Helix-Coil Transition of Alanine Peptides in Water: Force Field Dependence on the Folded and Unfolded Structures

S. Gnanakaran[1] and Angel E. García[1,2]*

[1]*Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, New Mexico*

[2]*Department of Physics, Applied Physics, and Astronomy, Rensselaer Polytechnic Institute, Troy, NY 12180*

ABSTRACT **The force fields used in classical modeling studies are semiempirical in nature and rely on their validation by comparison of simulations with experimental data. The all-atom replica-exchange molecular dynamics (REMD) methodology allows us to calculate the thermodynamics of folding/unfolding of peptides and small proteins, and provides a way of evaluating the reliability of force fields. We apply the REMD to obtain equilibrium folding/unfolding thermodynamics of a 21-residue peptide containing only alanine residues in explicit aqueous solution. The thermodynamics of this peptide is modeled with both the OPLS/AA/L and the A94/MOD force fields. We find that the helical content and the values for the helix propagation and nucleation parameters for this alanine peptide are consistent with measurements on similar peptides and with calculations using the modified AMBER force field (A94/MOD). The nature of conformations, both folded and unfolded, that contributes to the helix-coil transition profile, however, is quite different between these two force fields. Proteins 2005;59:773–782.** © 2005 Wiley-Liss, Inc.

Key words: folding; explicit solvent; replica exchange molecular dynamics; trialanine; poly-proline like (PP_{II}); OPLS; AMBER; Lifson-Roig; Zimm-Bragg; free energy

INTRODUCTION

Protein folding, a process by which a linear chain of amino acids comes together to form a three-dimensional structure, is still an unsolved problem in biophysics.[1–3] Now, the availability of computational resources such as Linux Beowulf clusters and distributed computing,[4] and experimental measurements on fast-folding proteins and peptides, have made atomic simulations an integral part of solving the folding problem.[5] The folding time scales measured in T-jump experiments, and the time scales sampled in all-atom MD simulations are similar, and present an excellent scenario for direct comparison.[6,7] Such a synergy between experiment and theory provides a structural model to interpret and explain experimental results and to validate simulation methodologies and predictions. However, a major obstacle for such direct comparison is the reliability of the force fields used in atomic simulations.

Currently, computational modeling utilizes a variety of force fields such as AMBER,[8,9] CHARMM,[10] GROMOS,[11] OPLS,[12] and MMFF94.[13] It is well known that each of these force fields has their own strengths in reproducing observations due to the data that were specifically used to parameterize them. Consequently, discrepancies in simulation results due to the choice of force fields has been well documented.[14–18] The folding of the C terminus of the beta-hairpin of protein G has been a subject of many simulations that utilize a variety of force fields with both implicit and explicit solvent representations.[4,15,16,19–23] For example, explicit solvent simulations of beta-hairpin utilizing the AMBER94 force field produce helix intermediates,[19] whereas those that utilize the OPLS/AA force field do not.[23] It is not straightforward to compare results obtained with different force fields because discrepancies in results may depend on other factors such as incomplete conformational sampling, system size, implicit/explicit treatment of solvent,[24] simulation conditions, the treatment of long-range electrostatics, or unavailability of conclusive structural and thermodynamic experimental information.

It is essential, therefore, to treat simulation factors other than the force field accurately, and to keep them consistent so that any observed bias could be attributed to the force field. Utilizing such an approach, two popular versions of the AMBER force field, AMBER94[9] and AMBER96,[8] were evaluated for reproducibility of helix-coil transitions of alanine (Ala)-based peptides with replica exchange molecular dynamics (REMD) in explicit solvent.[25] That study illuminated different secondary structural bias between these two force fields and provided avenues for improvement. A minor modification to AMBER force field (A94/MOD) was suggested based on reproducibility of experimental values for folding of helical peptides.[25] A94/MOD completely eliminates the bias for the backbone dihedral terms in AMBER94. The A94/MOD force field reproduced experimental observations of a shift in secondary structural preferences with increasing lengths of peptide size.[26] Importantly, that validation study illustrated that the AMBER force field is already parameterized reasonably well, and requires only minor refinements.

Grant sponsors: U.S. DOE and LANL LDRD funds.

*Correspondence to: Angel E. García, Theoretical Biology and Biophysics Group, T-10 MS K710, Los Alamos National Laboratory, Los Alamos, NM 87545. E-mail: axg@lanl.gov

Received 17 August 2004; Accepted 6 December 2004

Published online 6 April 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20439

© 2005 WILEY-LISS, INC.

Other studies of the AMBER and CHARMM force fields also suggest that refinement may involve only the backbone torsional energy parameters.[27,28]

In this study, we considered the OPLS/AA/L force field as a continuation of our efforts to test and validate force fields in model proteins. This work was motivated by past studies that have suggested that A94/MOD and OPLS/AA/L[12,29] should reproduce experimental thermodynamic values based on the observations that they favored less helical conformations compared to other force fields.[14,30] The original OPLS force field was based on united atoms,[31] and was followed by the all-atom version, OPLS/AA.[12] Recently, a refined version of OPLS/AA force field, OPLS/AA/L, for peptides and proteins developed based on high-level ab initio calculations that was used to refine the OPLS/AA torsional parameters belonging to the backbone and to the side chains of amino acids.[29] This is the version of the OPLS force field that is considered in our calculations.

Another motivation for the current study is concerned with the inconsistency of the parameterization when applied with explicit solvent models versus implicit solvent models based on the Generalized Born (GB) approach.[24,32,33] Recently, it was demonstrated that OPLS/AA/L led to improved side chain predictions with the rest of the protein held fixed at the native configuration in implicit solvent.[34] OPLS/AA force fields with a GB solvent model have been used to fold polypeptides,[35] and have provided an effective free energy function that was used as a scoring function to detect native protein folds among decoys.[36] Therefore, it is imperative to consider the transferability of this force field to folding of peptides in explicit solvent.

In a force field validity study such as this, three factors need special consideration: the system, solvent, and sampling. The system considered is the 21-residue alanine peptide, A21. This choice of systems provides the same control environment as previous studies so that the results can be compared to AMBER force fields and to measurements in similar sequences.[37] A21 is not soluble in water, but peptides consisting of alanine blocks (of length 7–13) flanked by charged amino acids (ornithine) have been studied experimentally.[37] The reliability of the OPLS/AA force field will be evaluated based on how well it describes the helix-coil transition of the A21 peptide, relative to other force fields and to model systems. The effect of hydration is taken into account by an explicit solvent model. The computational challenge is to capture the relevant structural ensembles of the peptide during the helix-coil transition, and requires an efficient conformational sampling technique that can gather the appropriate canonical population from the free energy landscape of the peptide. The REMD methodology, which has emerged as an efficient technique to sample the conformational space of short peptides in explicit solvent, offers an excellent platform to validate force fields.[19,25,26,38–41]

MATERIALS AND METHODS

Simulation Details

The A21, $(Ala)_{21}$, peptide (Ac-A_{21}-Nme) was blocked with acetyl and amino groups at the N- and C-terminus, respectively. This A21 peptide was simulated in explicit aqueous solution under periodic boundary conditions. The dimension of the solvated cubic box was 43.41 Å. The peptide was solvated with 2640 water molecules using the TIP3P model.[42] A time step of 2 fs was used. All bonds involving hydrogen atoms were constrained by using SHAKE with a tolerance of 0.0005 Å.[43] The suite of programs in AMBER 4.1[43] was modified to incorporate REMD.[25,44,45]

The REMD was implemented with 48 replicas at constant volume and a fixed number of atoms, and covered a temperature range of 278 to 502 K in intervals of 3–8 K. The temperatures of the replicas were chosen to maintain an exchange rate among replicas of 8–20%.[25] The system was coupled to an external heat bath with relaxation time of 0.1 ps.[46] The generalized reaction field treatment (GRF)[47] was used for electrostatic interactions with a cutoff of 8 Å. The 1,4 Lennard-Jones interactions were not scaled in simulations with A94/MOD, and were scaled by 0.5 in the simulations with OPLS/AA/L. Nonbonded interactions were updated every 10 integration steps.

All replicas were started from random conformations, and no structural biases were imposed in the sampling. Exchanges were attempted every 125 integration steps. Each replica was simulated for 8.6 ns, yielding a total sampling time of 412.8 ns. The production run was considered to be the last 5 ns/replica. Block averages over 1 ns were performed to estimate errors in the ensemble averages.

We also carried out REMD simulations on the following capped peptide systems in explicit TIP3P water; $(Ala)_3$ in 523 H_2O, and $(Ala)_7$ in 851 H_2O. The cubic box size was chosen to reproduce the correct density at 1 atm and 300 K. The REMD was implemented with 24 replicas with a temperature range of 276 to 469 K. Each replica was simulated for 10 ns, yielding a total sampling time of 240 ns. The production run was considered to be last 6 ns/replica, and error analysis was done with 1.5 ns/block averaging. The simulation for $(Ala)_7$ has also been carried out in a larger box with 2553 water molecules and A94/MOD force field in an effort to detect any box size effects on the probability of forming PP_{II}-like conformations. The calculated PP_{II} content was found to be weakly dependent of the box size, and the propensity to form PP_{II} increased by ~17% with a smaller box.

The OPLS/AA parameter and topology files[12] were obtained from Professor Jorgensen. The dihedral terms were corrected to incorporate the recent modifications suggested for amino acids.[29] The combining rules for the Lennard-Jones (LJ) parameters (ϵ and σ) were modified from the arithmetic mean combination rule used in AMBER to the geometric mean combination rule used in OPLS. To ensure convergence of the calculated averages, we extended the length of each replica simulation to be at least three limes longer than the equilibrium time needed

for the helical content to approach equilibrium exponentially.[48] During the REMD simulations all replicas explored the full range of temperatures at least once. Over the production segment of the simulation calculated block averages (over 1 ns per block) have similar values for helical content.

Calculation of Secondary Structural Content

Within the Lifson-Roig (LR) and Zimm-Bragg (ZB) models for the helix-coil transition, the intrinsic propensity of an amino acid to form a helix is a measure of the interactions of the amino acid with its own and the nearest-neighbor amino acid backbone.[49,50] We analyzed configurations generated by the REMD simulations in terms of backbone (ϕ, ψ) dihedral angles to obtain helical content that requires three intervening residues to be in the helical region [(ϕ, ψ) = (-65 ± 35, -37 ± 30)]. Acetylated and amidated peptides with a 21-amino acid sequence can have a maximum helical length of 19. In a similar way, the PP_{II} content is defined as the number of three consecutive residues found in the PP_{II} basin [($-105 < \phi < -30$) and ($60 < \psi < 180$)]. To minimize the artifacts due to dihedral angle boundaries, the ranges are chosen to cover most of the basin at low T.

Calculation of LR and ZB Parameters

The statistical mechanical models of LR and ZB characterize the helix-coil transition in terms of a helix nucleation parameter and a helix propagation parameter.[49,50,51] The fractional helical content and helical segments can be related to the partition function to obtain these parameters in the LR model. The details of this theory can be found elsewhere.[50,51] The partition function, Z, is given as

$$Z = (0\ 0\ 1)M^n\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \tag{1}$$

where the weight matrix M is

$$M = \begin{pmatrix} w & v & 0 \\ 0 & 0 & 1 \\ v & v & 1 \end{pmatrix}.$$

The number of residues is designated as n and it is the same as the number of amino acids for a capped peptide. w and v are the propagation parameter and the equilibrium constant for the helical conformation in the random coil, respectively. The nucleation parameter will be proportional to v^2, because it will require restriction of two residues before it can be compensated by a hydrogen bond. The parameter v is considerably less than unity due to low entropy requirement, and is assumed to be independent of temperature. The average fractional helical content is

$$\langle n_H \rangle = \frac{1}{(n-2)}\frac{w}{Z}\frac{\partial Z}{\partial w}, \tag{2}$$

and the average helical segment is

$$\langle n_s \rangle = \frac{v_{12}}{Z}\frac{\partial Z}{\partial v_{12}}, \tag{3}$$

where v_{12} is the value at row 1, column 2 of the weight matrix. At each temperature, the values for w and v that reproduce the calculated mean values for fractional helical content and segments are fitted using Equations (2) and (3).

The ZB model characterizes helix formation according to hydrogen bond formation of its peptide NH group. The ZB parameter s is the propagation equilibrium constant and σ is the nucleation parameter that represents the barrier to helix initiation. In the ZB model, the first helical hydrogen bond is given a weight of $\sigma^* s$, and the subsequent hydrogen bonds are given a weight of s. The ZB parameters are related to the LR parameters by[51]

$$s = \frac{w}{1+v};\ \sigma = \frac{v^2}{(1+v)^4} \tag{4}$$

Evaluation of the Free Energy

The free energy of folding a helix is associated with spatial restriction of three residues that requires reduction in entropy and compensation by a gain in enthalpy due to the formation of hydrogen bonds when adding peptide units to an existing helix.[52] The temperature dependence of the ZB helix propagation parameter can be fit to a thermodynamic model to obtain free energies [$\Delta G(T)$]. We fit s to the following functional form to extract the temperature dependent enthalpy [$\Delta H(T)$] and entropy [$\Delta S(T)$] changes and a temperature-independent heat capacity (ΔC_p) as follows

$$\begin{aligned}\Delta G(T) &= -RT\log s \\ &= \Delta H(T) - T\Delta S(T) \\ &= \Delta H_0 + \Delta C_p[T - T_0] - T\Delta S_0 - \Delta C_p T\log\left(\frac{T}{T_0}\right)\end{aligned} \tag{5}$$

where $\Delta H_0 = \Delta H(T_0)$ and $\Delta S_0 = \Delta S(T_0)$, for an arbitrary reference temperature T_0 (the lowest temperature used in the fitting).

RESULTS AND DISCUSSION

We compare the helix-coil thermodynamics of the A21 peptide described by the OPLS/AA/L force field with measurements. Figure 1 shows the Lifson-Roig (LR)-based average fractional helical content, and number of helical fragments as a function of temperature (T) for the A21 peptide. According to the LR definition,[50] the helical content is expressed in terms of helix backbone parameters, and involves restricting backbone dihedrals (ϕ, ψ) of three intervening residues to the helical region of the Ramachandran Plot. This peptide contains 41% helical content at 278 K. The pure Ala peptide is not accessible to the experiment due to its limited solubility. However, Ala_n ($n = 13$) flanked by two ornithine charged amino acids are found to be about 40% helical in water at 273 K.[37] The calculated helical content is in quantitative agreement with the above measurement.

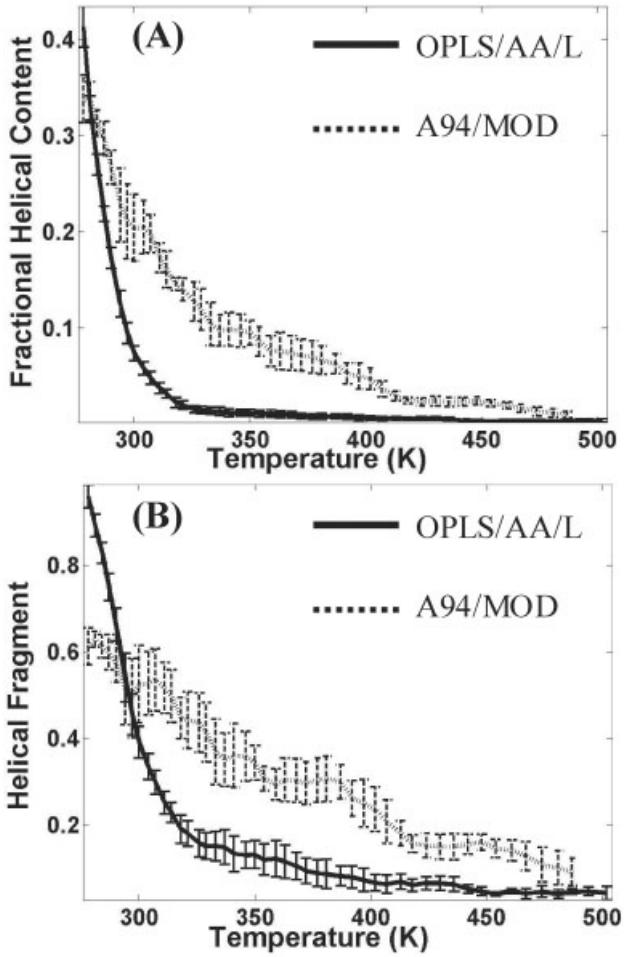

Fig. 1. The average fractional helical content (**a**) and the average number of helical segments (**b**) of A21 peptide as a function of temperature over the entire simulated temperature range. The solid line shows results obtained with the OPLS/AA/L force field. The results obtained with A94/MOD force field (dashed line) from an earlier study[25] is also provided for comparison. Error bars are calculated by block averaging as explained in the Methodology section.

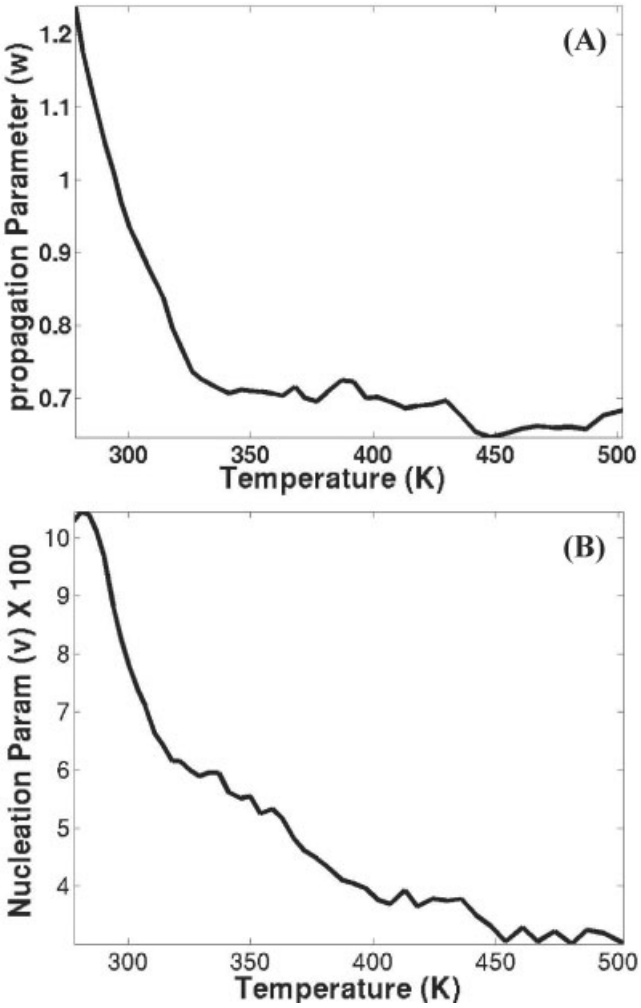

Fig. 2. The Lifson-Roig helix propagation (**a**) and nucleation (**b**) parameters for A21 peptide as a function of temperature.

Next we consider how well the OPLS/AA/L force field reproduces the measured nucleation and propagation parameters for helix formation. The LR helix propagation (w) and nucleation parameters (v) for the A21 peptide with the OPLS/AA/L force field are found to be $w = 1.24$ and $v = 0.103$ at 278 K. The fractional helical content and helical segments were related to the partition function to obtain these parameters in the LR model. The ratio of w/v provides the strength of an effective hydrogen bond and is ~ -1.25 kcal/mol. Figure 2 shows the temperature dependence of LR parameters, an important dependency that is often neglected. The LR parameters can be related to Zimm-Bragg (ZB) model parameters, which are classified according to hydrogen bond formation of its peptide NH group.[51] At 278 K the corresponding ZB parameters are, $s = 1.12$ and $\sigma = 0.007$, where s is the propagation equilibrium constant and σ is the nucleation parameter that represents the barrier to helix initiation. The value for σ is close to that of Yang et al. ($\sigma = 0.004 \pm 0.002$ for fixed values of $s = 1.4$ to 1.7)[53] and in disagreement by one order of magnitude with that of Platzer et al. ($s = 1.08$ at 273 K for fixed value σ of 0.0008 ± 0.0002).[54]

Most of the current protein force fields do not yield reasonable nucleation parameters that are close to the measured range as OPLS/AA/L is. Several published studies have estimated the ZB model parameters for solvated Ala peptides with a broad range of modeling sophistication and force fields.[24,25,30,55–57] Calculated σ values for various lengths of Ac-$(Ala)_n$-Nme peptides using REMD with the CHARMM/GB force field were an order of magnitude higher than the largest value reported from experiments.[30] Multicanonical Monte Carlo simulations of $(Ala)_{10}$ in explicit solvent with the ECEEP/2 force field obtained a large value (0.2) for σ.[56] With AMBER94, a large σ was found for the A21 peptide with the implicit and explicit solvent models.[24,25] With the explicit solvent, AMBER94 yielded a nucleation constant of 0.06.[25] However, A94/MOD reproduced the observed nucleation and propagation parameters of $\sigma = 0.004$ and $s = 1.30$ for A21 in explicit aqueous solution.[25] In Figure 1, for comparison, average

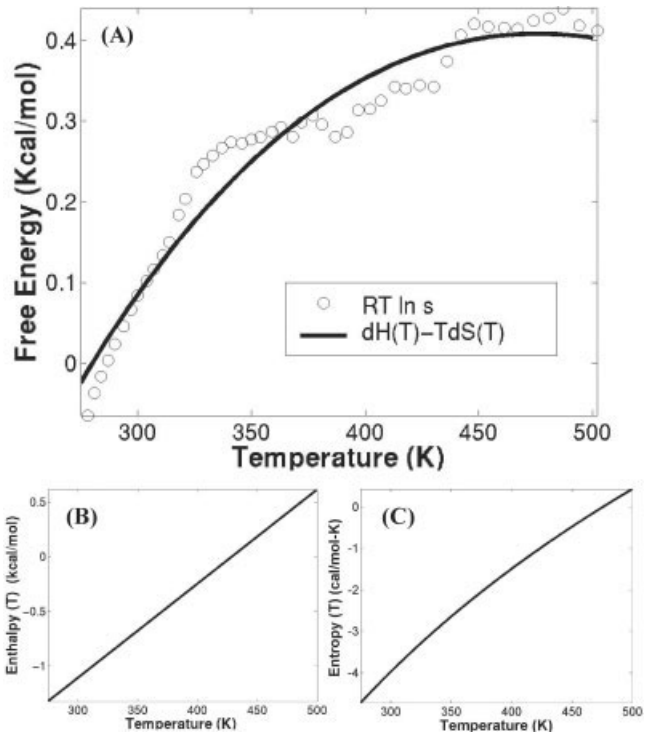

Fig. 3. Plot shows (**a**) the free energy of conversion of a coil alanine residue to a helical one at the end of long helix at 300 K. The temperature-dependent enthalpy (**b**) and entropy (**c**) differences are also shown. These were obtained by fitting the temperature dependence of the ZB helix propagation parameter to a thermodynamic model given in Equation (5).

fractional helical content and number of helical fragments obtained with A94/MOD are plotted alongside those obtained with OPLS/AA/L. The current study, which is carried out under similar conditions as that of AMBER force field studies, confirms that A94/MOD and OPLS/AA/L force fields are capable of providing reasonable helix-coil transition parameters for poly-Ala peptides in explicit aqueous solution. Molecular dynamics simulations by Sorin and Pande found that the A94/MOD force field has much larger propensity to form helices than was found by us.[58] Their calculations used GROMACS[59] and a different treatment of electrostatics. We conducted calculations of an Ala-based peptide using Ewald sums and found higher melting temperature than for the GRF, but similar nucleation parameters. The origin of these discrepancies is under study.

The free energy [$\Delta G(T)$] of folding the A21 peptide with OPLS/AA/L is compared to the calorimetric measurements and to the values obtained with A94/MOD. The T dependence of the ZB parameter s is fitted to a thermodynamic model with the T-dependent enthalpy [$\Delta H(T)$] and entropy [$\Delta S(T)$] changes and a T-independent heat capacity (ΔC_p). Figure 3 shows the free energy and its fit to the ΔH and ΔS components. The calculated $\Delta G(300\ \mathrm{K}) = 0.085$ Kcal/mol, corresponds to the ΔG of conversion of a coil (nonhelical) Ala residue to a helical one at the end of long helix at 300 K.[54] We get $\Delta H(300\ \mathrm{K}) = -1.1$ kcal/mol, $\Delta S(300\ \mathrm{K}) = -0.004$ kcal/(mol-K), and $\Delta C_p = 0.0086$ kcal/(mol-K). Assuming, $\Delta C_p = 0$, and fitting only up to 402 K, we obtain $\Delta H = -0.78$ kcal/mol, $\Delta S = -0.0029$ kcal/(mol-K). The calculated ΔH is close to the measured ΔH of -0.9 kcal/mol per residue (assuming $\Delta C_p = 0$) for a 50-residue peptide containing primarily Ala.[60] The calculated ΔS is close to the reported value of 0.0023 kcal/(mol-K) for helix unfolding.[61] The value and the sign of the calculated ΔC_p conform to the accepted range of -0.008 to $+0.008$ kcal/(mol-K) for helix formation.[62] Recent measurements based on direct determination of ΔH in a series of host peptides using differential scanning calorimetry reported $\Delta C_p = -0.0076$ kcal/(mol-K).[63] The ΔH and ΔS values are comparable to those obtained with A94/MOD force field.[25]

Even though both A94/MOD and OPLS/AA/L force fields predominantly produce similar helical population for A21 in reasonable agreement with the measurements, the structural composition of the conformations sampled along the helix-coil transition can be quite different. The conformational variability of A21 peptide with OPLS/AA/L is explored in terms of three properties; backbone (ϕ, ψ) dihedral angles of individual residues, end-to-end distances (d) and radius of gyration (R_g). The Ramachandran plot pertaining to the average (ϕ, ψ) is shown in Figure 4. It will be used to identify the three different well-defined conformations; α, β, and poly-proline like (PP_{II}).[64] In Figure 5, the R_g is plotted as a function of the end-to-end distances (d) at four different T. It is also useful to know the amount of helical structure within the distinct populations seen in these contour plots at low T. Therefore, bar graphs are provided at the top of the contour plots clarifying the contribution to the helical content from the individual d at respective T. The T dependence of d and R_g is used to characterize the overall structural attributes of conformations during the helix-coil transition.

First, we identify well-defined conformations that can exist at low T, which are not α-helical. With OPLS/AA/L, the A21 peptide predominantly samples α-helical conformations at 287 K. It also samples PP_{II} and β regions, with the latter exhibiting a slight preference [Fig. 4(a)]. Nonnegligible amounts of population are also found in the α' region ($-180 < \phi < -110$ and $-50 < \psi < 80$). With increasing T, the preference shifts predominantly to the β basin [Fig. 4(b) and (c)]. Even though the quantification of the non-α-helical populations at 300 K yields 13% PP_{II} population, no significant PP_{II} content ($<0.1\%$) is found. Whereas, $\sim 20\%$ PP_{II} content is found for the A21 peptide with A94/MOD at low T.[65]

Now, we explore whether the folded conformations sampled at low T by OPLS/AA/L and A94/MOD force fields are different. Two populations of distinct sizes dominate the distribution at low T [Fig.5(a)]; elongated conformations (R_g = 9–9.5 Å) with long d (25 Å) and compact conformations (R_g = 6.5 Å) with short d (10 Å). We see in Figure 1(a) that the fractional helical content drops from 0.41 to 0.28 when increasing temperature from 278 to 284 K. From Figure 5(a)–(b) it can be seen that the drop in helical content occurs for conformations that have long end-to-end distances. As T increases, major contribution to the helical content now arises from compact structures

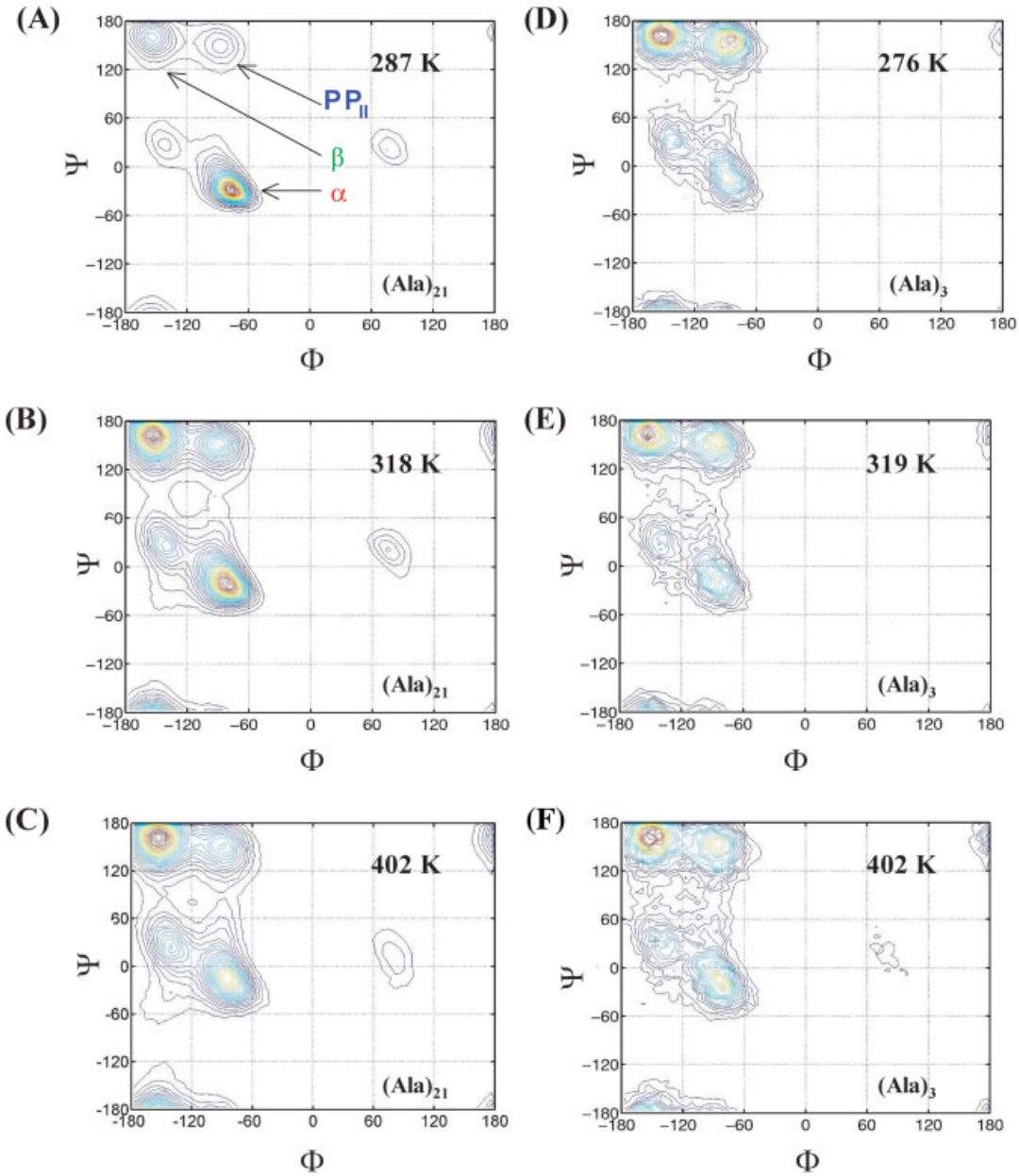

Fig. 4. The backbone (ϕ, ψ) conformational map of A21 peptide is shown at three different temperatures; 287 K (**a**), 318 K (**b**), and 402 K (**c**). The (ϕ, ψ) map of capped trialanine peptide, $(Ala)_3$, is shown at temperatures; 279 K (**d**), 319 K (**e**), and 402 K (**f**). Three well-defined conformations sampled are marked in the plot.

with short d. Comparison to the d distribution obtained with A94/MOD[65] illuminate two major differences; significant contributions to helical content at low T arises from conformations with d (35 Å) that are longer than those obtained with OPLS/AA/L and the existence of a well-defined peak at very long d (50–60 Å) corresponding to PP_{II} conformation. The dependence of the height of this peak on the simulation box size is not known.

The discrepancy in the PP_{II} content of $(Ala)_{21}$ between A94/MOD and OPLS/AA/L, and the availability of experimental evidence of PP_{II} content only in shorter peptides prompted the examination of the structural content of $(Ala)_3$ and $(Ala)_7$ with OPLS/AA/L force field under identical conditions permitting direct comparisons with those obtained with A94/MOD. Measurements using different experimental techniques on Ala peptides of lengths of 7 or less find PP_{II} as the dominant conformation and propose this conformation as a major distribution in unfolded or disordered proteins in aqueous solution.[17,66–71] The (ϕ, ψ) map of $(Ala)_3$ obtained using OPLS/AA/L [Fig. 4(d)–(f)] shows four basins; α, β, PP_{II}, and α′. At low T (276 K), predominantly extended

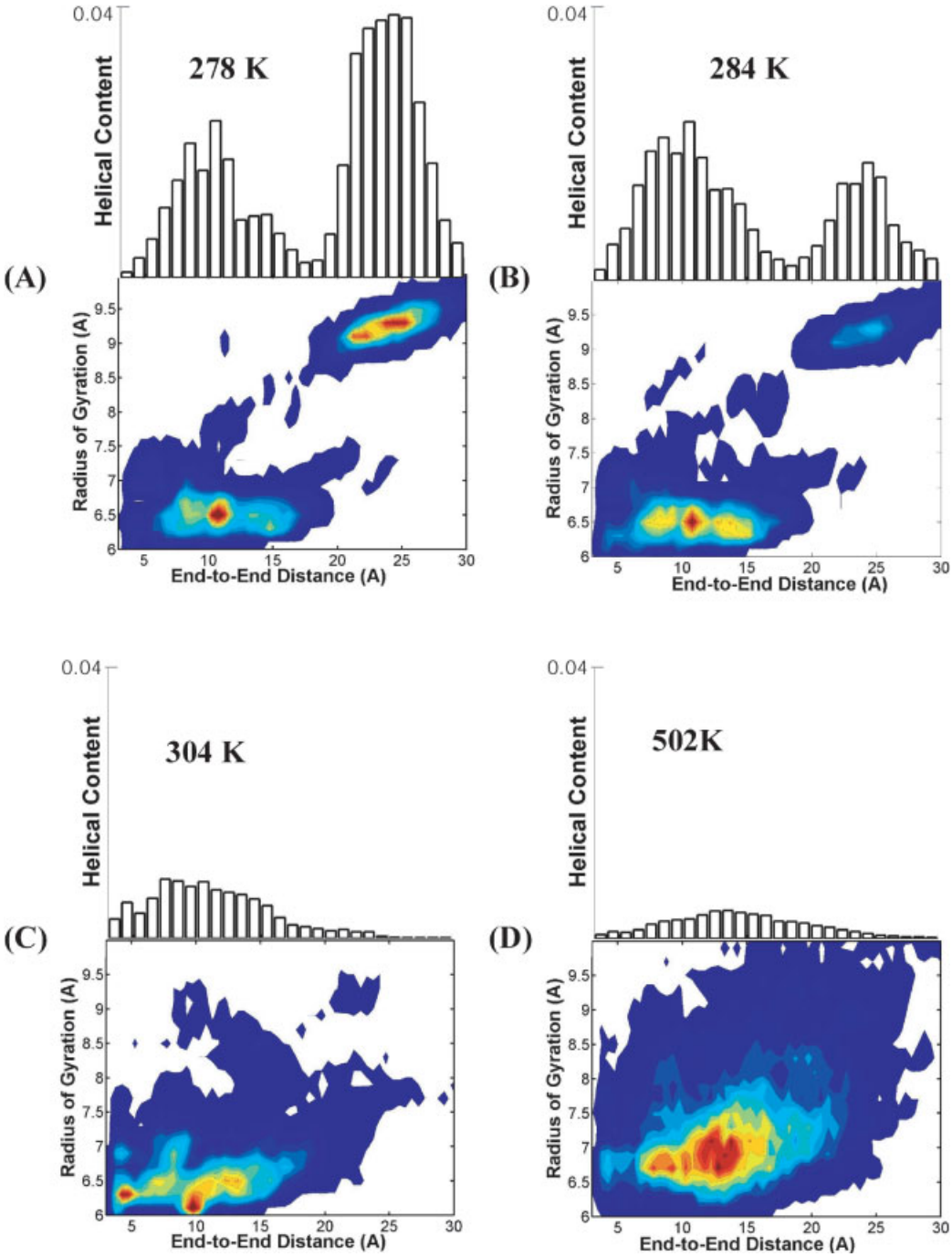

Fig. 5. The radius of gyration of A21 peptide is plotted as a function of the end-to-end distances at four different temperatures; 278 K (**a**), 284 K (**b**), 304 K (**c**), and 502 K (**d**). The color in the plot gradually changes from blue (least probable) to red (most probable). The bar graph at the top of the contour plot shows the contribution to the fractional helical content at the corresponding temperature from the individual end-to-end distances. All bar graphs are plotted on the same scale.

conformation (66%) with slightly more β (31%) than PP_{II} (25%) population is found. This result disagrees with the two-dimensional infrared (2D-IR) measurements,[71] but is in better agreement with the results from a different set of measurements that found a 50:50 mixture of β and PP_{II}.[66] The sampling of α′ basin is seen quite often with the OPLS compared to other force fields.[14,17,18] The PP_{II} content (1%) is also greatly diminished in the longer length $(Ala)_7$ peptide even though individual residues tend to sample noticeable amounts of PP_{II} (15%). Predominantly, the β basin is populated. With A94/MOD, predominantly PP_{II} population was found for $(Ala)_2$, $(Ala)_3$, and $(Ala)_7$ and, helical conformations were preferred over PP_{II} and β for larger

Ala peptides of lengths 14 and 17 (S. Gnanakaran and A.E. Garcia, unpublished results, 2004).[26] Our simulations suggest that the propensity of a force field to adopt nondominant conformations (e.g., PP_{II}) in longer sequences will depend on the ability of the force field to adopt these conformations in shorter peptides.

The $(Ala)_3$ model system also provides the opportunity to compare our results with other studies that have used OPLS/AA/L with explicit and implicit solvent models. Our results are consistent with the earlier studies, which found that, unlike OPLS/AA, the OPLS/AA/L version differentiates between the PP_{II} and β basins.[17,18]An MD simulation study of a cationic trialanine in TIP5P water found that OPLS/AA/L produced predominantly PP_{II} rather than β in agreement with 2D-IR measurements.[17] Even though the 2D-IR measurements were carried out on a cationic $(Ala)_3$, another experiment found that the conformational distribution of $(Ala)_3$ was independent of the terminal protonation states.[68] Results provided in a recent force field comparison study on $(Ala)_3$ with implicit solvent models shows that OPLS/AA/L yields higher PP_{II} population for cationic form (62%) compared with a neutral (capped) form (31%).[14] The relative population of PP_{II} and β of neutral trialanine obtained in that study is in agreement with our results. The same implicit solvent model, however, incorrectly predicts high α-helical conformation (45%) for $(Ala)_3$ with the A94/MOD force field.[14] Therefore, it appears that the OPLS/AA/L is capable of transferring the correct conformational preferential order found in explicit solvent to implicit solvent, whereas the A94/MOD is not.

Finally, we investigate whether the nature of unfolded configurations sampled at higher T by OPLS/AA/L and A94/MOD force fields is also different. One difference in the unfolded structures obtained with A94/MOD and OPLS/AA/L that is manifested in the end-to-end distance distribution (d) is the T dependence of the peak that arises from conformations with very short d (4 Å). Even though both force fields produce this peak, with OPLS/AA/L, the intensity of this peak increases significantly with T at relatively low T [Fig. 5(c)]. With A94/MOD, the same peak does not show this T dependence. At very high T, a wide range of conformations with a broad distribution of R_g and d is sampled [Fig. 5(d)]. The d distribution can be approximated by a Gaussian centered at 15 Å, and is similar to that of A94/MOD.[65]

The observed differences in the sampled secondary structures between OPLS/AA/L and A94/MOD can be attributed to parameters corresponding to backbone dihedrals.[15,25,65] It is also possible that parameters other than those of the backbone dihedrals can be responsible for some of the differences observed in this study. For example, the peak at the short d [Fig. 5(c)] occurs where the CH_3–CH_3 radial distribution function peaks in alkane liquids, which have been reproduced well by the OPLS/AA force field.[72] Comparison of methyl group parameters between OPLS/AA/L and A94/MOD indicate similar charge distributions but different LJ parameters, σ and ε. The ε_c and ε_H of A94/MOD are approximately two and a half times that of OPLS/AA/L, respectively (A94/MOD: ε_c = 0.1094, ε_H = 0.0157 and OPLS/AA/L: ε_c = 0.066, ε_H = 0.03). Hence, the discrepancies observed at short end-to-end distances with OPLS/AA/L and A94/MOD may be attributed to the ε parameter. Therefore, the difficult task of calibrating force fields over a temperature range will require experiments such as fluorescence resonance energy transfer measurements with a suitably chosen donor–acceptor pair that can capture the T dependence of conformations with short d. Importantly, the avenues for refinement may not necessarily involve only backbone dihedrals but also nonbonded parameters.

In summary, our structural and thermodynamic analysis indicate that even though both A94/MOD and OPLS/AA/L force fields yield similar helical content for the A21 peptide, the nature of conformations that contribute to the helix-coil transition profile is quite different.

CONCLUSION

We have used replica exchange molecular dynamics simulations to study the effect of the OPLS/AA/L force field on the helix formation for the A21 peptide. The helix-coil transition is in reasonable agreement with experimental data for similar short peptides. As with the A94/MOD force field, it yields reasonable Zimm-Bragg nucleation parameters that are close to the measured values. Enthalpic and entropic values obtained by fitting the temperature dependence of the ZB helix propagation parameter to a thermodynamic model are also comparable to measured values and to those obtained with A94/MOD force field under similar simulation conditions.

Even though both A94/MOD and OPLS/AA/L produce reasonable results for the helicity of A21 peptide, they sample quite different nonhelical conformations. The OPLS/AA/L force field produces no significant PP_{II} content in the A21 peptide. In shorter peptides, it yields β (31%) and PP_{II} (25%) populations for $(Ala)_3$ and predominantly β conformation for $(Ala)_7$. The A94/MOD force field, however, reproduced the observed shift in conformational preference with increasing peptide length (PP_{II} to α-helix) and temperature (increase in β).[26] The diminished PP_{II} content found in the simulations of the neutral short peptides of lengths 3 and 7 with OPLS/AA/L compared to A94/MOD may explain the differences in the non-α-helical conformational distribution in A21 peptide.

The OPLS/AA/L produces less PP_{II} content for short neutral Ala peptides, whereas A94/MOD may be producing higher than expected PP_{II} content for the same peptides. In conjunction with other force field comparison studies we find that it is possible that the PP_{II} content manifested through these force fields depends on the protonation of terminal groups and, therefore, warrants further examination, both experimental and theoretical, on the effect of capping groups for very short peptide systems. Overall, the enhanced sampling and the detailed comparison of simulation results with experimental data provide a way for validating force fields. In our opinion, this study also demonstrates that the gas phase ab initio methods based refinement of alanine dihedral potential is

justified, as exemplified by the reasonable conformational variability found for A21 peptide in condensed phase.

ACKNOWLEDGMENTS

We thank Professor Jorgensen for providing us with the OPLS/AA force field parameter and topology files.

REFERENCES

1. Dobson C. Protein folding and misfolding. Nature 2003;426:884–890.
2. Onuchic J, Nymeyer H, Garcia A, Chahine J, Socci N. The energy landscape theory of protein folding: insights into folding mechanisms and scenarios. Adv Protein Chem 2000;53:87–152.
3. Radford S. Protein folding: progress made and promises ahead. Trends Biochem Sci 2000;25:611–618.
4. Zagrovic B, Sorin E, Pande V. Beta-hairpin folding simulations in atomistic detail using an implicit solvent model. J Mol Biol 2001;313:151–169.
5. Snow C, Nguyen N, Pande V, Gruebele M. Absolute comparison of simulated and experimental protein-folding dynamics. Nature 2002;420:102–106.
6. Fersht A, Daggett V. Protein folding and unfolding at atomic resolution. Cell 2002;108:573–582.
7. Kubelka J, Hofrichter J, Eaton W. The protein folding "speed limit." Curr Opin Struct Biol 2004;14:76–88.
8. Kollman PA, Dixon RW, Cornell WD, Fox T, Chipot C, Pohorille A. In: Gunsteren Wv, editor. Computer simulations of biological systems. Dondrecht, The Netherlands: ESCOM; 1997.
9. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. A 2nd generation force-field for the simulation of proteins; nucleic-acids; and organic-molecules. J Am Chem Soc 1995;117:5179–5197.
10. MacKerell A, Bashford D, Bellott M, Dunbrack R, Evanseck J, Field M, Fischer S, Gao J, Guo H, Ha S, JosephMcCarthy D, Kuchnir L, Kuczera K, Lau F, Mattos C, Michnick S, Ngo T, Nguyen D, Prodhom B, Reiher W, Roux B, Schlenkrich M, Smith J, Stote R, Straub J, Watanabe M, WiorkiewiczKuczera J, Yin D, Karplus M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 1998;102:3586–3616.
11. Gunsteren WFV, Billeter SR, Eising AA, Hunenberger PH, Kruger P, Mark AE, Scott WRP, Tironi IG. GROMOS96 manual and user guide. Zurich, Switzerland: Vdf Hochschulverlag; 1996.
12. Jorgensen W, Maxwell D, TiradoRives J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. J Am Chem Soc 1996;118:11225–11236.
13. Halgren T. MMFF. VII. Characterization of MMFF94, MMFF94s, and other widely available force fields for conformational energies and for intermolecular-interaction energies and geometries. J Comput Chem (USA) 1999;20:730–748.
14. Zaman MH, Shen MY, Berry RS, Freed KF, Sosnick TR. Investigations into sequence and conformational dependence of backbone entropy, inter-basin dynamics and the flory isolated-pair hypothesis for peptides. J Mol Biol 2003;331:693–711.
15. Yoda T, Sugita Y, Okamoto Y. Comparisons of force fields for proteins by generalized-ensemble simulations. Chem Phys Lett 2004;386:460–467.
16. Zhou R, Berne B. Can a continuum solvent model reproduce the free energy landscape of a beta-hairpin folding in water? Proc Natl Acad Sci USA 2002;99:12777–12782.
17. Mu YG, Kosov DS, Stock G. Conformational dynamics of trialanine in water. 2. Comparison of AMBER, CHARMM, GROMOS, and OPLS force fields to NMR and infrared experiments. J Phys Chem B 2003;107:5064–5073.
18. Hu H, Elstner M, Hermans J. Comparison of a QM/MM force field and molecular mechanics force fields in simulations of alanine and glycine "dipeptides" (Ace-Ala-Nme and Ace-Gly-Nme) in water in relation to the problem of modeling the unfolded peptide backbone in solution. Proteins 2003;50:451–463.
19. Garcia A, Sanbonmatsu K. Exploring the energy landscape of a beta hairpin in explicit solvent. Proteins 2001;42:345–354.
20. Wu H, Wang S, Brooks B. Direct observation of the folding and unfolding of a beta-hairpin in explicit water through computer simulation. J Am Chem Soc 2002;124:5282–5283.
21. Pande V, Rokhsar D. Molecular dynamics simulations of unfolding and refolding of a beta-hairpin fragment of protein G. Proc Natl Acad Sci USA 1999;96:9062–9067.
22. Dinner A, Lazaridis T, Karplus M. Understanding beta-hairpin formation. Proc Natl Acad Sci USA 1999;96:9068–9073.
23. Zhou RH, Berne BJ, Germain R. The free energy landscape for beta hairpin folding in explicit water. Proc Natl Acad Sci USA 2001;98:14931–14936.
24. Nymeyer H, Garcia A. Simulation of the folding equilibrium of alpha-helical peptides: A comparison of the generalized born approximation with explicit solvent. Proc Natl Acad Sci USA 2003;100:13934–13939.
25. Garcia A, Sanbonmatsu K. Alpha-helical stabilization by side chain shielding of backbone hydrogen bonds. Proc Natl Acad Sci USA 2002;99:2782–2787.
26. Gnanakaran S, Garcia AE. Validation of an all-atom protein force field: from dipeptides to larger peptides. J Phys Chem B 2003;107:12555–12557.
27. Feig M, MacKerell AD Jr., Brooks CL III. Force field influence on the observation of pi -helical protein structures in molecular dynamics simulations. J Phys Chem B 2003;107:2831–2836.
28. Beachy MD, Chasman D, Murphy RB, Halgren TA, Friesner RA. Accurate ab initio quantum chemical determination of the relative energetics of peptide conformations and assessment of empirical force fields. J Am Chem Soc 1997;119:5908–5920.
29. Kaminski G, Friesner R, Tirado-Rives J, Jorgensen W. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. J Phys Chem B 2001;105:6474–6487.
30. Ohkubo Y, Brooks C. Exploring Flory's isolated-pair hypothesis: statistical mechanics of helix-coil transitions in polyalanine and the C-peptide from RNase A. Proc Natl Acad Sci USA 2003;100:13916–13921.
31. Jorgensen W, Tirado-Rives J. The OPLS potential functions for proteins. Energy minimizations for crystals of cyclic peptides and crambin. J Am Chem Soc 1988;110:1657–1666.
32. Zhou R. Free energy landscape of protein folding in water: explicit vs. implicit solvent. Proteins 2003;53:148–161.
33. Still WC, Tempczyk A, Hawley RC, Hendrickson T. Semianalytical treatment of solvation for molecular mechanics and dynamics. J Am Chem Soc 1990;112:6127–6129.
34. Jacobson M, Kaminski G, Friesner R, Rapp C. Force field validation using protein side chain prediction. J Phys Chem B 2002;106:11673–11680.
35. Ulmschneider J, Jorgensen W. Polypeptide folding using Monte Carlo sampling, concerted rotation, and continuum solvation. J Am Chem Soc 2004;126:1849–1857.
36. Felts A, Gallicchio E, Wallqvist A, Levy R. Distinguishing native conformations of proteins from decoys with an effective free energy estimator based on the OPLS all-atom force field and the surface generalized born solvent model. Proteins 2002;48:404–422.
37. Spek E, Olson C, Shi Z, Kallenbach N. Alanine is an intrinsic alpha-helix stabilizing amino acid. J Am Chem Soc 1999;121:5571–5572.
38. Garcia AE, Onuchic JN. Folding a protein in a computer: an atomic description of the folding/unfolding of protein A. Proc Natl Acad Sci USA 2003;100:13898–13903.
39. Gnanakaran S, Nymeyer H, Portman J, Sanbonmatsu KY, Garcia AE. Peptide folding simulations. Curr Opin Struct Biol 2003;13:168–174.
40. Gnanakaran S, Garcia AE. Folding of a highly conserved diverging turn motif from the SH3 domain. Biophys J 2003;84:1548–1562.
41. Gnanakaran S, Hochstrasser RM, Garcia AE. Nature of structural inhomogeneities on folding a helix and their influence on spectral measurements. Proc Natl Acad Sci USA 2004;101:9229–9234.
42. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys 1983;79:926–935.
43. Pearlman D, Case DA, Caldwell JW, Ross WS, Cheatam ITE, Ferguson DM, Singh UC, Weiner P, Kollman P. AMBER 41. San Francisco, CA: University of California; 1995.
44. Garcia AE, Sanbonmatsu KY. Exploring the energy landscape of a beta hairpin in explicit solvent. Proteins 2001;42:345–354.
45. Sugita Y, Okamoto Y. Replica-exchange molecular dynamics method for protein folding. Chem Phys Lett 1999;314:141–151.

46. Berendsen HJC, Postma JPM, Gunsteren WFv, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. J Chem Phys 1984;81:3684–3690.
47. Hummer G, Soumpasis D, Neumann M. Computer simulation of aqueous Na-Cl electrolytes. J Phys Condens Matter 1994;6:A141–A144.
48. Nymeyer H, Gnanakaran S, Garcia AE. Atomic simulations of protein folding, using the replica exchange algorithm. Methods Enzymol 2004;383:119–149.
49. Zimm BH, Bragg JK. Theory of the phase transition between helix and random coil in polypeptide chains. J Chem Phys 1959;31:526–535.
50. Lifson S, Roig A. On the theory of helix-coil transition in polypeptides. J Chem Phys 1961;34:1963–1974.
51. Qian H, Schellman J. Helix-coil theories: a comparative study for finite length polypeptides. J Phys Chem 1992;96:3987–3994.
52. Scholtz JM, Baldwin RL. The mechanism of alpha-helix formation by peptides. Annu Rev Biophys Biomol Struct 1992;21:95–118.
53. Yang J, Zhao K, Gong Y, Vologodskii A, Kallenbach N. Alpha-helix nucleation constant in copolypeptides of alanine and ornithine or lysine. J Am Chem Soc 1998;120:10646–10652.
54. Platzer KEB, Ananthanarayanan VS, Andreatta RH, Scheraga HA. Helix-coil stability-constants for naturally occurring amino-acids in water. 4. Alanine parameters from random poly(hydroxypropylglutamine-co-l-alanine). Macromolecules 1972;5:177–187.
55. Sung S, Wu X. Molecular dynamics simulations of synthetic peptide folding. Proteins 1996;25:202–214.
56. Mitsutake A, Okamoto Y. Helix-coil transitions of amino-acid homo-oligomers in aqueous solution studied by multicanonical simulations. J Chem Phys 2000;112:10638–10647.
57. Smith A, Hall C. Alpha-helix formation: discontinuous molecular dynamics on an intermediate-resolution protein model. Proteins 2001;44:344–360.
58. Sorin E, Pande V. Exploring the helic-coil transition via All-atom equilibrium ensemble simulations. Biophys J 2004, submitted. Available online in Biophys. J. BioFAST: January 21, 2005. doi:10.1529/biophysj.104.051938.
59. Lindahl E, Hess B, van der Spoel D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. J Mol Biol 2001;7:306–317.
60. Scholtz JM, Marqusee S, Baldwin RI, York EJ, Stewart JM, Santoro M, Bolen DW. Calorimetric determination of the enthalpy change for the alpha-helix to coil transition of an alanine peptide in water. Proc Natl Acad Sci USA 1991;88:2854–2858.
61. Taylor JW, Greenfield NJ, Wu B, Privalov PL. A calorimetric study of the folding-unfolding of an alpha-helix with covalently closed N and C-terminal loops. J Mol Biol 1999;291:965–976.
62. Lopez MM, Chin DH, Baldwin RL, Makhatadze GI. The enthalpy of the alanine peptide helix measured by isothermal titration calorimetry using metal-binding to induce helix formation. Proc Natl Acad Sci USA 2002;99:1298–1302.
63. Richardson JM, Makhatadze GI. Temperature dependence of the thermodynamics of helix-coil transition. J Mol Biol 2004;335:1029–1037.
64. Pappu RV, Rose GD. A simple model for polyproline II structure in unfolded states of alanine-based peptides. Protein Sci 2002;11:2437–2455.
65. Garcia AE. Characterization of non-alpha helical conformations in Ala peptides. Polymer 2004;45:669–676.
66. Eker F, Cao X, Nafie L, Schweitzer-Stenner R. Tripeptides adopt stable structures in water. a combined polarized visible Raman, FTIR, and VCD spectroscopy study. J Am Chem Soc 2002;124:14330–14341.
67. Shi Z, Olson C, Rose G, Baldwin R, Kallenbach N. Polyproline II structure in a sequence of seven alanine residues. Proc Natl Acad Sci USA 2002;99:9190–9195.
68. Schweitzer-Stenner R, Eker F, Griebenow K, Cao XL, Nafie LA. The conformation of tetraalanine in water determined by polarized raman, FT-IR, and VCD spectroscopy. J Am Chem Soc 2004;126:2768–2776.
69. McColl IH, Blanch EW, Hecht L, Kallenbach NR, Barron LD. Vibrational Raman optical activity charaterization of poly(L-proline) II helix in alanine oligopeptides. J Am Chem Soc 2004;126:5076–5077.
70. Woutersen S, Mu Y, Stock G, Hamm P. Subpicosecond conformational dynamics of small peptides probed by two-dimensional vibrational spectroscopy. Proc Natl Acad Sci USA 2001;98:11254–11258.
71. Woutersen S, Pfister R, Hamm P, Mu Y, Kosov D, Stock G. Peptide conformational heterogeneity revealed from nonlinear vibrational spectroscopy and molecular-dynamics simulations. J Chem Phys 2002;117:6833–6840.
72. Kaminski G, Duffy EM, Matsui T, Jorgensen WL. Free-energies of hydration and pure liquid properties of hydrocarbons from the OPLS all-atom model. J Phys Chem 1994;98:13077–13082.