SCIENTIFIC REP⚙RTS

OPEN

Correct folding of an α-helix and a β-hairpin using a polarized 2D torsional potential

Ya Gao[1], Yongxiu Li[1], Lirong Mou[2], Bingbing Lin[1], John Z. H. Zhang[1,3] & Ye Mei[1,3]

A new modification to the AMBER force field that incorporates the coupled two-dimensional main chain torsion energy has been evaluated for the balanced representation of secondary structures. In this modified AMBER force field (AMBER03^{2D}), the main chain torsion energy is represented by 2-dimensional Fourier expansions with parameters fitted to the potential energy surface generated by high-level quantum mechanical calculations of small peptides in solution. Molecular dynamics simulations are performed to study the folding of two model peptides adopting either α-helix or β-hairpin structures. Both peptides are successfully folded into their native structures using an AMBER03^{2D} force field with the implementation of a polarization scheme (AMBER03^{2D}p). For comparison, simulations using a standard AMBER03 force field with and without polarization, as well as AMBER03^{2D} without polarization, fail to fold both peptides successfully. The correction to secondary structure propensity in the AMBER03 force field and the polarization effect are critical to folding Trpzip2; without these factors, a helical structure is obtained. This study strongly suggests that this new force field is capable of providing a more balanced preference for helical and extended conformations. The electrostatic polarization effect is shown to be indispensable to the growth of secondary structures.

Unraveling the mechanism of protein folding has been a grand challenge in biological science for decades[1–3]. Relentless experimental effort has been devoted to the study of protein folding and unfolding under various conditions[4–6]. In parallel, theoretical and computational methods are playing an increasingly important role in this field. With continuous advances in computer technology, the lengthy time scales required for protein folding by computer simulation have lessened significantly. For example, a recent atomistic simulation of a 58-residue protein BPTI conducted by Shaw *et al.* achieved results in one millisecond[7]. However, the successful application of computer simulations is impeded by the accuracy of the force fields employed. Defects in potential energy can strongly bias the simulation result toward incorrect conformations[8,9].

Despite cases of excellent agreement between experiments and computer simulations reported in the literature[10–12], some inherent deficiencies of conventional force fields are well known. For example, the α-helical propensity of the AMBER03[13] force field is too high relative to experimental measurements, while that of AMBER99SB[14] is arguably too low[15]. CHARMM27 also has a helical propensity. A 10-μs simulation of the all β-structure of the Pin WW domain starting from an unfolded conformation resulted in a purely helical structure when run with CHARMM27[9]. Thus, developing a well-balanced force field for various secondary structures poses a challenge to computational biology. Best *et al.* calibrated the backbone rotation terms in the widely used AMBER03, AMBER99SB, and CHARMM22/ CMAP force fields, which resulted in a significant improvement in reproducing experimental residual

[1]State Key Laboratory of Precision Spectroscopy, Department of Physics and Institute of Theoretical and Computational Science, East China Normal University, Shanghai 200062 China. [2]Institutes for Advanced Interdisciplinary Research, East China Normal University, Shanghai 200062 China. [3]NYU-ECNU Center for Computational Chemistry at NYU Shanghai, Shanghai China 200062. Correspondence and requests for materials should be addressed to Y.M. (email: ymei@phy.ecnu.edu.cn)

dipolar couplings[16,17]. Lindorff *et al.* refined the χ_1 torsion potentials for amino acid side chains in AMBER99SB and found much closer agreement between the simulations and NMR experiments for the rotameric states of all four modified residues[18]. Other torsional corrections to CHARMM, GROMOS, and OPLS force fields have also been suggested[19–23]. Lindorff *et al.* performed a systematic validation of these refined force fields against experimental data, and suggested that force fields have improved over time. Nevertheless, they also highlighted residual deficiencies in force fields and noted areas for future improvement, such as temperature dependence[24]. Along with defects in torsional potential, the lack of an inhomogeneous polarization effect is also a likely reason for the failure of some protein folding simulations. The polarization effect has been shown to be critical to the success of protein folding[25–33]. The crucial role played by the polarization effect in protein self-assembly has also been observed[34,35].

In widely used AMBER force fields, the main chain torsion terms are treated individually, for ϕ (C-N-Cα-C) and ψ (N-Cα-C-N), as one-dimensional Fourier series. Due to the limited number of parameters for the torsion term, parameterization mainly focuses on a small portion of space near some energy basins. This leads to errors in torsion energy for inter-basin regions and barriers[36–38]. As these two main chain torsion energies are not physically separable in the potential energy map, a more desirable choice is to expand the torsion energy function by using orthogonal basis sets of both ϕ and ψ in a coupled fashion. Recently, we calibrated the AMBER03 and AMBER99SB force fields by replacing the one-dimensional Fourier series for ϕ and ψ with a coupled 2-dimensional (2D) main chain torsion energy (denoted AMBER03^{2D} and AMBER99SB^{2D})[39]. These two force fields are more balanced among various conformations than the original AMBER force fields. This idea has also been adopted by Lwin and Luo in their ff99ci force field[40]. However, because the model system used for parameterization is an alanine dipeptide, the electrostatic polarization effect along main chain hydrogen bonds has not been included, which is essential when ordered secondary structures such as α-helices and β-strands are formed. Therefore, the performance of AMBER03^{2D} and AMBER99SB^{2D} in molecular dynamics simulations of real proteins still remains to be examined.

In our previous work, we have assessed the performance of these force fields by checking the simulated conformation distribution and NMR J coupling of short alanine peptides. In this work, we evaluate the reliability of AMBER03^{2D} in the folding simulations of two short peptides representing, respectively, an α-helix and a β-sheet. Our results show that when combining this new force field and the electrostatic polarization effect, both peptides can fold into their native structures. For comparison, the AMBER03 force field is too helical to fold the β-hairpin peptide whether the polarization effect is included or excluded.

Methodology

2D torsion force field. A detailed explanation of the implementation and parameterization of this coupled 2D torsion potential can be found in our previous work[39]. Here we give a brief introduction. The alanine dipeptide (AD) was chosen as a model system for parameterization. The fitting of main chain torsion energy at a molecular mechanical (MM) level aims to reproduce the total energy of AD at a quantum mechanical (QM) level. Energy calculations are conducted on 24×24 grid points in the 2-dimensional space of main chain torsions with a 15° interval. All quantum mechanical calculations are performed at the M06 2X/aug-cc-pvtz//HF/6-31G** level by Gaussian 09[41,42]. At the QM level, the solvation effect is implemented by the integral equation formalism variant of the polarizable continuum model (IEFPCM)[43] in both optimization and single point calculations. The QM potential energy of AD can be written as

$$E_{QM} = E_{int} + G_{PCM}. \tag{1}$$

The potential energy of AD at the MM level is expressed as

$$E_{MM} = E_{mct} + E_{oth} + G_{GB}, \tag{2}$$

in which E_{mct} and E_{oth} are the main chain torsion related energy terms and other internal contributions, respectively, and G_{GB} is the generalized Born solvation free energy. By equalizing the MM and QM energies, the main chain torsion term can be calculated as

$$E_{mct} = E_{int} - E_{oth} + G_{PCM} - G_{GB}. \tag{3}$$

In the AMBER force field, the torsion energy is expressed as a series of separate Fourier expansions for ϕ and ψ. However, a more rational method involves decomposing the main chain torsion energy map with a double Fourier series, after which the expansion coefficients can be shown in matrix form as

$$\mathrm{C}(\mathrm{m}, \mathrm{n}) = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \mathrm{d}\phi \int_{-\pi}^{\pi} d\psi E(\phi, \psi) e^{\mathrm{im}\,\phi} e^{\mathrm{in}\psi}, \tag{4}$$

in which $E(\phi, \psi)$ are the main chain torsion potential energies (E_{mct}). At any value of (ϕ, ψ), the potential energy and the force acting on atom *k* involved in backbone torsions can be calculated, respectively, as

$$E(\phi, \psi) = \sum_{m=-N_\phi}^{N_\phi} \sum_{n=-N_\psi}^{N_\psi} C(m, n) e^{-im\phi} e^{-in\psi}, \tag{5}$$

and

$$F_k(\phi, \psi) = -\frac{\partial E(\phi, \psi)}{\partial R_k} = -\frac{\partial E(\phi, \psi)}{\partial \phi}\frac{\partial \phi}{\partial R_k} - \frac{\partial E(\phi, \psi)}{\partial \psi}\frac{\partial \psi}{\partial R_k}, \tag{6}$$

where R_k is the position of atom k.

Molecular dynamics simulations for benchmark systems. In our previous study, the accuracy of 2D force fields was validated with an implicit water model, showing close agreement with experimental results. In this work, an explicit water model (TIP3P) was used. We chose some well-studied systems for validations, for which experimental J coupling data are available. The benchmark systems used include capped dipeptides (Ace-X-NME, X≠P), tripeptides (XXX, X∈{A, G, V}; GYG, Y∈{A, V, F, L, S, E, K, M}), and alanine tetrapeptide. All systems used have available NMR data in the form of chemical shifts, J couplings, or both. Secondary structures for dipeptides are classified into "α_+" ($-160° < \phi < -20°$ and $-120° < \psi < 50°$), "α_h" with a more stringent definition ($-100° < \phi < -30°$ and $-67° < \psi < -7°$), "β" ($-180° < \phi < -90°$ and $50° < \psi < 240°$ or $160° < \phi < 180°$ and $110° < \psi < 180°$), and "PPII" ($-90° < \phi < -20°$ and $50° < \psi < 240°$) as used by Best *et al.*, and compared with experimental values from vibrational spectroscopy[44]. Along with the secondary structure population, $^3J(H_N, H_\alpha)$ coupling was calculated from the Karplus equation to inspect the intrinsic conformational distributions of different force fields. The parameters used in this work are taken from Hu *et al.*[45], which is an experimental investigation of ubiquitin averaged over various residues.

$$^3J(\theta) = A\cos^2(\theta + \Delta) + B\cos(\theta + \Delta) + C \tag{7}$$

Electrostatic polarization effect. In classic molecular mechanics, main chain hydrogen bonds stabilize the secondary structure through Coulomb interactions. However, when a hydrogen bond forms or breaks, the electron density redistributes to lower the total energy of the system. Therefore, electrostatic polarization plays an important role in protein folding. We employ a recently developed polarization model[31]. This model is based on the theoretical study of two alanine dipeptides that interact through a main chain hydrogen bond. By alternating the distance between the donor and the acceptor (d) from 2.5 Å to 6.5 Å and fixing the other degrees of freedom (DoF), a series of quantum mechanical calculations can be performed to obtain the electronic structures of the interacting dipeptide pair. In the calculation of each dipeptide, the other dipeptide is taken as a group of background charges located at the nuclei to polarize the electronic structure of the dipeptide studied. Electrostatic potentials on grids around the residue are calculated based on polarized wave functions, and new atomic charges are obtained using the restrained electrostatic potential (RESP) charge fitting method[46]. Only the charge transfers between the N and H atoms and between the C and O atoms involved in the hydrogen bond are allowed. Other atoms have their atomic charges fixed to the original AMBER charges. Quantum mechanical calculations and charge fitting are performed iteratively until convergence is reached. Finally, the charge alternations of N and O atoms can be fitted to single exponential functions of d as

$$\Delta q_N = -0.493 \times \exp(-0.455 \times d), \tag{8}$$

and

$$\Delta q_O = -0.334 \times \exp(-0.466 \times d). \tag{9}$$

This polarization model has proven to be quite effective at improving the thermodynamics of 2I9M[31] and the much longer helix 2KHK[32].

Folding simulations of α-helix and β-hairpin proteins. 2I9M is a "*de novo*" designed 17-residue helical peptide. Its native structure is determined by NMR experimentation at 283K and pH = 5.0. Trpzip2 is a 12-residue mini-protein (PDB ID: 1LE1) designed by Cochran *et al.*[47] This peptide adopts a β-hairpin structure enabling π-stacking between tryptophan residues. The salt bridge between GLU5 and LYS8 may play a role in stabilizing the turn region of this protein[48]. In this paper, we employ both replica exchange molecular dynamics (REMD) and direct molecular dynamics simulations to study these two peptides. In REMD simulations, high temperature replicas can easily surmount the energy barrier and explore the phase space more efficiently than standard molecular dynamics at room temperature[49]. Free energy is calculated by the Weighted Histogram Analysis Method (WHAM)[50,51] as
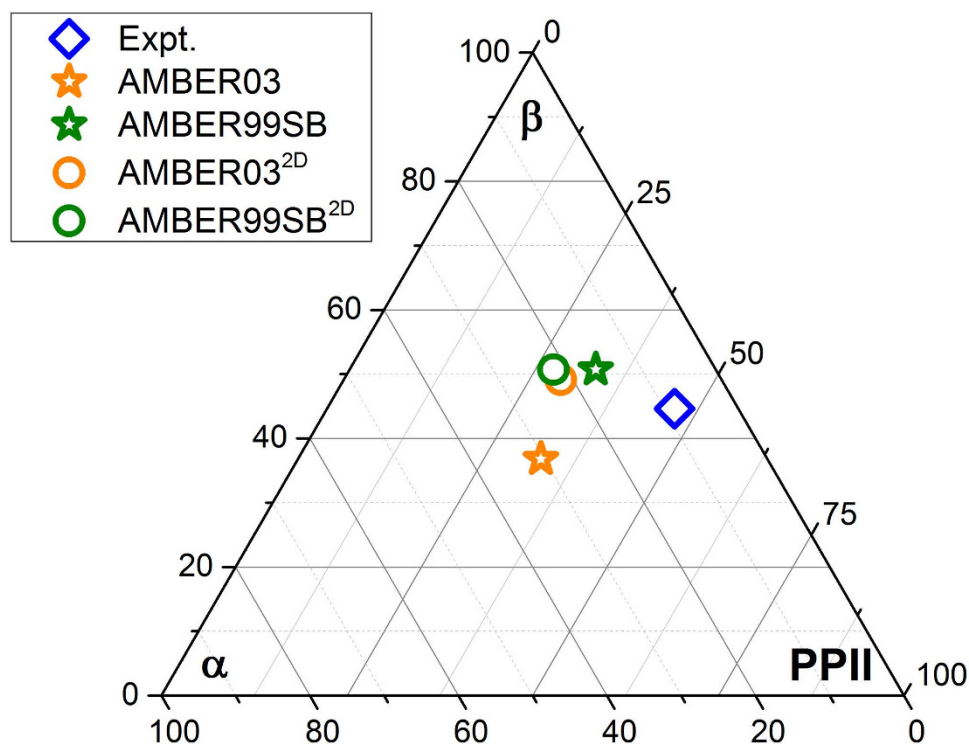
Figure 1. The simulated average conformational population for 19 dipeptides under AMBER03, AMBER03^{2D}, AMBER99SB, and AMBER99SB^{2D} force fields. Each axis denotes the population of α, β, and PPII from the bottom, clockwise. The corners of the triangle represent distributions with 100% α, β, or PPII, respectively.

$$\Omega_{mk} = \frac{H_{mk}}{\sum_{l=1}^{L} N_{kl} \Delta U e^{[f_l - \beta_l U_m]}}. \qquad (10)$$

Direct molecular dynamics simulations are performed at room temperature. More simulation details can be found in the section **Simulation Details**. The force fields employed include AMBER03, AMBER03 plus polarization (denoted as AMBER03p), AMBER03^{2D}, and AMBER03^{2D} plus polarization (denoted as AMBER03^{2D}p).

Results and Discussion

Accuracy examination of AMBER^{2D} force fields for benchmark systems in an explicit water model.

The torsional parameters used are derived from an alanine dipeptide. Thus, a key question is whether the performance of these 2D force fields is consistent among different amino acids. First, we evaluate the quality of AMBER03^{2D} and AMBER99SB^{2D} force fields for 19 amino acids using molecular simulations and compare these results with those of AMBER03 and AMBER99SB force fields. The torsional parameters for Proline are left intact as in the original AMBER force field. The conformational populations from simulations using the AMBER03, AMBER03^{2D}, AMBER99SB, and AMBER99SB^{2D} force fields are shown in Fig. 1. Because of the uncertainties in structure identification and the fitting procedure used in experimental analysis, this comparison is less direct than comparisons of NMR couplings. Despite this limitation, population analysis provides several insights. The α population in dipeptide simulations under the AMBER03 force field is nearly one fold larger than that under the AMBER99SB force field. After the 2D torsional potential correction, the secondary structural populations for AMBER03^{2D} and AMBER99SB^{2D} force fields are nearly the same, which indicates the well-balanced character of these 2D force fields. There is still a large discrepancy between calculations and experimental results, which may be due to systematic error in the experiment.

Along with these 19 dipeptides, we selected other benchmark systems for validation. Following Pande et al.[52], we calculated the $^3J(H_N, H_\alpha)$, which indicates the intrinsic secondary distribution of force fields. The result is shown in Fig. 2 for various systems. From this figure we can see that the torsional parameters derived from single amino acids are not readily applicable to all residue types. Large deviations can be observed for specific residues. The difference in solvent models between QM and MM calculations is
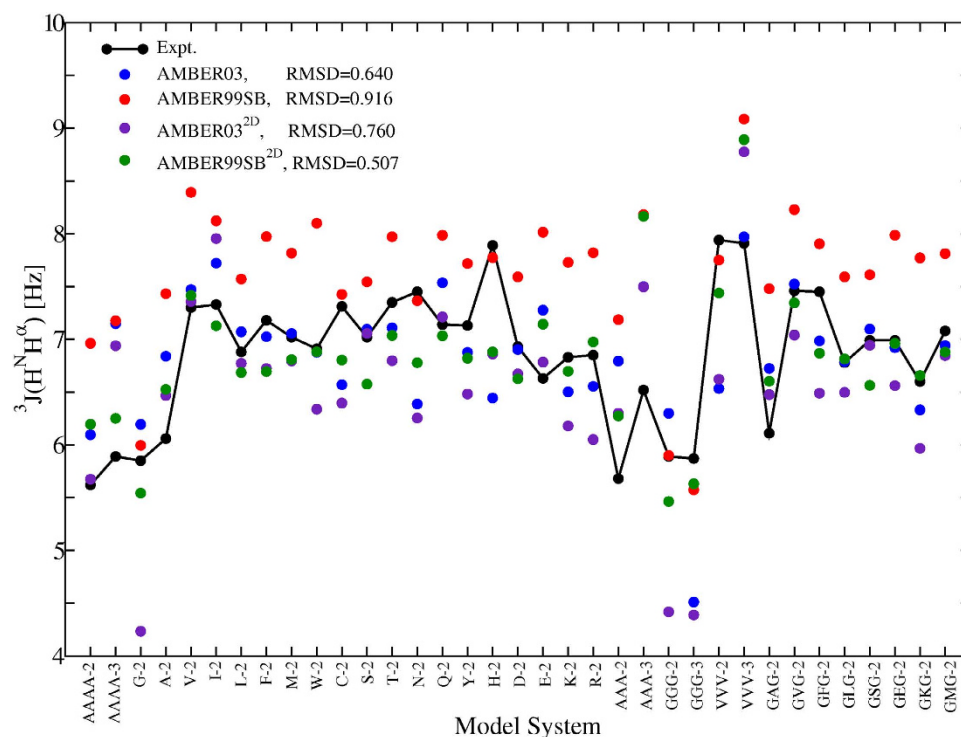
Figure 2. The simulated $^3J(H_N, H_\alpha)$ under AMBER03, AMBER03^{2D}, AMBER99SB, and AMBER99SB^{2D} force fields, and the RMSDs from experimental measurements.

also a possible source error in the revised force fields. For the AMBER03^{2D} force field, glycine is the largest contributor of error, which is consistent with previous observations[53]. However, the AMBER99SB^{2D} force field performs significantly better than its predecessor, leading to noticeable improvements in force field accuracy.

REMD simulations with the AMBER03^{2D} force field. Nucleation and helical growth are favored by the formation of main chain hydrogen bonds but are retarded by the desolvation of polar atoms and the loss of conformational entropy. The model peptide primarily adopts extended conformations, including β and PPII, under the AMBER03^{2D} force field[39]. This intrinsic conformational propensity of the AMBER03^{2D} force field results directly from the quantum mechanical potential energy surface of the alanine dipeptide and agrees with experimental observations[44,54,55]. The folding simulation of 2I9M can verify whether the enthalpic gain in the formation of a helix overcomes hindrances and shifts the population from extended conformations to helical structures. Conformation clustering analysis shows that there is no nucleus formed in the major clusters shown in Fig. 3. The free energy landscape shows that 2I9M covers a broad area in the space of RMSD and Rg under this force field, illustrating that there is no strongly preferential conformation adopted by this peptide. The minimal free energy structure has a much larger Rg than that of the native structure. The failure of peptide 2I9M to fold indicates that hydrogen bonding is not strong enough to compete with desolvation and the loss of entropy.

Clustering analysis of the AMBER03^{2D} trajectory shows that the peptide trpzip2 has a large portion of its population (42%) in conformations with a turn structure, as shown in Fig. 3. These conformations are close to the native structure, but further advancement to the folded state by the formation of main chain hydrogen bonds and π-stacking of tryptophan side chains is not achieved. There is only one minimum in the free energy landscape mapped to RMSD and Rg. Although located in the unfolded domain, the compactness of the conformations in this minimum is very close to the native structure. The free energy of the native structure is approximately 2 kcal/mol higher than that of this minimum. Therefore, there is a potential energy penalty in dewetting the backbone polar atoms, and entropy loss impedes the folding of trpzip2.

REMD simulations with the AMBER03^{2D}p force field. During the parameterization of the AMBER03^{2D} force field, quantum mechanical (QM) calculations are performed for various conformations of the alanine dipeptide. The main chain torsion term is fitted by minimizing the difference between the QM and MM potential energies. However, this model system is too small to form main chain hydrogen bonds as in typical helices or β-strands. Therefore, the electrostatic polarization effect due to hydrogen bond formation cannot be captured in the AMBER03^{2D} force field. The pairwise AMBER charge is
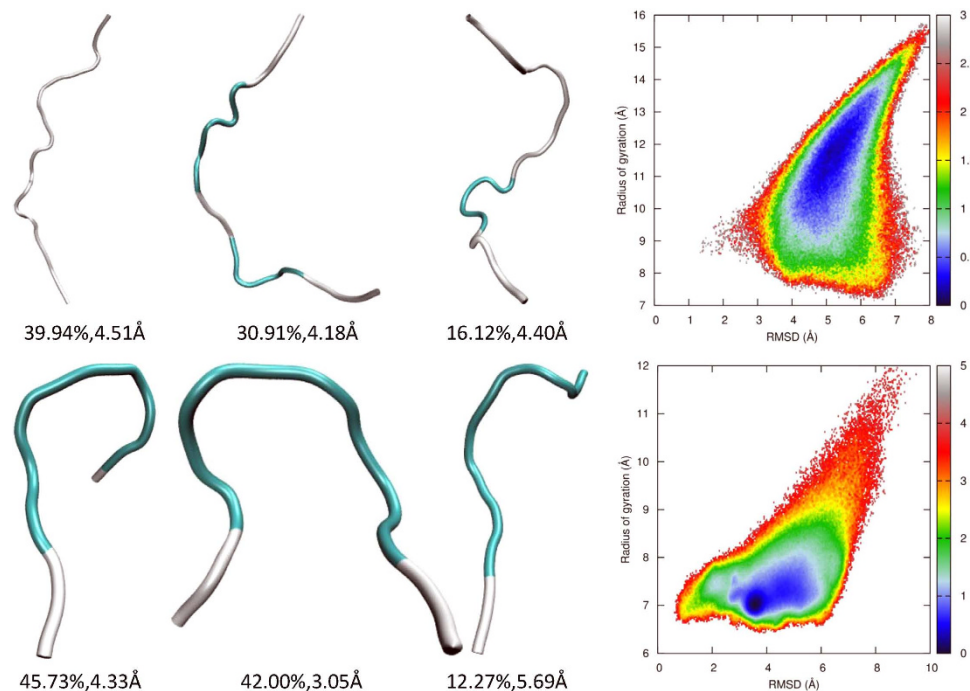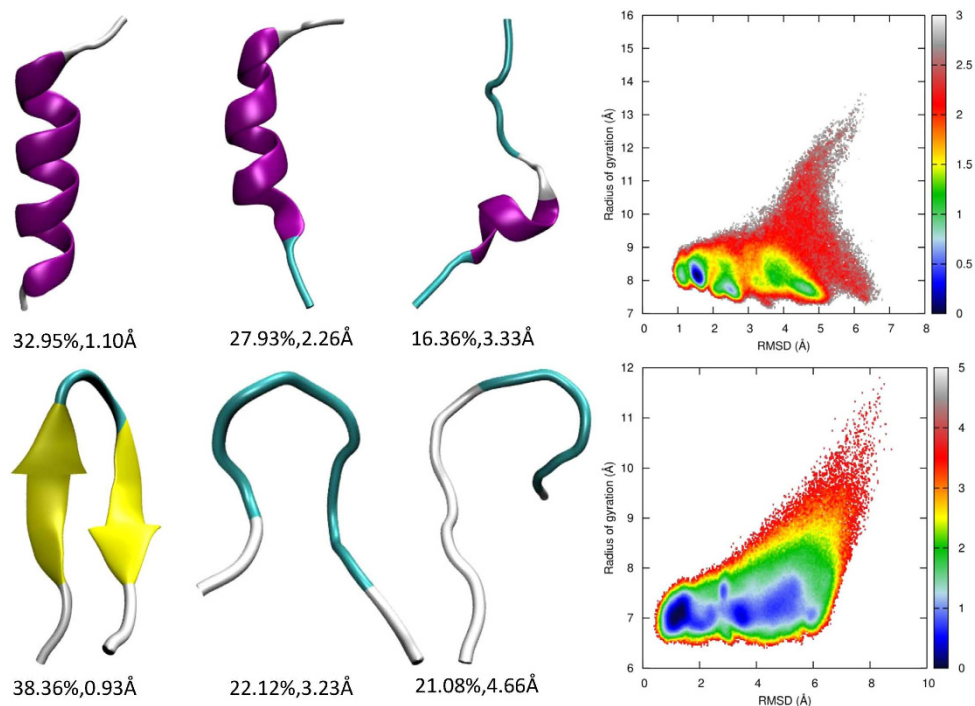
Figure 3. Representative structures from clustering analysis and the free energy landscape in the space of RMSD from the native structure and the radius of gyration for 2I9 M (top panel) and trpzip2 (bottom panel) using the AMBER03^{2D} force field at 300 K, from the REMD simulations.

based on a mean-field approximation. The heterogeneous polarization in a hydrogen bonded residue pair is not explicitly included. When hydrogen bonds occur, the electron density redistributes, giving rise to enhanced attraction between the donor and the acceptor. Therefore, it is necessary to incorporate a polarization effect into the AMBER03^{2D} force field. The polarization effect over the AMBER03^{2D} force field shifts the population of 2I9M significantly toward the native conformation. The RMSD distribution peaks at approximately 1.7 Å, which is conventionally regarded as the folded region. Nonetheless, there is still a large population residing in the unfolded region. As shown in Fig. 4, the largest cluster contains 32.95% of conformations with the centroid structure a fully folded helix. The second largest cluster contains 27.93% of the conformations, of which the centroid structure is partially unfolded. The peptide covers a narrow phase space with the global minimum in the folded domain and two local minima in the unfolded domain. These two local minima adopt even more compact conformations (with smaller Rg) than the native structure, corresponding to two partially folded structures. One has two nascent helix fragments in the middle of the chain, and the other has a longer helix segment with 2.5 consecutive turns (data not shown). The specific heat C_v can be calculated from the thermal fluctuation of total energy, which peaks at the melting temperature. Under the AMBER03^{2D}p force field, the simulated C_v of 2I9M is 315 K.

As the AMBER03^{2D} force field can generate a considerable population of conformations with a turn structure, a polarization effect can help these structures achieve a native fold by strengthening the main chain hydrogen bonds in trpzip2. As expected, clustering analysis shows that the dominant cluster is a reservoir of folded structures, as shown in Fig. 4, which cover over 38% of conformations. The RMSD of the centroid structure is below 1 Å. Other clusters are archives of unfolded structures. A large populations of unfolded structures indicates that the native structure is not the exclusively sampled conformation for trpzip2 in aqueous solution at 300 K. The global free energy minimum conformation under AMBER03^{2D} is now only a local minimum. The global minimum under the AMBER03^{2D}p force field is located in the folded region. These two minima are separated by a low barrier of approximately 1 kcal/mol.

The folding pathway of trpzip2 is still an ongoing debate[40,56–65], because the simulated pathway relies on the force field that is employed. Whether the formation of inter-strand hydrogen bonds occurs prior to the formation of π-stacking is still unknown. We map the free energy landscape into the hydrogen bond distance d and the stacking distance l (see the section **Simulation Details** for definitions) as shown in Fig. 5. The experimental structures (shown as black circles) are very close to the global minimum of the free energy landscape. The unfolding of trpzip2 begins with the separation of the π-stacked tryptophan side chains. As can be read from the free energy landscape, this intermediate state has a longer stacking distance but a similar hydrogen bond distance to the native structure. The next step of unfolding is the breaking of the main chain hydrogen bonds of trpzip2, during which it crosses a free energy
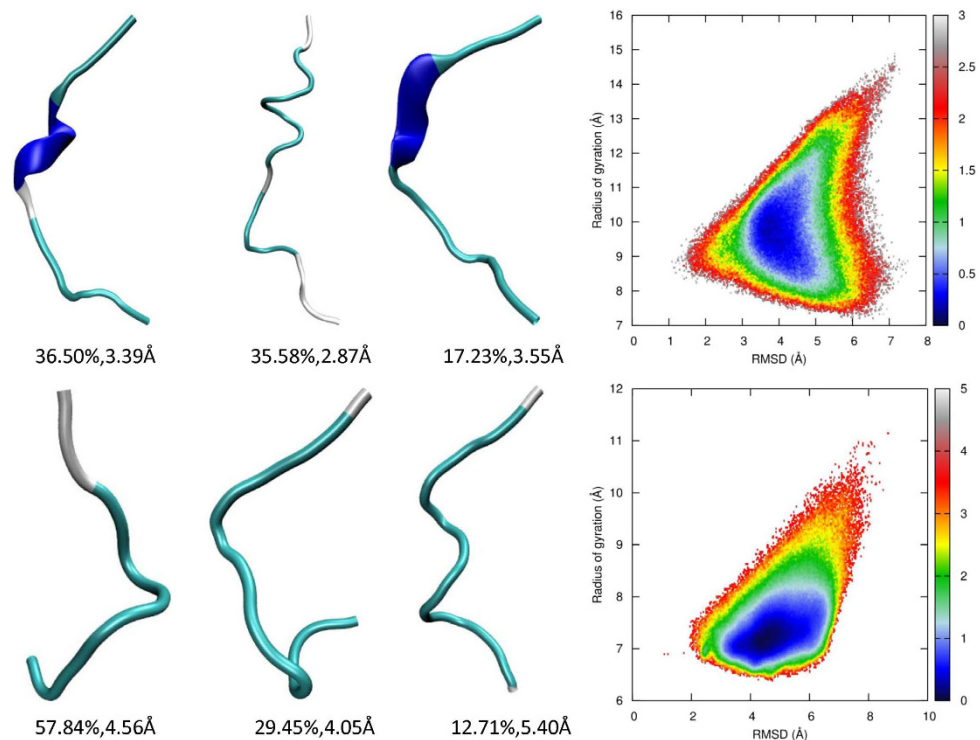
Figure 4. Representative structures from clustering analysis and the free energy landscape in the space of RMSD from the native structure and the radius of gyration for 2I9 M (top panel) and trpzip2 (bottom panel) using the AMBER03^{2D}p force field at 300 K, from the REMD simulations.
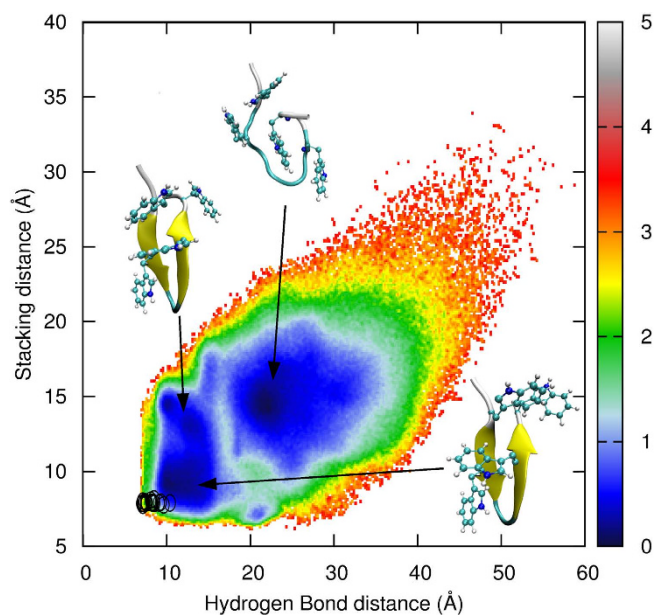
Figure 5. Free energy landscape of trpzip2 in the space of the length of inter-strand main chain hydrogen bonds (d) and the distance between stacking tryptophan side chains (l) at 300 K from the REMD simulation using the AMBER03^{2D}p force field. The locations of the experimental structures in the PDB file are shown as black circles.

barrier and enters the unfolded region. In the unfolded region, trpzip2 can adopt certain conformations with formed π-stacking but without the native main chain hydrogen bonds. However, if beginning from these π-stacking conformations, trpzip2 must surmount a barrier higher than that mentioned above
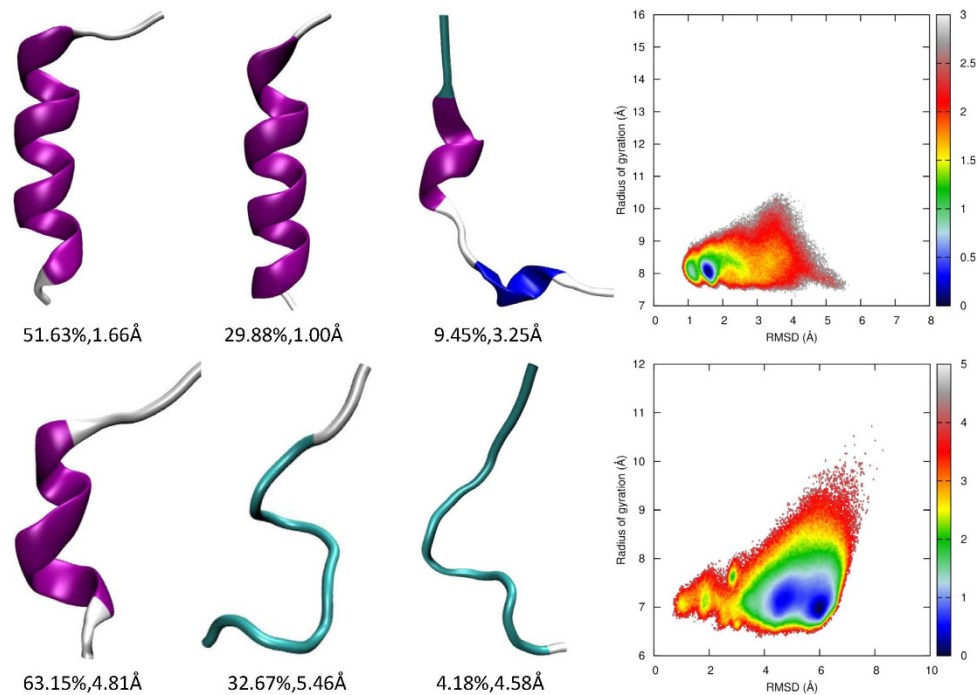
Figure 6. Representative structures from clustering analysis and the free energy landscape in the space of RMSD from the native structure and the radius of gyration for 2I9 M (top panel) and trpzip2 (bottom panel) using the AMBER03 force field at 300 K, from the REMD simulations.

before entering the folded region. Therefore, our REMD simulation employing the AMBER03^{2D}p force field supports the hydrogen bond-first mechanism of trpzip2 folding. The simulated C_v under the AMBER03^{2D}p force field is 335 K, which is close to the experimentally observed value of 345 K[47]. This 10 K difference can be explained by the underestimated viscous drag in the GB solvation model.

Comparative study of the AMBER03 and AMBER03p force fields.

The AMBER03 force field is biased toward helical conformations. It might not be difficult for protein 2I9M to fold under the AMBER03 force field. However, in our previous study[31], the simulated melting temperature of 2I9M under AMBER03 and the GB^{OBC} solvation model[66] was much lower than the experimentally measured value, despite the GB^{OBC} solvation model also being biased toward helical conformations[67]. In this study, we employ the GB^{MSMCO} solvation model[68], which is more balanced among conformations. As expected, peptide 2I9M cannot fold to its native state under the AMBER03 force field. The root mean square deviation (RMSD) from the native structure fluctuates between 2 and 7 Å. Clustering analysis of the folding trajectory at 300 K shows that, in the majority of trajectories, the peptide has only a helical nucleus and most of its residues are in nonstructured coil conformations, as shown in Fig. 6. This indicates that hydrogen bonds are not strong enough to compensate for the desolvation penalty and entropy loss during the growth of the helix. The conformations achieved are distributed very broadly along Rg and RMSD ranges. The free energy minimum has a less compact structure with a larger Rg (9.5 Å) than those of the native structures (Rg ≈ 8.57 Å). While employing the AMBER03p force field, the first two dominant conformation clusters are both reservoirs of folded structures that sum to over 80% of the achieved conformations, as shown in Fig. 7, which illustrates that polarization effectively shifts the conformation distribution of this peptide to its native structure. The free energy landscape shows that peptide 2I9M covers a much narrower phase space under AMBER03p than it does under the pairwise AMBER03 force field. The Rg of the minimal free energy conformation is very close that of the native structure. Therefore, the AMBER03p force field is very effective at folding this short helical peptide, and the electrostatic polarization effect plays a critical role in maintaining the secondary structure.

As the AMBER03 force field is biased toward helical conformations, the simulated folding of a protein in a β conformation is a challenge to this force field. Conformation clustering analysis, which shows that trpzip2 adopts mainly nonstructured conformations, clearly indicates that the native conformation of trpzip2 is not the preferred conformation under the AMBER03 force field at 300 K, as shown in Fig. 6. Turning on the polarization effect does not improve the result. Indeed, conformations with a short helical fragment appear as observed in Fig. 7. This failure is consistent with the previous observations that

Figure 7. Representative structures from clustering analysis and the free energy landscape in the space of RMSD from the native structure and the radius of gyration for 2I9M (top panel) and trpzip2 (bottom panel) using the AMBER03p force field at 300 K, from the REMD simulations.
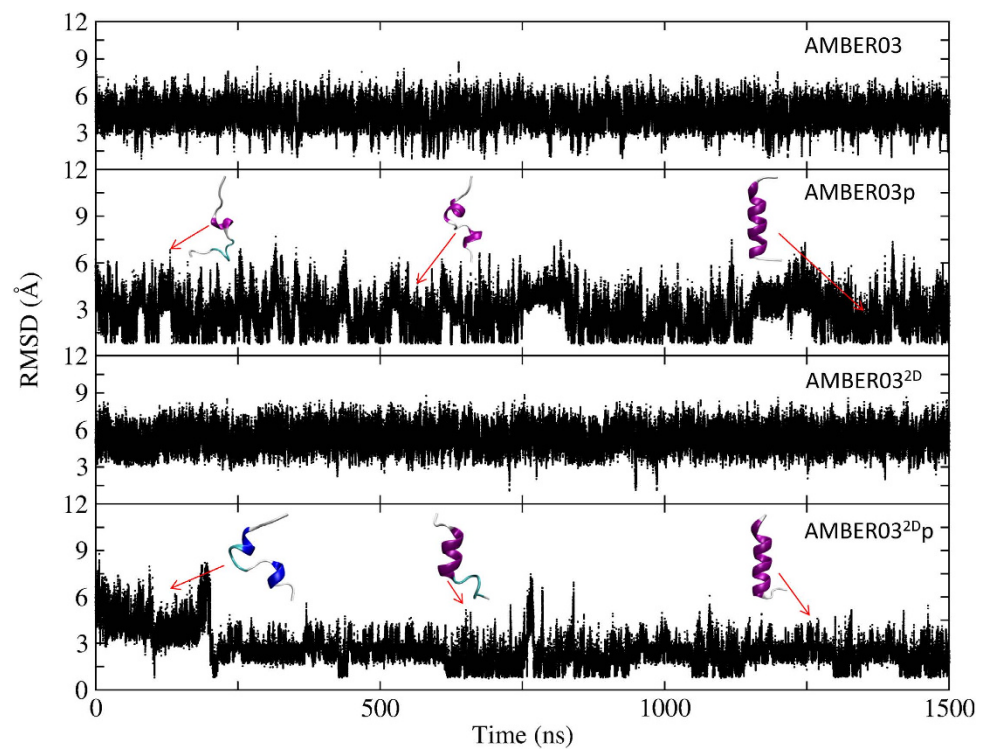
Figure 8. Time evolution of the backbone RMSD from the native structure at 300 K for 2I9M, calculated from direct molecular dynamics simulations at room temperature using the AMBER03, AMBER03p, AMBER03^{2D}, and AMBER03^{2D}p force fields, respectively.
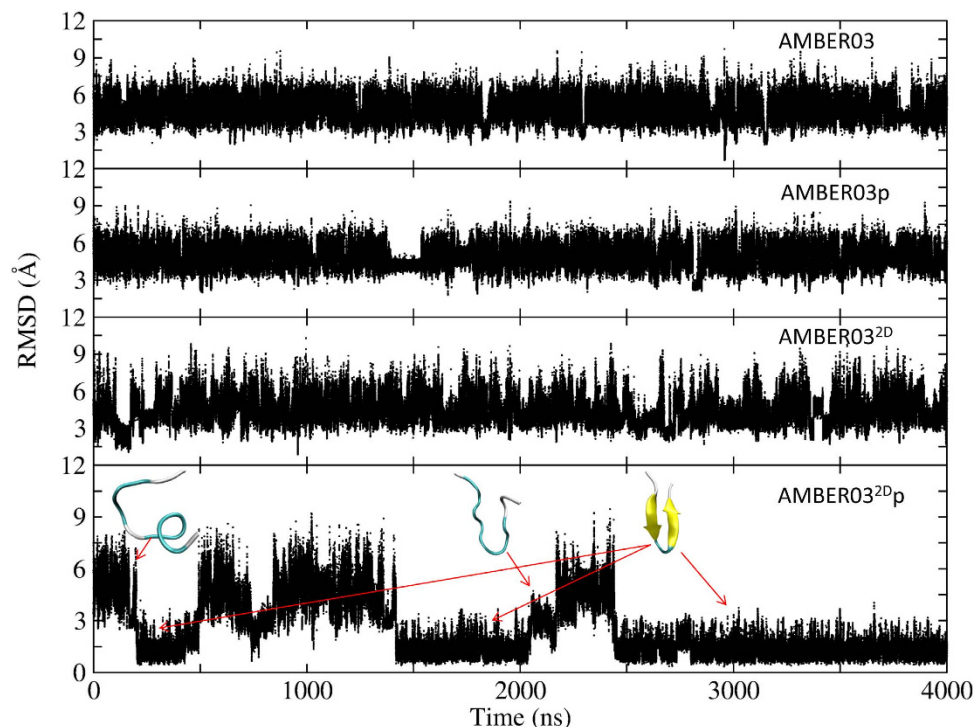
Figure 9. Time evolution of the backbone RMSD from the native structure at 300 K for trpzip2, calculated from direct molecular dynamics simulations at room temperature using the AMBER03, AMBER03p, AMBER03^{2D}, and AMBER03^{2D}p force fields, respectively.

AMBER03 is biased toward helical conformations[15], and that the polarization effect further stabilizes misfolded structures. There is only one minimum in the free energy landscape under the AMBER03 potential, as shown in Fig. 6. Polarization splits this free energy well into two not-well-separated minimums that are still located in the unfolded zone (see Fig. 7).

Direct MD simulation of 2I9M and trpzip2.   Direct molecular dynamics simulations that do not explore the entire phase space are also performed for 2I9M and trpzip2 to verify the results of REMD simulations and elucidate the folding mechanism. For 2I9M, the result is depicted in Fig. 8. Starting from a linear structure, the peptide does not fold to its native structure when using the AMBER03 or AMBER03^{2D} force fields. The average RMSDs are approximately 4.3 Å and 5.3 Å, respectively. When the polarization effect is incorporated, the peptide successfully folds to its native state with an RMSD distribution that peaks at 1.6 Å and 1.2 Å, respectively, for AMBER03p and AMBER03^{2D}p force fields. Moreover, the peptide under the AMBER03p force field reaches a folded conformation earlier than under the AMBER03^{2D}p force field, which is as expected due to the reduced helical propensity of this two-dimensional backbone potential. The representative structures under the AMBER03^{2D}p force field reveal that the nucleation of the helix occurs near the termini and gradually extends toward the central residues. Finally, the peptide reaches its native state.

The peptide trpzip2 also cannot fold into its native structure under AMBER03, AMBER03p, and AMBER03^{2D} force fields, with an average RMSD of 4.87, 4.91, and 4.44 Å, respectively, as shown in Fig. 9. However, under the AMBER03^{2D}p force field, the peptide reaches a folded state for the first time at approximately 200 ns with a backbone RMSD below 1.6 Å. Residing in the folded structure for approximately 300 ns, the peptide unfolds and the RMSD increases to approximately 8.0 Å. The folded structure appears again at 1.4 μs and remains there for 0.6 μs. After another 0.5 μs in the unfolded state, this peptide stays in its folded structure until the end of the simulation. The representative structures in the folding process show that the central residues adopt either bended structures or β-turn structures, which is the direct consequence of the secondary structure preference of the AMBER03^{2D} force field and the polarization effect. However, β strands are not very stable. This result is quite consistent with the REMD simulation result showing that the hairpin is the dominant but not unique structure adopted by this peptide. This result is also consistent with the REMD simulation result showing that the folded state is preferred under the AMBER03^{2D}p force field. The movies for 2I9M and Trpzip2 folding from direct molecular dynamics simulations at room temperature using the AMBER03^{2D}p force field are shown in the Supplementary Information.

Conclusion

The power of molecular dynamics simulations of proteins is limited by the accuracy and reliability of the force fields used. An unbalanced secondary structure propensity leads to failure in protein folding simulations. Moreover, the lack of an explicit polarization effect in pairwise force fields impedes the occurrence of ordered secondary structures. Recently, we optimized the AMBER force fields by replacing the uncoupled main chain torsion energy with a set of two-dimensional Fourier expansions, of which the parameters are fitted to the potential energy map resulting from high level quantum mechanical calculations of an alanine dipeptide[39]. Other force field parameters are kept intact as in the original AMBER force field. We name these new force fields AMBER99SB[2D] and AMBER03[2D]. Their performance in producing the secondary structure distribution of an alanine dipeptide has been investigated in a previous study. Although the accuracy of this force field for this model peptide is guaranteed, its application to other dipeptides and long peptide chains with ordered structures may not be straightforward. This is not surprising because main chain torsions are coupled to side chains, and the electrostatic polarization effect along main chain hydrogen bonds does not emerge during parameterization. Specific torsion parameters for each residue and the potential refitting of $\chi 1$ and $\chi 2$ torsions are necessary.

In this work, folding simulations were performed for two peptides adopting either an α-helix or a β-hairpin in their native structures, although only limited improvement was observed in residues other than ALA. The AMBER03 force field was used for comparison. Both replica exchange and direct molecular dynamics simulations were conducted, and their results are consistent. In REMD simulations, high temperature replicas can easily surmount the energy barrier and explore the phase space more efficiently than standard molecular dynamics at room temperature. The intrinsic helical propensity of the AMBER03[2D] force field is very low. Under this force field, peptide 2I9M primarily adopts random structures. No nucleation event was detected for the majority of its trajectory. For trpzip2, the preference of AMBER03[2D] for extended structures leads to a large population of bended structures. The accumulation of this conformation is essential for folding. However, further advancement to the native structure has not been observed. Therefore, nonadditive effects must play a critical role in the folding of ordered secondary structures. These nonadditive effects can be captured with the AMBER03[2D]p force field, in which the native structures of both peptides dominate, indicating that this force field is well balanced between helical and β conformations and that the electrostatic polarization effect is indispensable.

The AMBER03 force field is biased toward helical structures, but the folding simulation of a short helical peptide, 2I9M, using the AMBER03 force field still fails. With the polarization effect enabled, this peptide successfully folds into its native structures as the global minimum in the free energy landscape. Folding a peptide with β-strands poses a challenge to the AMBER03 force field due to its excessive helical tendency. Only disordered structures can be observed in a simulation of trpzip2-folding employing the AMBER03 force field. After enabling the polarization effect, structures with short helical fragments can be detected.

This study implies that both secondary structure propensity and electrostatic polarization play critical roles in determining the folding structure of a protein. The lack of either of these two effects during parameterization may cause inconsistencies in systems outside the training set. We adopt the tactic of fitting the secondary structure propensity according to high level quantum mechanical calculations of short model peptides and accounting for the non-pairwise effect in long peptide chains by using electrostatic polarization. The feasibility of this tactic and the reliability of this new force field have been justified in this study as well as in our previous work[39]. These studies pave the way for the future development of polarizable force fields.

Simulation Details. For dipeptides, tripeptides, and alanine tetrapeptide, the starting conformations used were linear structures generated using the LEaP module in AmberTools. The initial structure was immersed in the center of a truncated octahedral box of TIP3P water molecules, and all of the peptide atoms were no less than 12 Å from the boundary of the water box. To remove bad contacts before the simulation, we ran 20,000 steps of steepest descent followed by 20,000 steps of conjugate gradient energy minimizations. The relaxed structure was heated to its target temperature in 100 ps, during which all atoms in the protein were restrained with a force constant of 10 kcal/mol·Å². The target temperatures were chosen to match experimental conditions; dipeptides were held at 303 K, GXG tripeptides at 298 K, homotripeptides at 300 K, and alanine tetrapeptide at 300 K. All bonds with hydrogen atoms were fixed using the SHAKE algorithm[69]. The particle mesh Ewald method with a 10 Å cutoff in real space was used to calculate electrostatic interaction. A Langevin thermostat with a collision frequency of 1.0 ps^{-1} was used to regulate temperature[70]. Isotropic pressure coupling with a relaxation time of 2 ps was used to maintain the pressure to 1 atm. The integration time step was set to 2 fs. Trajectories were saved every 1 ps, and MD simulations were extended to 25 ns for each system (20 ns for dipeptides).

REMD simulations of 2I9M were performed with 12 replicas at 267.0, 283.0, 300.0, 328.0, 353.0, 380.0, 409.0, 439.0, 471.0, 506.0, 542.0, and 577.0 K. Those of trpzip2 were performed with the same number of replicas but at 255.00, 277.09, 300.63, 325.74, 352.50, 381.04, 411.51, 443.96, 478.65, 515.62, 555.04, and 597.08 K. All simulations started from a linear structure generated by the LEaP module in AmberTools. Each replica was heated to the target temperature in 100 ps after a local energy relaxation. The solvation effect was modeled using the generalized Born (GB) model developed by Mongan, Simmerling,

McCammon, Case, and Onufriev[68]. This GB model is believed to have a local conformational propensity close to that of the TIP3P water model[67]. The salt concentration was set to 0.2 M, and the integration time step used was 1 fs. Surface area was computed using the LCPO model[71]. All bonds with hydrogen atoms were fixed using the SHAKE algorithm. Nonbonded interactions were fully counted without any truncations. Temperature was regulated using Langevin dynamics with the collision frequency of $1 \, ps^{-1}$. Swaps were attempted every 0.25 ps. After 2-ns of equilibration, simulations were extended to 160 ns per replica for 2I9M and 400 ns for trpzip2. Snapshots were saved every 0.25 ps. Direct simulations were performed at 300 K, and the simulation durations used were 1.5 and $4 \, \mu s$ for 2I9M and trpzip2, respectively. Other control options were chosen as in the REMD simulations. Clustering analysis was performed by the MMTSB[72] package using the K-means clustering algorithm[73]. The MD simulations were performed using the AMBER11 Package with in-house modifications[74].

Two reaction coordinates are defined in the study of trpzip2-folding. One is the collective hydrogen bond distance d, which is calculated by

$$d = \sqrt{\sum_{i=1}^{6} d_i^2} \, , \tag{11}$$

in which d_i is the distance of the ith native hydrogen bond. This quantity is a metric of the formation of inter-strand hydrogen bonds. The other reaction coordinate is the stacking distance l defined as

$$l = \sqrt{l_1^2 + l_2^2} \, , \tag{12}$$

where l_1 and l_2 are the distance between TRP2 and TRP11 and that between TRP4 and TRP11, respectively. This coordinate is a metric of the stabilization force of π-stacking.

References

1. Sifers, R. N. Defective protein-folding as a cause of disease. *Nat. Struct. Biol.* **2,** 355–357 (1995).
2. Duan, Y. & Kollman, P. A. Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science* **282,** 740–744 (1998).
3. Slepoy, A. *et al.* Statistical mechanics of prion diseases. *Phys. Rev. Lett.* **87,** 058101 (2001).
4. Dobson, C. M. Experimental investigation of protein folding and misfolding. *Methods* **34,** 4–14 (2004).
5. Engelhard, M. & Evans, P. A. Experimental investigation of sidechain interactions in early folding intermediates. *Folding & Design* **1,** R31–R37 (1996).
6. Germann, H. P. & Heidemann, E. A synthetic model of collagen: an experimental investigation of the triple-helix stability. *Biopolymers* **27,** 157–163 (1988).
7. Shaw, D. E. *et al.* Atomic-Level Characterization of the Structural Dynamics of Proteins. *Science* **330,** 341–346 (2010).
8. Klepeis, J. L., Lindorff-Larsen, K., Dror, R. O. & Shaw, D. E. Long-timescale molecular dynamics simulations of protein structure and function. *Curr. Opin. Struct. Biol.* **19,** 120–127 (2009).
9. Freddolino, P. L., Liu, F., Gruebele, M. & Schulten, K. Ten-microsecond molecular dynamics simulation of a fast-folding WW domain. *Biophys. J.* **94,** L75–L77 (2008).
10. Shirts, M. R. & Pande, V. S. Solvation free energies of amino acid side chain analogs for common molecular mechanics water models. *J. Chem. Phys.* **122,** 134508 (2005).
11. Deng, Y. & Roux, B. Computations of Standard Binding Free Energies with Molecular Dynamics Simulations. *J. Phys. Chem. B* **113,** 2234–2246 (2009).
12. Buck, M., Bouguet-Bonnet, S., Pastor, R. W. & MacKerell, A. D. Importance of the CMAP correction to the CHARMM22 protein force field: Dynamics of hen lysozyme. *Biophys. J.* **90,** L36–L38 (2006).
13. Duan, Y. *et al.* A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **24,** 1999–2012 (2003).
14. Hornak, V. *et al.* Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins: Struct. Funct. Bioinform.* **65,** 712–725 (2006).
15. Best, R. B., Buchete, N.-V. & Hummer, G. Are Current Molecular Dynamics Force Fields too Helical? *Biophys. J.* **95,** 4494–4494 (2008).
16. Best, R. B. & Hummer, G. Optimized Molecular Dynamics Force Fields Applied to the Helix-Coil Transition of Polypeptides. *J. Phys. Chem. B* **113,** 9004–9015 (2009).
17. Best, R. B. *et al.* Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone phi, psi and Side-Chain chi(1) and chi(2) Dihedral Angles. *J. Chem. Theory Comput.* **8,** 3257–3273 (2012).
18. Lindorff-Larsen, K. *et al.* Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Struct. Funct. Bioinform.* **78,** 1950–1958 (2010).
19. Mackerell, A. D., Feig, M. & Brooks, C. L. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* **25,** 1400–1415 (2004).
20. MacKerell, A. D. *et al.* All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102,** 3586–3616 (1998).
21. Bjelkmar, P. *et al.* Implementation of the CHARMM Force Field in GROMACS: Analysis of Protein Stability Effects from Correction Maps, Virtual Interaction Sites, and Water Models. *J. Chem. Theory Comput.* **6,** 459–466 (2010).
22. Jorgensen, W. L., Maxwell, D. S. & TiradoRives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **118,** 11225–11236 (1996).
23. Kaminski, G. A., Friesner, R. A., Tirado-Rives, J. & Jorgensen, W. L. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B* **105,** 6474–6487 (2001).
24. Lindorff-Larsen, K. *et al.* Systematic Validation of Protein Force Fields against Experimental Data. *Plos One* **7,** e32131 (2012).

25. Guallar, V., Jarzecki, A. A., Friesner, R. A. & Spiro, T. G. Modeling of ligation-induced helix/loop displacements in myoglobin: Toward an understanding of hemoglobin allostery. *J. Am. Chem. Soc.* **128,** 5427–5435 (2006).

26. Ren, P. Y. & Ponder, J. W. Consistent treatment of inter- and intramolecular polarization in molecular mechanics calculations. *J. Comput. Chem.* **23,** 1497–1506 (2002).

27. Wang, Z. X. *et al.* Strike a balance: Optimization of backbone torsion parameters of AMBER polarizable force field for simulations of proteins and peptides. *J. Comput. Chem.* **27,** 781–790 (2006).

28. Meng, E. C., Cieplak, P., Caldwell, J. W. & Kollman, P. A. Accurate solvation free-energies of acetate and methylammonium ions calculated with a polarizable water model. *J. Am. Chem. Soc.* **116,** 12061–12062 (1994).

29. Ji, C. G., Mei, Y. & Zhang, J. Z. H. Developing polarized protein-specific charges for protein dynamics: MD free energy calculation of pK(a) shifts for Asp(26)/Asp(20) in thioredoxin. *Biophys. J.* **95,** 1080–1088 (2008).

30. Duan, L. L. *et al.* Folding of a Helix at Room Temperature Is Critically Aided by Electrostatic Polarization of Intraprotein Hydrogen Bonds. *J. Am. Chem. Soc.* **132,** 11159–11164 (2010).

31. Gao, Y. *et al.* Polarization of Intraprotein Hydrogen Bond Is Critical to Thermal Stability of Short Helix. *J. Phys. Chem. B* **116,** 549–554 (2012).

32. Gao, Y. *et al.* Direct folding simulation of a long helix in explicit water. *Appl. Phys. Lett.* **102,** 193706 (2013).

33. Wei, C. *et al.* Communication: The electrostatic polarization is essential to differentiate the helical propensity in polyalanine mutants. *J. Chem. Phys.* **134,** 171101 (2011).

34. Chen, Y. F. & Dannenberg, J. J. The Effect of Polarization on Multiple Hydrogen-Bond Formation in Models of Self-Assembling Materials. *J. Comput. Chem.* **32,** 2890–2895 (2011).

35. Li, Y., Ji, C. G., Xu, W. X. & Zhang, J. Z. H. Dynamical Stability and Assembly Cooperativity of beta-Sheet Amyloid Oligomers - Effect of Polarization. *J. Phys. Chem. B* **116,** 13368–13373 (2012).

36. Zaman, M. H. *et al.* Investigations into sequence and conformational dependence of backbone entropy, inter-basin dynamics and the flory isolated-pair hypothesis for peptides. *J. Mol. Biol.* **331,** 693–711 (2003).

37. Liu, Z., Ensing, B. & Moore, P. B. Quantitative Assessment of Force Fields on Both Low-Energy Conformational Basins and Transition-State Regions of the (phi-psi) Space. *J. Chem. Theory Comput.* **7,** 402–419 (2011).

38. MacKerell, A. D., Feig, M. & Brooks, C. L. Improved treatment of the protein backbone in empirical force fields. *J. Am. Chem. Soc.* **126,** 698–699 (2004).

39. Li, Y. X. *et al.* A Coupled Two-Dimensional Main Chain Torsional Potential for Protein Dynamics: Generation and Implementation. *J. Mol. Model.* **29,** 3647–3657 (2013).

40. Lwin, T. Z. & Luo, R. Force field influences in β-hairpin folding simulations. *Protein Sci.* **15,** 2642–2655 (2006).

41. Zhao, Y. & Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **120,** 215–241 (2008).

42. Gaussian 09, Revision B.01 (Gaussian, Inc., Wallingford, CT, 2010).

43. Tomasi, J., Mennucci, B. & Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **105,** 2999–3093 (2005).

44. Grdadolnik, J., Mohacek-Grosev, V., Baldwin, R. L. & Avbelj, F. Populations of the three major backbone conformations in 19 amino acid dipeptides. *Proc. Natl. Acad. Sci. USA* **108,** 1794–1798 (2011).

45. Hu, J.-S. & Bax, A. Determination of φ and χ1 Angles in Proteins from 13C−13C Three-Bond J Couplings Measured by Three-Dimensional Heteronuclear NMR. How Planar Is the Peptide Bond? *J. Am. Chem. Soc.* **119,** 6360–6368 (1997).

46. Cornell, W. D., Cieplak, P., Bayly, C. I. & Kollman, P. A. Application of RESP charges to calculate conformational energies, hydrogen-bond energies, and free-energies of solvation. *J. Am. Chem. Soc.* **115,** 9620–9631 (1993).

47. Cochran, A. G., Skelton, N. J. & Starovasnik, M. A. Tryptophan zippers: Stable, monomeric beta-hairpins. *Proc. Natl. Acad. Sci. USA* **98,** 5578–5583 (2001).

48. Shao, Q., Wei, H. & Gao, Y. Q. Effects of Turn Stability and Side-Chain Hydrophobicity on the Folding of beta-Structures. *J. Mol. Biol.* **402,** 595–609 (2010).

49. Nymeyer, H. How efficient is replica exchange molecular dynamics? An analytic approach. *J. Chem. Theory Comput.* **4,** 626–636 (2008).

50. Chodera, J. D. *et al.* Use of the weighted histogram analysis method for the analysis of simulated and parallel tempering simulations. *J. Chem. Theory Comput.* **3,** 26–41 (2007).

51. Kumar, S. *et al.* The weighted histogram analysis method for free-energy calculations on biomolecules.1. The method. *J. Comput. Chem.* **13,** 1011–1021 (1992).

52. Beauchamp, K. A., Lin, Y.-S., Das, R. & Pande, V. S. Are Protein Force Fields Getting Better? A Systematic Benchmark on 524 Diverse NMR Measurements. *J. Chem. Theory Comput.* **8,** 1409–1414 (2012).

53. Nerenberg, P. S. & Head-Gordon, T. Optimizing Protein−Solvent Force Fields to Reproduce Intrinsic Conformational Preferences of Model Peptides. *J. Chem. Theory Comput.* **7,** 1220–1230 (2011).

54. Hagarman, A. *et al.* Intrinsic Propensities of Amino Acid Residues in GxG Peptides Inferred from Amide I ' Band Profiles and NMR Scalar Coupling Constants. *J. Am. Chem. Soc.* **132,** 540–551 (2010).

55. Avbelj, F., Grdadolnik, S. G., Grdadolnik, J. & Baldwin, R. L. Intrinsic backbone preferences are fully present in blocked amino acids. *Proc. Natl. Acad. Sci. USA* **103,** 1272–1277 (2006).

56. Xiao, Y., Chen, C. & He, Y. Folding Mechanism of Beta-Hairpin Trpzip2: Heterogeneity, Transition State and Folding Pathways. *Int. J. Mol. Sci.* **10,** 2838–2848 (2009).

57. Chen, C. & Xiao, Y. Observation of multiple folding pathways of beta-hairpin trpzip2 from independent continuous folding trajectories. *Bioinformatics* **24,** 659–665 (2008).

58. Liu, Y., Kellogg, E. & Liang, H. J. Canonical and micro-canonical analysis of folding of trpzip2: An all-atom replica exchange Monte Carlo simulation study. *J. Chem. Phys.* **137,** 045103 (2012).

59. Nymeyer, H. Energy Landscape of the Trpzip2 Peptide. *J. Phys. Chem. B* **113,** 8288–8295 (2009).

60. Pitera, J. W., Haque, I. & Swope, W. C. Absence of reptation in the high-temperature folding of the trpzip2 beta-hairpin peptide. *J. Chem. Phys.* **124,** 141102 (2006).

61. Schlamadinger, D. E., Leigh, B. S. & Kim, J. E. UV resonance Raman study of TrpZip2 and related peptides: p-p interactions of tryptophan. *J. Raman Spectrosc.* **43,** 1459–1464 (2012).

62. Snow, C. D. *et al.* Trp zipper folding kinetics by molecular dynamics and temperature-jump spectroscopy. *Proc. Natl. Acad. Sci. USA* **101,** 4077–4082 (2004).

63. Song, J. *et al.* Investigating the Structural Origin of Trpzip2 Temperature Dependent Unfolding Fluorescence Line Shape Based on a Markov State Model Simulation. *J. Phys. Chem. B* **116,** 12669–12676 (2012).

64. Yang, L. J., Shao, Q. & Gao, Y. Q. Thermodynamics and Folding Pathways of Trpzip2: An Accelerated Molecular Dynamics Simulation Study. *J. Phys. Chem. B* **113,** 803–808 (2009).

65. Yang, W. Y., Pitera, J. W., Swope, W. C. & Gruebele, M. Heterogeneous folding of the trpzip hairpin: Full atom simulation and experiment. *J. Mol. Biol.* **336,** 241–251 (2004).

66. Onufriev, A., Bashford, D. & Case, D. A. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins: Struct. Funct. Bioinform.* **55,** 383–394 (2004).

67. Roe, D. R. *et al.* Secondary structure bias in generalized born solvent models: Comparison of conformational ensembles and free energy of solvent polarization from explicit and implicit solvation. *J. Phys. Chem. B* **111,** 1846–1857 (2007).

68. Mongan, J. *et al.* Generalized Born model with a simple, robust molecular volume correction. *J. Chem. Theory Comput.* **3,** 156–169 (2007).

69. Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. Numerical-integration of Cartesian equations of motion of a system with constraints: molecular-dynamics of n-alkanes. *J. Comput. Phys.* **23,** 327–341 (1977).

70. Uberuaga, B. P., Anghel, M. & Voter, A. F. Synchronization of trajectories in canonical molecular-dynamics simulations: Observation, explanation, and exploitation. *J. Chem. Phys.* **120,** 6363–6374 (2004).

71. Weiser, J., Shenkin, P. S. & Still, W. C. Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO). *J. Comput. Chem.* **20,** 217–230 (1999).

72. Feig, M., Karanicolas, J. & Brooks, C. L. MMTSB Tool Set: enhanced sampling and multiscale modeling methods for applications in structural biology. *J. Mol. Graphics Model.* **22,** 377–395 (2004).

73. Hartigan, J. A. & Wong, M. A. Algorithm AS 136: A K-Means Clustering Algorithm. *Appl. Stat.* **28,** 100–108 (1979).

74. AMBER11 (University of California, San Francisco, CA, U. S. A., 2010).

Acknowledgments

Author Contributions

Y.G. wrote the main text of the manuscript and prepared all the figures. Y.G. and Y.X.L. performed molecular dynamics simulations. L.R.M. and B.B.L. helped with data analysis. Y.M. and J.Z.H.Z. designed this study. All authors revised the manuscript.

Additional Information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Gao, Y. *et al.* Correct folding of an α-helix and a β-hairpin using a polarized 2D torsional potential. *Sci. Rep.* **5**, 10359; doi: 10.1038/srep10359 (2015).