doi:10.1016/j.jmb.2005.11.058 J. Mol. Biol. (2006) 356, 248–256

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

The Solvation Interface is a Determining Factor in Peptide Conformational Preferences

Eric J. Sorin[1], Young Min Rhee[4], Michael R. Shirts[5] and Vijay S. Pande[1,2,3]*

[1]*Department of Chemistry Stanford University, Stanford CA 94305-5080, USA*

[2]*Department of Structural Biology, Stanford University Stanford, CA 94305-5080 USA*

[3]*Stanford Synchrotron Radiation Laboratory, Stanford University, Stanford, CA 94305-5080, USA*

[4]*Department of Chemistry University of California at Berkeley, Berkeley, CA 94720 USA*

[5]*Department of Chemistry Columbia University, New York NY 10027, USA*

**Corresponding author*

The 21 residue polyalanine-based F_s peptide was studied using thousands of long, explicit solvent, atomistic molecular dynamics simulations that reached equilibrium at the ensemble level. Peptide conformational preference as a function of hydrophobicity was examined using a spectrum of explicit solvent models, and the peptide length-dependence of the hydrophilic and hydrophobic components of solvent-accessible surface area for several ideal conformational types was considered. Our results demonstrate how the character of the solvation interface induces several conformational preferences, including a decrease in mean helical content with increased hydrophilicity, which occurs predominantly through reduced nucleation tendency and, to a lesser extent, destabilization of helical propagation. Interestingly, an opposing effect occurs through increased propensity for 3_{10}-helix conformations, as well as increased polyproline structure. Our observations provide a framework for understanding previous reports of conformational preferences in polyalanine-based peptides including (i) terminal 3_{10}-helix prominence, (ii) low π-helix propensity, (iii) increased polyproline conformations in short and unfolded peptides, and (iv) membrane helix stability in the presence and absence of water. These observations provide physical insight into the role of water in peptide conformational equilibria at the atomic level, and expand our view of the complexity of even the most "simple" of biopolymers. Whereas previous studies have focused predominantly on hydrophobic effects with respect to tertiary structure, this work highlights the need for consideration of such effects at the secondary structural level.

Keywords: protein folding; explicit solvent; distributed computing; hydrophobicity; molecular dynamics

Introduction

The protein folding problem has garnered much attention over the past few decades. At the most fundamental level, folding is defined by the underlying distribution of thermodynamic states, both stable and otherwise, that determine the interesting and experimentally detectable features of the free energy landscape, as well as the measurable kinetic modes of conformational change. While it is accepted that many factors contribute to experimentally observed folding behaviors, a predominant theme in contemporary theories is that of hydrophobicity and hydrophilicity,[1] and several studies have been published recently that examine the role of water in self-assembly.[2–13]

It is intuitive to assume that the conformational preferences of a given protein or peptide are influenced significantly by local environmental conditions, and the most important environmental condition is the character of the surrounding solvent. More specifically, we refer herein to the solvation interface, which communicates bulk properties of the solvent (i.e. temperature, pressure, dielectric/salt effects) to the peptide and determines localized effects about the peptide due to specific solute–solvent interactions (i.e. hydrophobic, water-mediated, and ion-binding interactions). Understanding how the solvation

Abbreviations used: RMS, root-mean-squared; LR, Lifson–Roig.

E-mail address of the corresponding author: pande@stanford.edu

interface affects peptide conformational preferences is thus a fundamental concept in current and future efforts to understand and model the protein folding problem on the residue and atomic scales.

Although gaining experimental insight into the effects of solvation on protein folding remains a challenge,[14] a new era of computational power and complexity now allows the simulation of protein dynamics in all-atom detail using explicit representations of aqueous solvent. In many ways, such efforts began with the seminal work by Duan and Kollman, whose 1 μs explicit solvent simulation of the villin headpiece[15] set the stage for simulating protein conformational change in all-atom detail. More recently, Zhou *et al.* reported on hydrophobic collapse between two domains of the BphC enzyme, demonstrating a dependence of the observed hydrophobic collapse kinetics upon solute–solvent electrostatic interactions.[5] Rhee *et al.* studied complete folding of the BBA5 mini-protein in explicit solvent using ensemble dynamics methodology,[16] describing water-induced effects not captured by implicit solvation models, including a concurrent core collapse and desolvation mechanism,[9] and we recently reported a similar need for explicit solvation in simulations of the folding of a small RNA.[8]

While these and other recent works have shown the utility of simulation in furthering our understanding of the effects of solvation on configurational preferences, the limiting factor in contemporary simulation studies is most often that of sampling true equilibrium. We have recently reported true ensemble-level equilibrium sampling of the capped 21 residue α-helical F_s peptide in all-atom detail,[17,18] using Folding@Home, a worldwide distributed computing network.[19] The sampling reported in our previous studies,[17,18] orders of magnitude greater than the experimental folding time, allowed us to assess several contemporary protein potentials including AMBER-94, AMBER-96, AMBER-99, and AMBER-GS (modified AMBER-94 of Garcia and Sanbonmatsu), as well as a new variant dubbed AMBER-99φ, which reduces barriers along the φ torsion in AMBER-99, thereby making helical conformations more accessible to the peptide. While both the original and modified versions of AMBER-94 exhibited significant helix-friendly equilibria, and the AMBER-96 and AMBER-99 force fields favored extended non-helical conformations, the AMBER-99φ potential was shown to yield the strongest agreement with experimental measurements on this system, including folding rate, Lifson–Roig (LR) parameters, and radius of gyration.[17] The notable agreement with several experiments, the ability to simulate this system to true equilibrium, and the fact that α-helices compose the most elementary subunits of protein structure make this system an ideal candidate with which to examine effects of the solvation interface on elementary protein configurational preferences.

While classical helix–coil theories treated solely the statistics of the peptide,[20,21] it has long been appreciated that solvent effects play an important role in peptide and protein structure:[1] it is well established that polyalanine is helix-stabilizing,[22,23] that charged blocking groups stabilize even "short" polyalanine helices in aqueous solvent,[24] and that the hydrophobic effect plays a significant role in helix stability.[25,26] Herein, we consider the molecular basis of the effect of hydrophobicity on peptide conformation. Indeed, a number of conformational types are known within the "helical spectrum," which includes α, 3_{10}, π, and polyproline regions of the Ramachandran map, and preferences toward each of these specific forms of helix have been identified, as discussed below. This study serves as the first to probe the connection between those preferences and the character of the solvent. We have identified a clear correlation between such structural preferences and the strength of the hydrophobic effect, and the observed conformational response to hydrophobicity rationalizes the complex conformational preferences observed in simple polyalanine-based peptides.

Shirts and Pande recently showed that minor modifications to the TIP3P water model[27] decrease the mean error in hydration free energies of common amino acid sidechain analogs to 0.0 ± 0.1 kcal/mol while reducing the root-mean-squared (RMS) error below that of any of the published water models tested, with measured bulk liquid water properties remaining almost constant.[28] To consider the effect of the solvation interface on protein conformational preference in a simple and systematic way, we therefore examine the equilibrium thermodynamics of the F_s peptide in a spectrum of water models including TIP3P, TIP3P-MOD,[29] and the six TIP3P-M2*X* models described previously[28] and outlined in Table 1. A recent study by Rose and co-workers showed that solvation free energy in TIP3P explicit solvent simulations correlated strongly with the interaction energy between the peptide and coordinated (i.e. first shell) water molecules.[30] With this in mind, we conducted a computational thought experiment in which the solute–solvent interaction energy is varied by employing the eight TIP3P variants listed above, thereby studying the effects of "hydrophobic titration" on peptide conformational equilibrium.

In the solvent models listed above, the energetic benefit of forming close van der Waals contact between solute and solvent is increased by increasing the Lennard-Jones potential well depth, ε, by ~65%. To compensate for the increase in atomic repulsion at a given distance, the Lennard-Jones radius, σ, is decreased to maintain the experimental density of water, with changes of ≤1.5% in the optimal oxygen–oxygen separation. These changes thus allow the bulk character of water to remain essentially unchanged while varying the solute–solvent attraction greatly at the solvation interface. In this way, we perturb the solute–solvent interface to study resulting changes in the conformational preferences of

Table 1. TIP3P and MXX water models and simulation sampling

Water model	ε (kcal/mol)	σ (Å)	ρ (g cm^{-3})	Solvation error[a] (kcal/mol)	Max (ns)	Total time (μs)	>EQ (μs)
TIP3P	0.1521	3.15061	0.9859	0.50	95	78.197	13.352
TIP3P-MOD	0.1900	3.12171	0.9998	0.23	110	99.005	26.954
TIP3P-M20	0.2000	3.12000	0.9976	0.18	145	128.870	51.837
TIP3P-M21	0.2100	3.11800	0.9963	0.15	125	110.050	40.633
TIP3P-M22	0.2200	3.11500	0.9976	0.06	120	99.990	39.578
TIP3P-M23	0.2300	3.11300	0.9970	0.04	105	93.185	40.121
TIP3P-M24	0.2400	3.11100	0.9976	0.00	110	106.355	31.381
TIP3P-M25	0.2500	3.11000	0.9969	0.01	105	92.755	40.504
Total						808.407	284.360

[a] Error in the solvation free energy of the amino acid side-chain analogs, weighted by the amino acid frequency.[28]

polyalanine-based peptides with corresponding changes in hydrophilicity/hydrophobicity (i.e. "solute–solvent attraction/repulsion" or "binding strength at the solvation interface").

Figure 1 demonstrates the equilibrium sampled for each ensemble in our study, and we choose 40 ns as a conservative starting time for calculation of equilibrium properties. In comparison to the coil-to-helix folding time of ~16 ns, determined by experiment[31] and simulation,[17] a total sampling time of over 800 μs and a total equilibrium sampling of over 250 μs are included in the following analyses. We report below several key features of the hydrophobic effect with regard to the simple helix–coil system in equilibrium, and consider how these features arise from simple consideration of the hydrophilic and hydrophobic components of the solvent-accessible surface area (SASA) as a function of peptide length for several ideal conformational types.

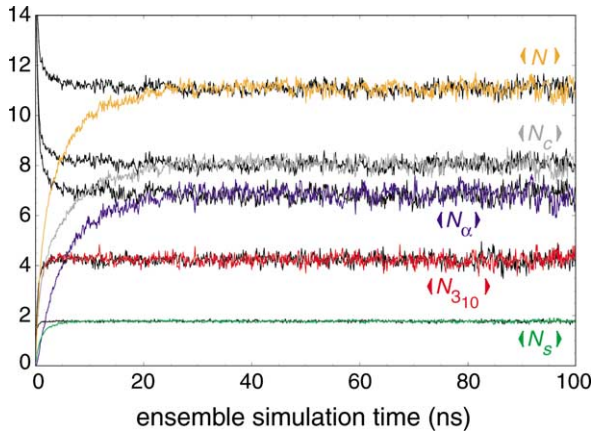

Figure 1. Example of helix–coil ensemble simulation convergence using DSSP for the ensemble averaged total helicity ⟨*N*⟩, contiguous helicity ⟨N_c⟩, number of helical segments ⟨N_s⟩, and mean α and 3_{10} helicities. Curves representing ensembles starting from the native state (unfolding) are shown in black, while colors represent folding ensembles. Agreement at long timescales of ensembles started folded and unfolded reflects the convergence of the kinetics as well as the high degree of sampling possible through distributed computing.

Results

Hydrophobicity influences peptide conformational preferences

The primary result we report herein is a shift in helical tendency and water coordination to the peptide backbone with changes in the attraction energy across the solvation interface (Figure 2). That is, increased hydrophilicity (decreased hydrophobicity) leads to an increase in the mean coordination of water to the peptide backbone and the helix–coil equilibrium shifts in favor of less helical content. As Figure 2 demonstrates, this effect is nearly position-independent along the sequence, with terminal residues showing slightly less sensitivity to hydrophobicity than central residues. Interestingly, the decline in α-helical tendency is accompanied by an increased preference for 3_{10} helical structure (Figure 2). This increase in 3_{10} structure is not as

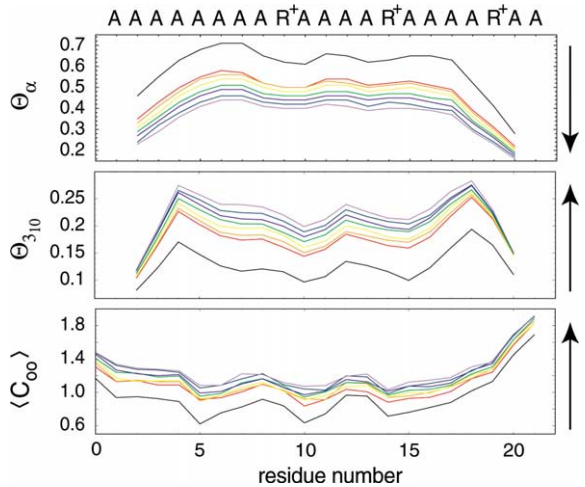

Figure 2. Fractional α and 3_{10} helicity and water coordination (C_{oo}) to the peptide backbone per residue. Standard TIP3P water is shown in black, with TIP3P-MOD in red and models M20 thru M25 spanning the color spectrum from orange to violet, respectively. Residue 0 represents the acetyl cap, and terminating alanine residues (1 and 21) are not considered by DSSP. Arrows represent the trend from most hydrophobic (TIP3P) to the most hydrophilic water model (TIP3P-M25).

dramatic as the decrease in α structure, however, resulting in a net decrease in overall helicity.

We attribute these changes in conformational preference directly to the increase in water coordination along the peptide. That is, 3_{10} helical conformations allow greater solvent access to the peptide, and are thus favored when hydrophilicity is increased. At terminal positions, less steric shielding of solvent by side-chains and/or neighboring backbone regions is occurring, which allows increased interactions between the aqueous solvent and the peptide. Terminal residues thus exist in a local environment that is different from those farther along the sequence, and it should be no surprise that hydrophobic "end-effects" would be present. The end-effect in this case is a lower sensitivity to changes in solute–solvent binding strength. As shown in Figure 3, the coordination of water to the peptide backbone and the preferences for α and 3_{10} conformations are correlated strongly with hydrophilicity (attraction energy, ε) and with one another.

Because the DSSP algorithm assigns secondary structure based on hydrogen bonding within the peptide, we assessed the possibility that the resulting changes in observed helicity were artifacts

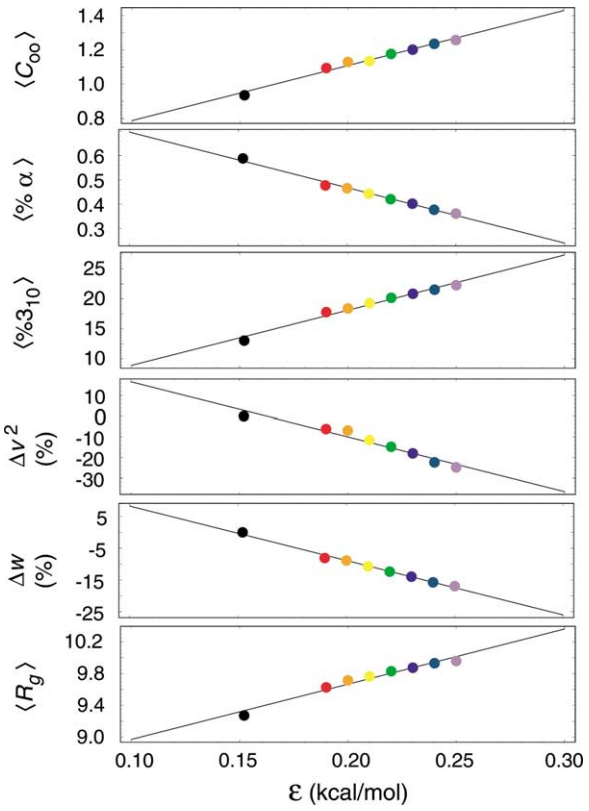

Figure 3. Structural properties as a function of hydrophobicity: from top to bottom are the mean backbone water coordination, the mean fractional α and 3_{10} helicities, the changes in LR helix–coil nucleation (v^2) and propagation (w) parameters from the TIP3P model, and the mean radius of gyration. The color coding matches that in Figure 2; hydrophobicity decreases from black (TIP3P) to violet (TIP3P-M25).

based solely on changes in the hydrogen bonding network in our simulations (due to added water coordination, and thus greater likelihood of peptide–water hydrogen bonds) rather than actual changes in the backbone conformational distribution on the Ramachandran map. This was done by examining the changes in LR helix–coil parameters with changes in solute–solvent attraction (ε), which also allowed us to determine to what degree the nucleation and propagation components of helix formation were affected by the change in hydrophobicity. The LR nucleation and propagation parameters (v^2 and w, respectively) were calculated using Qian-Schellman theory[32] as done by Garcia and Sanbonmatsu,[12] and described in our previous works.[17,18] This method of quantifying helicity inherently considers both α and 3_{10} conformations as contributing to the helical regime, yet considers only backbone ϕ/ψ torsion angles in assigning helical secondary structure to each residue in the sequence, thus offering a secondary method of assessing helix–coil structure completely independent of DSSP assignments.

As shown in Figure 3, it is evident that the trend established using DSSP assignments is not artifactual. Based solely on ϕ/ψ degrees of freedom, nucleation and propagation tendencies both decrease significantly with decreased hydrophobicity, and the change in nucleation dominates the decrease in helicity seen here (25% lower with a 65% increase in attraction energy), particularly as propagation depends upon successful nucleation events. Anti-correlations between the LR parameters and water coordination are also strong (R^2 values of 0.940 and 0.996). An examination of the equilibrium ϕ/ψ distributions in each water model shows this decrease in sampling of the helical basin and an increased sampling of turn conformations with decreasing hydrophobicity (Supplementary Data). The increased water coordination, and the resulting decrease in helicity also affect the bulk molecular size of the peptide, as shown by the strong correlation between hydrophilicity and the mean radius of gyration (R_g) in Figure ($R^2 = 0.983$).

How does the solvation interface help to predetermine secondary structure?

To better understand the nature of the effects reported above, we turn to a somewhat simplistic view of peptide conformational properties. We consider the hydrophobic and hydrophilic components of the SASA[33,34] within the "helical spectrum" of peptide conformations, which ranges from the fully extended peptide ($\phi = 180°$, $\psi = 180°$) to the helix of largest diameter ($\phi = -57°$, $\psi = -80°$). For capped polyalanine peptides Ace-ALA_X-NH_2 from one to 25 residues in length, ideal structures for each of these conformational types were generated and the SASA components were evaluated. Ideal structures for capped Ace-ALA_{10}-NH_2 peptides are shown in Figure 4(a) for visual clarity. Figure 4(b)

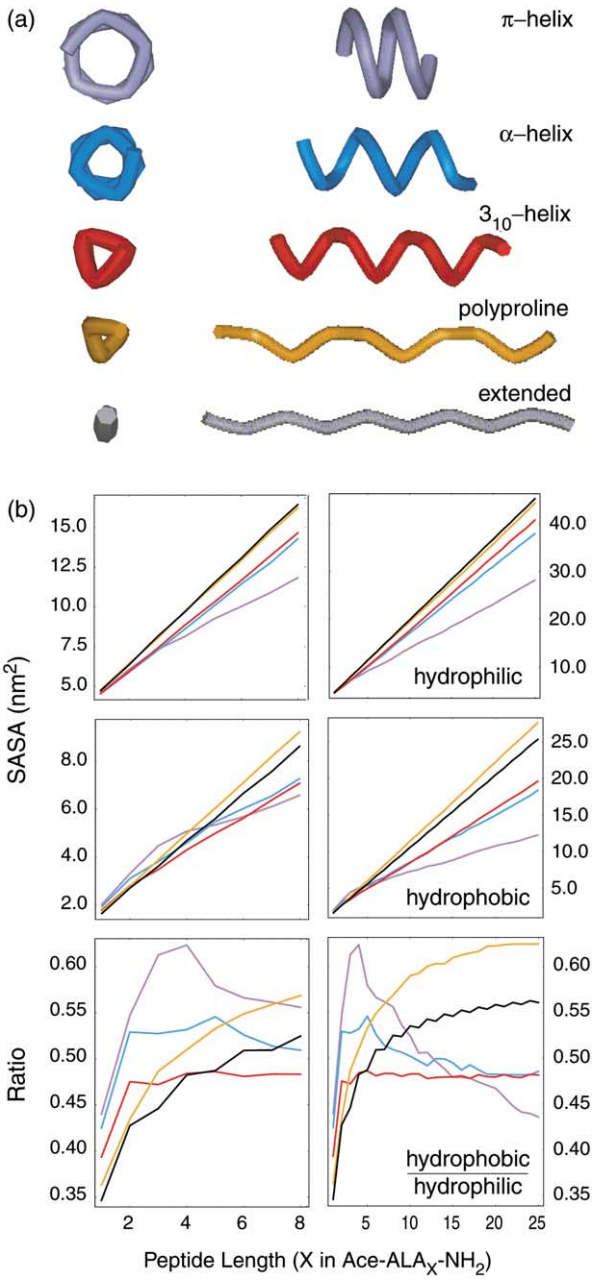

Figure 4. (a) Ideal conformations of the "helical spectrum" for capped polyalanine 10-mers. (b) Hydrophilic and hydrophobic solvent-accessible surface area as a function of Ace-ALA$_X$-NH$_2$ peptide length. The ratio of these components is shown in the third panel, and plot colors in (b) correspond to backbone structures represented in (a).

shows the resulting SASA profiles for each conformational type as a function of peptide length. Frames on the left of each panel show a magnification of these profiles for shorter peptide lengths.

While this analysis does not consider the complex energetic (i.e. hydrogen bond, torsional effects, etc.) and entropic factors of peptide configurational preference, we seek herein only to examine the hydrophobic effect at the solvation interface, and the plots in Figure 4(b) offer several key features. At all peptide lengths simulated, hydrophilic SASA scales inversely with helix diameter: extended conformations maximize this solvophilic component, which decreases with increasing helix diameter. A similar trend is observed for the hydrophobic component for peptide lengths over about ten residues, and three regimes appear to separate these conformational types: extended and polyproline structures maximize hydrophobic SASA and π-helix minimizes it, with the more ubiquitous α and 3_{10} helices lying midway between these regions.

To better compare these conformational types relative to one another, the final panel in Figure 4(b) plots the ratio of hydrophobic to hydrophilic SASA for each. Because these two components are inverse metrics of the same physical phenomenon (solvent–peptide attraction/repulsion), this ratio allows us to gauge which conformational types are best suited to an aqueous environment, with a high ratio suggesting a poor fit in water and a low ratio suggesting a water-friendly conformational type. One striking feature is the length-independent low SASA ratio for the 3_{10} helix, which is the most solvophilic of the helices in the most biologically and chemically relevant region (5–15 residues). In contrast, α-helical conformations are less solvophilic in local regions (i.e. helix nuclei with five or fewer residues participating) and become more suited to the aqueous environment with increased helical length. The same is true for the π-helix, which becomes the most solvophilic conformer only in stretches outside the biologically relevant regime (i.e. over 15 residues).

These profiles may explain, without consideration of entropy or inter-peptide interaction energetics, the lack of π-helices in polyalanine-based systems, and the relatively low π abundance in peptide/protein systems in general. Indeed, we observe negligible π content in our simulations, as was reported for polyalanine-based peptides,[35] and to our knowledge no observation of this conformational type in polyalanine-based systems has been reported. In comparison, 3_{10} helices would be predicted by these SASA profiles alone to be predominant in polyalanine-based peptides, and accounting for the energetic differences between 3_{10} and α conformations is clearly a necessity in assessing accurately the propensity of these types relative to one another. The fact that 3_{10} is a more water-friendly conformational type than α supports the observations reported above: in terminal regions, where solvent exposure is greatest, 3_{10} helices should be prevalent.

Complex conformational equilibrium in simple polyalanine-based peptides: a model for larger proteins?

A significant literature has developed around the study of polyalanine-based peptides in recent years, and this literature could be the subject of an extensive review. Several interesting facets of the conformational preferences of such peptides have

been published, and a comparison of recent publications made in our previous work suggested that even simple polyalanine peptides exhibit significant length-dependent conformational preferences.[18] Below, we discuss how the observations above relate to, and rationalize, the most significant of these conformational preferences in terms of hydrophobic arguments.

To begin with, our results rationalize the findings reported by Lewis *et al.*,[26] who observed that the lysine-blocked polyalanine 24-mer (Ace-K_2-A_{24}-K_2-NH_2) forms stable helical structure in phospholipids, but is disrupted when water is present. That is, in the absence of water the hydrophobic effect stabilizes helical conformations, as described above. However, when water is present in such an environment, the hydrophilic regions of the peptide surface can act to draw water away from the hydrophobic phospholipids. The peptide–environment interface is thereby dominated by the attraction of water to the peptide, mimicking the effect of heightened hydrophilicity, and helical propensity decreases. The hydrophobic interior of folded proteins could be another example of such an environment in which secondary structure is stabilized in the absence of water and grows less stable as tertiary contacts are broken and solvent access increases.

Unlike this protein unfolding scenario, however, solvated polyalanine-based helices do not enshroud a significant non-polar core or produce a dramatic change in long-range hydrophobic interactions when losing helical content. Additionally, the hydrogen bond network about helix-forming peptides includes an averaging of solute and solvent participation, and changes in the hydrogen bond network with respect to the very rapid and dynamic helix–coil conformational change take place only in localized regions near the peptide. It is therefore not surprising that a small apparent change in heat capacity has been detected for helix formation and denaturation.[22]

While the 3_{10}-helix has been suggested as potentially playing a role as nucleus or intermediate in α-helix formation,[36] few experimental probes differentiate between 3_{10} and α-helical forms, and few studies have been conducted specifically on 3_{10} helices. Of most interest with respect to the current study is the work by Millhauser *et al.*, whose nuclear Overhauser effect spectroscopy studies of the alanine-based peptides 3K (Ace-$(A_4K)_3$A-NH_2) and MW (Ace-AMAAKAWAAKAAAARA-NH_2) suggested that 3_{10}-helix populations were significant, particularly in the terminal regions.[36] While this finding remains somewhat controversial, previous simulation results have supported Millhauser's report.[17,37] As discussed above, added solvation in terminal regions favors 3_{10} conformations due to their smaller helical diameter, which can more easily accommodate surrounding solvent, and these regions are less sensitive to changes in hydrophobicity due to their inherent "end-effect" status.

Finally, we consider polyproline structure, which has been suggested as a dominant conformational form in the unfolded state of proteins and, in particular, polyalanine peptides.[30,38–40] Some previous reports have assessed polyproline character in polyalanine peptides.[30,38–43] Notably, Shi *et al.* offered convincing experimental evidence that the polyalanine A_7 peptide, blocked by two charged hydrophilic residues at each terminus, adopts a predominantly polyproline structure at biologically relevant temperatures with little α-helical structure.[38] We have recently studied this same blocked peptide[44] using small-angle X-ray scattering, which showed that the molecular size of this system ($R_g \approx 7.4$ Å) is much smaller than would be expected of the ideal polyproline conformation ($R_g \approx 13.1$ Å). The picture that emerged from this observation, in light of the study by Shi *et al.*, is one in which short, localized polyproline regions act independently and likely fluctuate on relatively fast timescales, much like what has been reported for the helix–coil transition in the F_s peptide.[17]

This picture is justified by the SASA profiles in Figure 4(b) and the discussion above. While the polyproline conformation leaves more hydrophobic area exposed than the fully extended peptide at all lengths, making it the least hydrophilic conformation for peptide lengths greater than about eight residues, this is not the case for shorter peptides, and the hydrophilicity of short polyproline stretches lies between 3_{10} and α conformations. Although the extended peptide appears to be more solvophilic than polyproline, it is also entropically disfavored, which likely leads to larger detectable polyproline propensity due to the neighboring positions of these conformations on the Ramachandran map.

The results described above lend insight into the length-dependent conformational preferences inherent to polyalanine-based peptides by considering the effect of the solvation interface. Indeed, the solvophilic nature of short polyproline stretches explains the observations made by Shi *et al.* in the context of our recent small-angle X-ray scattering work. As we have stated previously, this presents a new challenge in peptide simulations: potential sets that produce the best agreement with experimental results for larger helix–coil systems (such as the capped 21-mer F_s peptide), predict little polyproline content. While the correlation between polyproline content and binding strength across the solvation interface ($R^2 \sim 0.7$) is not as strong as those reported above, our current simulations predict increasing populations in the polyproline regime of the Ramachandran map with increasing hydrophilicity (~35% increase in polyproline between the most hydrophobic and hydrophilic solvent models), in agreement with the observation that shorter peptides are highly solvophilic in extended conformational states.

Simple polyalanine-based systems thus appear to exhibit complex conformational preferences

based on the character of the local solvation interface, with α-helices favored in hydrophobic environments, 3_{10}-helical regions favored in hydrophilic regions and near peptide termini, and polyproline conformations favored in short peptides or short regions of longer peptides. Still, experimental and/or computational studies on a wide variety of peptides and proteins have flushed out several other facets of folding behavior, and the overall importance of hydrophobicity appears to be highly system-specific. Indeed, the formation of secondary structure is not universally known to precede formation of tertiary structure, and larger systems often exhibit several phases of folding, some with final secondary structure forming near the end of the unfolded-to-folded transition. Understanding the complexity of polyalanine-based peptides may therefore offer insight into the highly variable secondary structural preferences of larger systems at various stages throughout folding

What are the relevant "states" in the folding of a given sequence/structure? In the context of larger systems, long-range motions can cause significant changes in local solvent–peptide interactions in many ways (side-chain packing, intra-polymer collisions, etc.). Shifts in conformational preferences over a portion of the sequence thus seem unavoidable, adding further complexity to the assessment of "states" along the protein folding reaction. Indeed, according to the observations above, the hierarchical view of protein folding, which assumes that secondary structure formation precedes tertiary folding, may be a reasonable approximation that simply misses the underlying detail at the residue level. Furthermore, the results above make it somewhat intuitive that as hydrophobic collapse occurs, polyproline structure will become less advantageous, supporting hypotheses that polyproline structure is more dominant in the unfolded state of proteins than observed in native structures.[30,38–40]

Concluding Discussion

The results of our equilibrium simulations offer a new look at the role of water in the protein folding process. Indeed, experimentally observed conformational preferences in polyalanine-based peptides can be explained solely on the basis of the effective interaction with the solvent at the water–peptide boundary. Helix–coil equilibrium, which represents the structurally most simple and kinetically most rapid portion of the protein folding problem, is greatly affected by the binding strength at the solvation interface ("hydrophobicity"). Water coordination is correlated strongly with hydrophilicity and thus affects peptide conformational preferences. The propensity for α-helical conformations decreases with decreased hydrophobicity, as suggested by Yang & Honig,[25] and total helicity decreases both in nucleation and propagation tendencies, with the decrease in nucleation being predominant. In contrast, the propensity of 3_{10}-helices increases with decreased hydrophobicity as this conformational type allows greater coordination of water to the peptide backbone than the α-helix, thus explaining the observations of significant 3_{10} populations in terminal regions where solvent exposure is highest.

Considerations of the hydrophilic and hydrophobic components of the SASA for the ideal conformational types shown in Figure 4(a), and the dependence of those profiles on peptide length, reinforce these observations and offer additional insight into the conformational preferences in polyalanine-based peptides. While π-helix conformations are relatively solvophobic for peptide lengths that are of general biological and chemical relevance, and the more common α and 3_{10} types are solvophilic in this regime, extended conformations (including polyproline structure) are most solvophilic in short peptide segments. This latter observation explains previous experimental reports in which a short polyalanine peptide was observed to favor polyproline conformations, yet also exhibited a molecular size nearly half that of the ideal polyproline structure. Indeed, the model presented by Zagrovic *et al.*, in which short, independent polyproline segments are present,[44] appears to be well supported by the picture offered herein.

As reported previously, our results suggest that the inclusion of an explicit representation of the solvent is necessary to best model biomolecular dynamics.[5–10] Indeed, this report demonstrates the sensitivity of observed thermodynamic equilibria to minor changes in the character of the solvation interface, which can currently be modeled only in the presence of explicit molecular water. We hope that the relationship observed between conformational preference and hydrophobicity/SASA may serve as a benchmark for further refinement of solvation models, both implicit and explicit alike, around peptide and protein systems.

Methods

The capped F_s peptide (Ace-A_5[AAAR$^+$A]$_3$A-NMe) was simulated under the AMBER-99φ all-atom potential in explicit ionic solvent as described,[17] using the GROMACS molecular dynamics suite,[33] as modified for the Folding@Home infrastructure†. A canonical helix ($\phi = -57°$, $\psi = -47°$) and a random coil configuration with no helical content were generated and centered in 40 Å cubic boxes. Electroneutrality was gained by placing three chloride ions randomly around the solute with minimum ion–ion and ion–solute separations of 5 Å. The helix and coil conformations were then solvated with 2075 and 2065 TIP3P water molecules,[27] respectively, energy minimized using a steepest descent algorithm, and annealed for 500 ps of molecular dynamics with the peptide conformation held fixed.

† http://folding.stanford.edu

Starting configurations were then generated for the additional water models tested in this study, including TIP3P-MOD[29] and the TIP3P-M2*X* models,[28] by modifying the water parameters (σ, ε) and rerunning the energy minimization and annealing steps described above. Each of the resulting conformations served as the starting point for 1000 independent molecular dynamics trajectories, yielding a total of 16,000 trajectories simulated on ~20,000 CPUs within the Folding@Home supercluster. Table 1 summarizes these TIP3P variants, including the frequency-weighted average error in the solvation free energy of amino acid side-chain analogs reported previously,[28] and the sampling achieved for each solvation model.

All simulations reported herein were conducted as described,[17] using standard AMBER scaling of 1–4 interactions[18] in the NPT ensemble[45] at 1 atm (101,325 Pa) and 305 K, the approximate F_s midpoint temperature detected by circular dichroism[46] and ultraviolet resonance Raman.[47] Long-range electrostatic interactions were treated using the reaction field method with a dielectric constant of 80, and 9 Å cutoffs were imposed on all Coulombic and Lennard-Jones interactions. Non-bonded pair lists were updated every ten steps, and covalent bonds involving hydrogen atoms were constrained with the LINCS algorithm.[48] An integration step size of 2 fs was used with coordinates stored every 100 ps. As noted above, this study includes a total sampling time of over 800 μs, and a total equilibrium sampling of over 250 μs. For simplicity, secondary structure was assessed primarily using the widely recognized Dictionary of Secondary Structure in Proteins, DSSP.[49]

Acknowledgements

We thank Bojan Zagrovic for his critical reading of this manuscript. This work would not have been possible without the worldwide Folding@Home volunteers, who contributed invaluable processor time (http://folding.stanford.edu). E.J.S., M.R.S. and Y.M.R. were supported by Veatch, Fannie and John Hertz, and Stanford Graduate fellowships, respectively. The computation was supported by ACS-PRF (36028-AC4), NSF Molecular Biophysics, NSF MRSEC CPIMA (DMR-9808677), and a gift from Intel.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2005.11.058

References

1. Rose, G. D. & Wolfenden, R. (1993). Hydrogen bonding, hydrophobicity, packing, and protein folding. *Annu. Rev. Biophys. Biomol. Struct.* **22**, 381–415.
2. ten Wolde, P. R. & Chandler, D. (2002). Drying-induced hydrophobic polymer collapse. *Proc. Natl Acad. Sci. USA*, **99**, 6539–6543.
3. Cheung, M. S., Garcia, A. E. & Onuchic, J. N. (2002). Protein folding mediated by solvation: water expulsion and formation of the hydrophobic core occur after the structural collapse. *Proc. Natl Acad. Sci. USA*, **99**, 685–690.
4. Shea, J. E., Onuchic, J. N. & Brooks, C. L. (2002). Probing the folding free energy landscape of the src-SH3 protein domain. *Proc. Natl Acad. Sci. USA*, **99**, 16064–16068.
5. Zhou, R., Huang, X., Margulis, C. J. & Berne, B. J. (2004). Hydrophobic collapse in multidomain protein folding. *Science*, **305**, 1605–1609.
6. Zhou, R., Berne, B. J. & Germain, R. (2001). The free energy landscape for beta hairpin folding in explicit water. *Proc. Natl Acad. Sci. USA*, **98**, 14931–14936.
7. Zhou, R. (2003). Free energy landscape of protein folding in water: explicit *vs*. implicit solvent. *Proteins: Struct. Funct. Genet.* **53**, 148–161.
8. Sorin, E. J., Rhee, Y. M. & Pande, V. S. (2005). Does water play a structural role in the folding of small nucleic acids? *Biophys. J.* **88**, 2516–2524.
9. Rhee, Y. M., Sorin, E. J., Jayachandran, G., Lindahl, E. & Pande, V. S. (2004). Simulations of the role of water in the protein-folding mechanism. *Proc. Natl Acad. Sci. USA*, **101**, 6456–6461.
10. Nymeyer, H. & Garcia, A. E. (2003). Simulation of the folding equilibrium of α-helical peptides: a comparison of the generalized Born approximation with explicit solvent. *Proc. Natl Acad. Sci. USA*, **100**, 13934–13939.
11. Hummer, G., Garcia, A. E. & Garde, S. (2001). Helix nucleation kinetics from molecular simulations in explicit solvent. *Proteins: Struct. Funct. Genet.* **42**, 77–84.
12. Garcia, A. E. & Sanbonmatsu, K. Y. (2001). α-Helical stabilization by side chain shielding of backbone hydrogen bonds. *Proc. Natl Acad. Sci. USA*, **99**, 2782–2787.
13. Garcia, A. E. & Sanbonmatsu, K. Y. (2001). Exploring the energy landscape of a beta-hairpin in explicit solvent. *Proteins: Struct. Funct. Genet.* **42**, 345–354.
14. Makarov, V., Pettitt, B. M. & Feig, M. (2002). Solvation and hydration of proteins and nucleic acids: a theoretical view of simulation and experiment. *Accts Chem. Res.* **35**, 376–384.
15. Duan, Y. & Kollman, P. A. (1998). Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science*, **282**, 740–744.
16. Pande, V. S., Baker, I., Chapman, J., Elmer, S., Kaliq, S., Larson, S. *et al*. (2003). Atomistic protein folding simulations on the submillisecond timescale using worldwide distributed computing. *Biopolymers*, **68**, 91–109.
17. Sorin, E. J. & Pande, V. S. (2005). Exploring the helix–coil transition via all-atom equilibrium ensemble simulations. *Biophys. J.* **88**, 2472–2493.
18. Sorin, E. J. & Pande, V. S. (2005). Empirical force field assessment: the interplay between backbone torsions and non-covalent term scaling. *J. Comput. Chem.* **26**, 682–690.
19. Zagrovic, B., Sorin, E. J. & Pande, V. (2001). β-Hairpin folding simulations in atomistic detail using an implicit solvent model. *J. Mol. Biol.* **313**, 151–169.
20. Zimm, B. H. & Bragg, J. K. (1959). Theory of the phase transition between helix and random coil in polypeptide chains. *J. Chem. Phys.* **31**, 526–535.

21. Lifson, S. & Roig, A. (1961). Theory of helix–coil transition in polypeptides. *J. Chem. Phys.* **34**, 1963–1974.
22. Lopez, M. M., Chin, D.-H., Baldwin, R. L. & Makhatadze, G. I. (2002). The enthalpy of the alanine peptide helix measured by isothermal titration calorimetry using metalbinding to induce helix formation. *Proc. Natl Acad. Sci. USA*, **99**, 1298–1302.
23. Spek, E. J., Olson, C. A., Shi, Z. & Kallenbach, N. R. (1999). Alanine is an intrinsic α-helix stabilizing amino acid. *J. Am. Chem. Soc.* **121**, 5571–5572.
24. Marqusee, S., Robbins, V. H. & Baldwin, R. L. (1989). Unusually stable helix formation in short alanine based peptides. *Proc. Natl Acad. Sci. USA*, **86**, 5286–5290.
25. Yang, A.-S. & Honig, B. (1995). Free energy determinants of secondary structure formation. I. α-Helices. *J. Mol. Biol.* **252**, 351–365.
26. Lewis, R. N. A. H., Zhang, Y.-P., Hodges, R. S., Subczynski, W. K., Kusumi, A., Flach, C. R. *et al.* (2001). A polyalanine-based peptide cannot form a stable transmembrane α-helix in fully hydrated phospholipid bilayers. *Biochemistry*, **40**, 12103–12111.
27. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935.
28. Shirts, M. R. & Pande, V. S. (2005). Solvation free energies of amino acid side chain analogs for common molecular mechanics water models. *J. Chem. Phys.* **122**, 134508.
29. Sun, Y. & Kollman, P. A. (1995). Hydrophobic solvation of methane and nonbond parameters of the TIP3P water model. *J. Comput. Chem.* **16**, 1164–1169.
30. Mezei, M., Fleming, P. J., Srinivasan, R. & Rose, G. D. (2004). Polyproline II helix is the preferred conformation for unfolded polyalanine in water. *Proteins: Struct. Funct. Genet.* **55**, 502–507.
31. Williams, S., Causgrove, T. P., Gilmanshin, R., Fang, K. S., Callender, R. H., Woodruff, W. H. & Dyer, R. B. (1996). Fast events in protein folding: helix melting and formation in a small peptide. *Biochemistry*, **35**, 691–697.
32. Qian, H. & Schellman, J. A. (1992). Helix–coil theories: a comparative study for finite length polypeptides. *J. Phys. Chem.* **96**, 3987–3994.
33. Lindahl, E., Hess, B. & van der Spoel, D. (2001). GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Mod.* **7**, 306–317.
34. Eisenhaber, F., Lijnzaad, P., Argos, P., Sander, C. & Scharf, M. (1995). The double cubic lattice method: efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. *J. Comput. Chem.* **16**, 273–284.
35. Hiltpold, A., Ferrara, P., Gsponer, J. & Caflisch, A. (2000). Free energy surface of the helical peptide $Y(MEARA)_6$. *J. Phys. Chem. B*, **104**, 10080–10086.
36. Millhauser, G. L., Stenland, C. J., Hanson, P., Bolin, K. A. & van de Ven, F. J. M. (1997). Estimating the relative populations of 3_{10}-helix and α-helix in ala-rich peptides: a hydrogen exchange and high field NMR study. *J. Mol. Biol.* **267**, 963–974.
37. Armen, R., Alonso, D. O. V. & Daggett, V. (2003). The role of α-, 3_{10}-, and π-helix in helix–coil transitions. *Protein Sci.* **12**, 1145–1157.
38. Shi, Z., Olson, C. A., Rose, G. D., Baldwin, R. L. & Kallenbach, N. R. (2002). Polyproline II structure in a sequence of seven alanine residues. *Proc. Natl Acad. Sci. USA*, **99**, 9190–9195.
39. Garcia, A. E. (2004). Characterization of non-alpha helical conformations in Ala peptides. *Polymer*, **45**, 669–676.
40. Asher, S. A., Mikhonin, A. V. & Bykov, S. (2004). UV Raman demonstrates that α-helical polyalanine peptides melt to polyproline II conformations. *J. Am. Chem. Soc.* **126**, 8433–8440.
41. Drozdov, A. N., Grossfield, A. & Pappu, R. V. (2003). Role of solvent in determining conformational preferences of alanine dipeptide in water. *J. Am. Chem. Soc.* **126**, 2574–2581.
42. Kentsis, A., Mezei, M., Gindin, T. & Osman, R. (2004). Unfolded state of polyalanine is a segmented polyproline II helix. *Proteins: Struct. Funct. Genet.* **55**, 493–501.
43. Weise, C. F. & Weisshaar, J. C. (2003). Conformational analysis of alanine dipeptide from dipolar couplings in a water-based liquid crystal. *J. Phys. Chem. B*, **107**, 3265–3277.
44. Zagrovic, B., Lipfert, J., Sorin, E. J., Millett, I. S., Van Gunsteren, W. F., Doniach, S. & Pande, V. S. (2005). Unusual compactness of a polyproline type II structure. *Proc. Natl Acad. Sci. USA*, **102**, 11698–11703.
45. Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., Dinola, A. & Haak, J. (1984). Molecular-dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690.
46. Thompson, P. A., Eaton, W. A. & Hofrichter, J. (1997). Laser temperature jump study of the helix–coil kinetics of an alanine peptide interpreted with a 'kinetic zipper' model. *Biochemistry*, **36**, 9200–9210.
47. Ianoul, A., Mikhonin, A., Lednev, I. K. & Asher, S. A. (2002). UV resonance Raman study of the spatial dependence of α-helix unfolding. *J. Phys. Chem. A*, **106**, 3621–3624.
48. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. (1997). LINCS: A Linear Constraint Solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472.
49. Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577–2637.

Edited by G. von Heijne

(Received 7 October 2005; received in revised form 15 November 2005; accepted 18 November 2005)
Available online 5 December 2005