered proteins in human diseases: introducing the D2 concept. *Annu. Rev. Biophys.* **2008**, *37*, 215–246.
- (26) Trojanowski, J. Q.; Virginia, M. Parkinson's disease and related α -synucleinopathies are brain amyloidoses. *Ann. N.Y. Acad. Sci.* **2006**, *991*, 107–110.
- (27) Goedert, M. Filamentous nerve cell inclusions in neurodegenerative diseases: tauopathies and alpha-synucleinopathies. *Philos. Trans. R. Soc. London, Ser. B* **1999**, *354*, 1101–1118.
- (28) Brangwynne, C. P.; Eckmann, C. R.; Courson, D. S.; Rybarska, A.; Hoege, C.; Gharakhani, J.; Jülicher, F.; Hyman, A. A. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **2009**, *324*, 1729–1732.
- (29) Hyman, A. A.; Weber, C. A.; Jülicher, F. Liquid-liquid phase separation in biology. *Annu. Rev. Cell Dev. Biol.* **2014**, *30*, 39–58.
- (30) Alberti, S.; Gladfelter, A.; Mittag, T. Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates. *Cell* **2019**, *176*, 419–434.
- (31) Alberti, S.; Dormann, D. Liquid-liquid phase separation in disease. *Annu. Rev. Genet.* **2019**, *53*, 171–194.
- (32) Ozenne, V.; Schneider, R.; Yao, M.; Huang, J.-r.; Salmon, L.; Zweckstetter, M.; Jensen, M. R.; Blackledge, M. Mapping the potential energy landscape of intrinsically disordered proteins at amino acid resolution. *J. Am. Chem. Soc.* **2012**, *134*, 15138–15148.
- (33) Gomes, G.-N. W.; Krzeminski, M.; Namini, A.; Martin, E. W.; Mittag, T.; Head-Gordon, T.; Forman-Kay, J. D.; Grdinariu, C. C. Conformational ensembles of an intrinsically disordered protein consistent with NMR, SAXS, and single-molecule FRET. *J. Am. Chem. Soc.* **2020**, *142*, 15697–15710.
- (34) Schuler, B.; Müller-Späth, S.; Soranno, A.; Nettels, D. *Intrinsically Disordered Protein Analysis*; Springer, 2012; pp 21–45.
- (35) Piana, S.; Donchev, A. G.; Robustelli, P.; Shaw, D. E. Water dispersion interactions strongly influence simulated structural properties of disordered protein states. *J. Phys. Chem. B* **2015**, *119*, 5113–5123.
- (36) Best, R. B.; Zheng, W.; Mittal, J. Balanced protein–water interactions improve properties of disordered proteins and non-specific protein association. *J. Chem. Theory Comput.* **2014**, *10*, 5113–5124.
- (37) Nerenberg, P. S.; Jo, B.; So, C.; Tripathy, A.; Head-Gordon, T. Optimizing solute–water van der Waals interactions to reproduce solvation free energies. *J. Phys. Chem. B* **2012**, *116*, 4524–4534.
- (38) Henriques, J.; Cragnell, C.; Skepo, M. Molecular dynamics simulations of intrinsically disordered proteins: force field evaluation and comparison with experiment. *J. Chem. Theory Comput.* **2015**, *11*, 3420–3431.
- (39) Song, D.; Wang, W.; Ye, W.; Ji, D.; Luo, R.; Chen, H.-F. ff14IDPs force field improving the conformation sampling of intrinsically disordered proteins. *Chem. Biol. Drug Des.* **2017**, *89*, 5–15.
- (40) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
- (41) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; et al. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **2003**, *24*, 1999–2012.
- (42) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712–725.
- (43) Abascal, J. L.; Vega, C. A general purpose model for the condensed phases of water: TIP4P/2005. *J. Chem. Phys.* **2005**, *123*, 234505.
- (44) Best, R. B.; Mittal, J. Protein simulations with an optimized water model: cooperative helix formation and temperature-induced unfolded state collapse. *J. Phys. Chem. B* **2010**, *114*, 14916–14923.
- (45) Best, R. B.; de Sancho, D.; Mittal, J. Residue-Specific α -Helix Propensities from Molecular Simulation. *Biophys. J.* **2012**, *102*, 1462–1467.
- (46) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D., Jr. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **2017**, *14*, 71–73.
- (47) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- (48) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The missing term in effective pair potentials. *J. Phys. Chem.* **1987**, *91*, 6269–6271.
- (49) Robustelli, P.; Piana, S.; Shaw, D. E. Developing a molecular dynamics force field for both folded and disordered protein states. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, E4758–E4766.
- (50) Piana, S.; Robustelli, P.; Tan, D.; Chen, S.; Shaw, D. E. Development of a force field for the simulation of single-chain proteins and protein–protein complexes. *J. Chem. Theory Comput.* **2020**, *16*, 2494–2507.
- (51) Robustelli, P.; Piana, S.; Shaw, D. E. Mechanism of coupled folding-upon-binding of an intrinsically disordered protein. *J. Am. Chem. Soc.* **2020**, *142*, 11092–11101.
- (52) Izadi, S.; Anandakrishnan, R.; Onufriev, A. V. Building water models: a different approach. *J. Phys. Chem. Lett.* **2014**, *5*, 3863–3871.

- (53) Shabane, P. S.; Izadi, S.; Onufriev, A. V. General purpose water model can improve atomistic simulations of intrinsically disordered proteins. *J. Chem. Theory Comput.* **2019**, *15*, 2620–2634.
- (54) PED3: Protein Ensemble Database. <http://pedb.vib.be/> (accessed Aug 01, 2019).
- (55) Persson, R. A.; Pattini, V.; Singh, A.; Kast, S. M.; Heyden, M. Signatures of solvation thermodynamics in spectra of intermolecular vibrations. *J. Chem. Theory Comput.* **2017**, *13*, 4467–4481.
- (56) Heyden, M. 3D-2PT, 2022. <https://github.com/MatthiasHeyden/3D-2PT> (Accessed June 2, 2022), DOI: 10.5281/zenodo.7311555.
- (57) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1–2*, 19–25.
- (58) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys.* **2004**, *120*, 9665–9678.
- (59) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Struct., Funct., Bioinf.* **2010**, *78*, 1950–1958.
- (60) Lazar, T.; Martínez-Pérez, E.; Quaglia, F.; Hatos, A.; Chemes, L. B.; Iserte, J. A.; Méndez, N. A.; Garrone, N. A.; Saldaño, T. E.; Marchetti, J.; et al. PED in 2021: a major update of the protein ensemble database for intrinsically disordered proteins. *Nucleic Acids Res.* **2021**, *49*, D404–D411.
- (61) Krivov, G. G.; Shapovalov, M. V.; Dunbrack, R. L., Jr. Improved prediction of protein side-chain conformations with SCWRL4. *Proteins: Struct., Funct., Bioinf.* **2009**, *77*, 778–795.
- (62) Wohl, S.; Jakubowski, M.; Zheng, W. Salt-dependent conformational changes of intrinsically disordered proteins. *J. Phys. Chem. Lett.* **2021**, *12*, 6684–6691.
- (63) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- (64) Hess, B.; Bekker, H.; Berendsen, H. J.; Fraaije, J. G. LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
- (65) Miyamoto, S.; Kollman, P.-A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **1992**, *13*, 952–962.
- (66) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.
- (67) Berendsen, H.-J.-C.; Postma, J.-P.-M.; van Gunsteren, W.-F.; DiNola, A.; Haak, J.-R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- (68) Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **1984**, *52*, 255–268.
- (69) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981**, *52*, 7182–7190.
- (70) Pattini, V.; Heyden, M. Pressure effects on protein hydration water thermodynamics. *J. Phys. Chem. B* **2019**, *123*, 6014–6022.
- (71) Pattini, V.; Vasilevskaya, T.; Thiel, W.; Heyden, M. Distinct protein hydration water species defined by spatially resolved spectra of intermolecular vibrations. *J. Phys. Chem. B* **2017**, *121*, 7431–7442.
- (72) Heyden, M. Heterogeneity of water structure and dynamics at the protein-water interface. *J. Chem. Phys.* **2019**, *150*, 094701.
- (73) Fajardo, T. N.; Heyden, M. Dissecting the conformational free energy of a small peptide in solution. *J. Phys. Chem. B* **2021**, *125*, 4634–4644.
- (74) Waskasi, M. M.; Lazaric, A.; Heyden, M. Solvent-mediated forces in protein dielectrophoresis. *Electrophoresis* **2021**, *42*, 2060–2069.
- (75) Hoover, W.-G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A: At., Mol., Opt. Phys.* **1985**, *31*, 1695–1697.
- (76) Ben-Naim, A. Hydrophobic interaction and structural changes in the solvent. *Biopolymers* **1975**, *14*, 1337–1355.
- (77) Ben-Amotz, D. Water-mediated hydrophobic interactions. *Annu. Rev. Phys. Chem.* **2016**, *67*, 617–638.
- (78) Stelzl, L. S.; Pietrek, L. M.; Holla, A.; Oroz, J.; Sikora, M.; Köfinger, J.; Schuler, B.; Zweckstetter, M.; Hummer, G. Global structure of the intrinsically disordered protein tau emerges from its local structure. *JACS Au* **2022**, *2*, 673–686.
- (79) Mukrasch, M. D.; Biernat, J.; von Bergen, M.; Griesinger, C.; Mandelkow, E.; Zweckstetter, M. Sites of tau important for aggregation populate β -structure and bind to microtubules and polyanions. *J. Biol. Chem.* **2005**, *280*, 24978–24986.
- (80) Hofmeister, F. *Zur Lehre von der Wirkung der Salze*; Springer-Verlag: Berlin, New York, 1888; Vol. 25, pp 1–30.
- (81) Roux, B.; Yu, H. A.; Karplus, M. Molecular basis for the Born model of ion solvation. *J. Phys. Chem.* **1990**, *94*, 4683–4688.
- (82) Åqvist, J.; Hansson, T. On the validity of electrostatic linear response in polar solvents. *J. Phys. Chem.* **1996**, *100*, 9512–9521.
- (83) Monroe, J. I.; Jiao, S.; Davis, R. J.; Robinson Brown, D.; Katz, L. E.; Shell, M. S. Affinity of small-molecule solutes to hydrophobic, hydrophilic, and chemically patterned interfaces in aqueous solution. *Proc. Natl. Acad. Sci. U.S.A.* **2021**, *118*, No. e2020205118.
- (84) Lynden-Bell, R. M.; Rasaiah, J. C. Mobility and solvation of ions in channels. *J. Chem. Phys.* **1996**, *105*, 9266–9280.
- (85) Papoian, G. A.; Ulander, J.; Wolynes, P. G. Role of water mediated interactions in protein–protein recognition landscapes. *J. Am. Chem. Soc.* **2003**, *125*, 9170–9178.
- (86) Ambadipudi, S.; Biernat, J.; Riedel, D.; Mandelkow, E.; Zweckstetter, M. Liquid–liquid phase separation of the microtubule-binding repeats of the Alzheimer-related protein Tau. *Nat. Commun.* **2017**, *8*, 275.

Recommended by ACS

Dynamics of Hydrogen Bonds between Water and Intrinsically Disordered and Structured Regions of Proteins

Korey M. Reid, David M. Leitner, et al.
SEPTEMBER 06, 2023
THE JOURNAL OF PHYSICAL CHEMISTRY B

READ

Benchmarking Molecular Dynamics Force Fields for All-Atom Simulations of Biological Condensates

Kumar Sarthak, Aleksei Aksimentiev, et al.
MAY 03, 2023
JOURNAL OF CHEMICAL THEORY AND COMPUTATION

READ

Optimized OPEP Force Field for Simulation of Crowded Protein Solutions

Stepan Timr, Fabio Sterpone, et al.
APRIL 18, 2023
THE JOURNAL OF PHYSICAL CHEMISTRY B

READ

Cumulative Millisecond-Long Sampling for a Comprehensive Energetic Evaluation of Aqueous Ionic Liquid Effects on Amino Acid Interactions

Till El Harrar and Holger Gohlke
DECEMBER 15, 2022
JOURNAL OF CHEMICAL INFORMATION AND MODELING

READ

Get More Suggestions >