Reconciling the Solution and X-ray Structures of the Villin Headpiece Helical Subdomain: Molecular Dynamics Simulations and Double Mutant Cycles Reveal a Stabilizing Cation−π Interaction†

Lauren Wickstrom,[‡,§] Yuan Bi,[§,||] Viktor Hornak,[⊥] Daniel P. Raleigh,*[,‡,||,@] and Carlos Simmerling*[,‡,||,⊥]

Biochemistry and Structural Biology Program, Department of Chemistry, Center for Structural Biology, and Graduate Program in Biophysics, State University of New York, Stony Brook, New York 11794-3400

Received August 29, 2006; Revised Manuscript Received January 15, 2007

ABSTRACT: The 36-residue helical subdomain of the villin headpiece, HP36, is one of the smallest cooperatively folded proteins, folding on the microsecond time scale. The domain is an extraordinarily popular model system for both experimental and computational studies of protein folding. The structure of HP36 has been determined using X-ray crystallography and NMR spectroscopy, with the resulting structures exhibiting differences in helix packing, van der Waals contacts, and hydrogen bonding. It is important to determine the solution structure of HP36 with as much accuracy as possible since this structure is widely used as a reference for simulations and experiments. We complement the existing data by using all-atom molecular dynamics simulations with explicit solvent to evaluate which of the experimental models is the better representation of HP36 in solution. After simulation for 50 ns initiated with the NMR structure, we observed that the protein spontaneously adopts structures with a backbone conformation, core packing, and C-capping motif on the third helix that are more consistent with the crystal structure. We also examined hydrogen bonding and side chain packing interactions between D44 and R55 and between F47 and R55, respectively, which were observed in the crystal structure but not in the NMR-based solution structure. Simulations showed large fluctuations in the distance between D44 and R55, while the distance between F47 and R55 remained stable, suggesting the formation of a cation−π interaction between those residues. Experimental double mutant cycles confirmed that the F47−R55 pair has a larger energetic coupling than the D44−R55 interaction. Overall, these combined experimental and computational studies show that the X-ray crystal structure is the better reference structure for HP36 in solution at neutral pH. Our analysis also shows how detailed molecular dynamics simulations combined with experimental validation can help bridge the gap between NMR and crystallographic methods.

The villin headpiece helical subdomain (HP36),[1] the C-terminal portion of the villin headpiece, is the shortest naturally occurring sequence which has been shown to fold cooperatively (Figure 1). Infrared temperature jump (*1*), laser fluorescence (*2*, *3*), and NMR line shape analysis (*4*) techniques have been used to determine the folding of HP36 occurs on the microsecond time scale. Its rapid folding, small size, and simple topology of three helices have made this domain an extremely popular system for experimental (*1*−*13*) and computational (*14*−*30*) studies. Much of this work relies on using the folded structure as a reference, and thus, the accuracy of the known HP36 structure is particularly important.

Several structures have been determined for HP36, one by NMR and the others by X-ray crystallographic methods (*6*, *7*). These structures vary in the hydrophobic core packing, in interhelical H-bonds, and in the length of the helices. In addition, two potentially important side chain contacts differ significantly between the NMR and X-ray structures: F47−R55, 4.3 (X-ray) and 6.3 Å (NMR); and D44−R55, 2.7 (X-ray) and 7.9 Å (NMR) (Figure 2A,B). In the X-ray structure, the F47−R55 pair forms a van der Waals contact which could be particularly stabilizing as a cation−π interaction, while the D44−R55 pair forms a hydrogen bond (D44 Oδ1

† Support for this project was provided by the National Institutes of Health (Grant GM61678 to C.S.), by the Donors of the Petroleum Research Fund of the American Chemical Society [ACS-PRF (D.P.R.)], the National Science Foundation (MCB-0614365 to D.P.R.), and National Computational Science Alliance (NCSA) Grant MCA02N028 (C.S.), which provided resources at NCSA.

* To whom correspondence should be addressed. C.S.: e-mail, carlos.simmerling@stonybrook.edu; phone, (631) 632-1336; fax, (631) 632-1555. D.P.R.: e-mail, draleigh@notes.cc.sunysb.edu; phone, (631) 632-9547; fax, (631) 632-7960.

‡ Biochemistry and Structural Biology Program.

§ These authors contributed equally to this work.

|| Department of Chemistry.

⊥ Center for Structural Biology.

@ Graduate Program in Biophysics.

[1] Abbreviations: HP36, villin headpiece helical subdomain; DSSP, dictionary of secondary structure prediction; rmsd, root-mean-square deviation; GB, generalized Born; PME, particle mesh Ewald; IPS, isotropic periodic sum; MD, molecular dynamics; PMF, potential of mean force; D44N, Asp44 to Asn mutant of WT HP36; F47L, Phe47 to Leu mutant of WT HP36; R55M, Arg55 to Met mutant of WT HP36; D44N/R55M, Asp44 to Asn, Arg55 to Met double mutant of WT HP36; F47L/R55M, Phe47 to Leu, Arg55 to Met double mutant of WT HP36; RP-HPLC, reverse phase high-performance liquid chromatography; ESI-MS, electrospray ionization mass spectrometry; CD, circular dichroism; TMAO, trimethylamine *N*-oxide; $\Delta G°_U$, free energy of unfolding at 25 °C; $\Delta C_p°$, heat capacity change of protein unfolding.

10.1021/bi061785+ CCC: $37.00 © 2007 American Chemical Society
Published on Web 03/06/2007

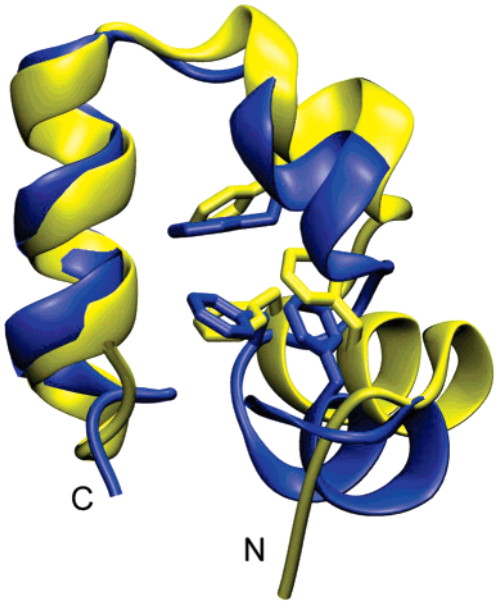

FIGURE 1: Two experimental structures of the villin helical subdomain, showing only the backbone (ribbons) and heavy atoms for the three phenylalanines in the core (F47, F51, and F58). The NMR structure of HP36 (PDB entry 1VII) is colored blue, and the X-ray structure (PDB entry 1YRF) is colored yellow. Differences in the backbone and the phenylalanine core packing are highlighted using a best fit alignment on backbone residues L62−F76.

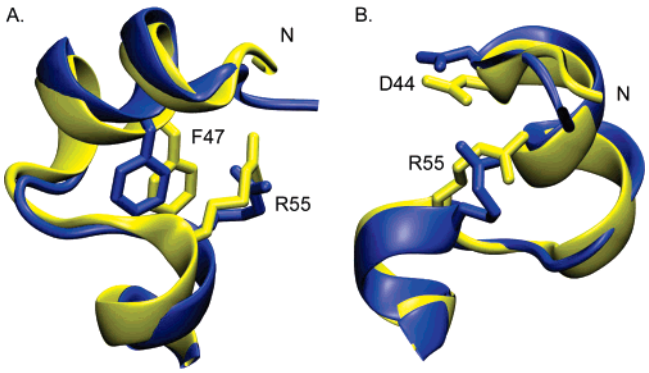

FIGURE 2: Comparison of side chain interactions in the X-ray and NMR structures, using a best fit alignment on residues L42−P62. (A) The R55 and F47 side chains are shown in both the NMR (blue) and X-ray structure (yellow). In the X-ray structure, R55 is involved in a van der Waals contact with F47 and a hydrogen bond with D44. (B) In the X-ray structure, Nϵ of R55 forms a hydrogen bond with Oδ1 of D44 in contrast to the NMR structure where the atoms are almost 8 Å apart. The N-terminus is labeled.

and R55 Nϵ). Neither contact is present in the NMR structure. These differences may arise from changes in the HP36 sequence used in the two sets of experiments, although this seems unlikely. The crystallographic study employed the N68H mutant of HP36 and also lacks the N-terminal methionine incorporated by the expression system used for the NMR study [note that we adopt the typical numbering convention (*5*, *6*) for HP36, in which L42 follows the N-terminal methionine]. Another possible reason for the structural differences could be the variation of experimental conditions such as pH or temperature. There was significant deviation in the pH between structural determinations; the NMR structure was determined at pH 3.7 in contrast to the more neutral conditions of the crystallography experiment (pH 6.7). An alternate explanation for the observed structural differences is that they arise from methodological limitation conditions; these frequently give rise to differences in structures of the same protein determined using different techniques. In general, NMR structures are less precise than X-ray structures, particularly if only homonuclear methods are used. Nevertheless, X-ray structures can suffer from effects due to crystal packing; the resulting contacts may have a local influence on conformational preferences. The small size of HP36 and its correspondingly large surface area to volume ratio could make crystal contacts play an important role. On the other hand, crystallographic data are often collected at low temperatures which might result in the dampening of thermal motions that are present under physiological conditions.

Many computational studies have used HP36 as a model system for the development and validation of protein folding methods and for optimization of force field parameters (*14*−*30*). If the native reference structure is not correct, the basis of these studies may not be valid. For example, the structure of the first helix and the structure of the C-terminus vary in the ensemble of NMR structures (*6*, *27*), and many MD studies have therefore neglected these regions of the experimental structure when evaluating their success. Nevertheless, most simulations are performed at neutral pH, and thus, it is not clear if the simulations should be compared to the NMR structure determined at pH 3.7. A better structural model for neutral conditions would be invaluable in pursuing further work in understanding the folding and stability of this important model system for protein folding.

Accurate computational studies can provide an alternate method of studying conformational behavior and alleviate the uncertainty about which structure is the better representation of the folded state in solution. In principle, molecular dynamics (MD) simulations can supply detailed information with spatial and time resolution that exceed the ability of NMR and X-ray experiments, providing insight into the role of specific interactions that may not be readily accessible through experiments that probe averages over rapidly interconverting ensembles.

Here, we conducted all-atom MD simulation in explicit solvent using the NMR and X-ray structures of HP36 to gain insight into the details of the folded state in solution. The simulations diverge from the initial NMR structure and spontaneously adopt a structure that is much more similar to the X-ray structure, suggesting that the X-ray structure is a more accurate representation of the structure in solution at neutral pH. In addition, two residue pairs, D44−R55 and F47−R55, spontaneously formed contacts during the simulation, with the F47−R55 pair appearing to be more stable. These interactions were reported in the crystal structure but were not present in the ensemble of structures generated by the NMR studies. Thus, we conclude that the F47−R55 pair may play an important role in stabilizing HP36 in solution. We acknowledge that simulation models can be limited in accuracy, and any predictions should be tested through direct experimentation. To validate our computational observations, we employed an experimental double mutant cycle analysis. The results are consistent with our simulation data and suggest that the interaction between F47 and R55 plays a role in stabilizing the native state through a cation−π interaction. Overall, the results show how properly validated

MD simulations can provide an avenue to testing the stability and validity of structural models that were derived from experimental data.

MATERIALS AND METHODS

Computational. The numbering system corresponds to that used for the full-length villin headpiece, with the sequence of residues M41−F76 (MLSDEDFKAVFGMTRSAFANLPWKQQNLKKEKGLF). HP36 has free N- and C-termini that were modeled in the charged state. This sequence and termini correspond to those used in the experimental studies. All side chains for Asp, Glu, Lys, and Arg were charged during the simulation. All calculations employed Amber version 8 (*31*) and used the ff99SB modification (*32*) of the Amber ff99 force field (*33*, *34*). SHAKE (*35*) was used to constrain bonds involving hydrogen. The time step was 2 fs. The temperature was maintained using the weak coupling algorithm (*36*) with a thermostat of 37 °C (310 K), and the pressure was equilibrated to 1 atm. All production simulations were performed using the NVT ensemble. An independent simulation using the NPT ensemble provided similar results (data not shown).

Solvation plays a key role in biomolecular structural preferences, and thus, accurate treatment of solvation is essential for the investigation of structural propensities in simulations. Explicit solvent models can be highly effective, particularly when water has nonbulk properties and interacts directly with the solute (*37*). Implicit models such as the semianalytical generalized Born model (GB) (*38*) are attractive because they are computationally less expensive and can converge more rapidly than simulations in explicit water due to a lack of solvent viscosity. While the GB has been widely used for protein folding studies by a number of groups, other investigators have reported poor results, including secondary structural bias and ion pairing issues (*39*−*41*). Our previous studies on fragments of HP36 have shown that the use of explicit water produced results which were much more consistent with experimental trends than those obtained with implicit solvent (*30*). Consequently, we used explicit solvent in our simulations of HP36, in a truncated octahedral box using periodic boundary conditions with particle mesh Ewald (PME) (*42*) and a direct space cutoff of 8 Å. To investigate the influence of long-range periodicity, two additional simulations were run: one with the isotropic periodic sum (IPS) (*43*) nonlattice method with a cutoff of 8 Å and another with an atom-based nonbonded cutoff of 12 Å with no smoothing function. Simulations were initiated from the NMR structure (PDB entry 1VII) surrounded by 2327 TIP3P (*44*) waters molecules and equilibrated at 310 K for 50 ps with harmonic restraints on solute atoms, followed by minimization with gradually reduced positional restraints and three 5 ps MD simulations with gradually reduced restraints. The production simulations of the NMR structure were 50 ns in length for two PME simulations with different random seeds for assignment of velocities and 30 ns for the IPS and cutoff simulations, respectively. As a control, the X-ray structure (PDB entry 1YRF) was set up with the same amount of waters and equilibrated in a similar fashion. This simulation was run for 30 ns using PME.

Data Analysis. The last 5 ns of the simulation was used for cluster analysis and DSSP calculations. Cluster analysis was performed with Moil-view (*46*) using all atoms as a similarity criterion with average linkage. Clusters were formed with the bottom−up approach using a similarity cutoff of 2.5 Å. DSSP analysis and calculation of distances, rmsd values, and radii of gyration were carried out using the ptraj module in Amber. Distances between side chains were calculated using selected heavy atoms as indicated in the text. PMFs were calculated according to eq 1:

$$\Delta G^\circ = -RT \ln(N_i/N_0) \quad (1)$$

where N_i is the population of a particular histogram bin along the reaction coordinates that were employed and N_0 is the most populated bin. Error bars were estimated for the PMF by averaging two independent simulations and subsequently subtracting the PMF of an individual simulation from the average PMF.

Cloning, Expression, and Purification. The plasmid (pET3a-NTL9-FXa-HP36) containing the gene for HP36 was prepared as described previously (*45*). The primers for generation of the D44N, F47L, and R55M mutants were from Operon. The fusion proteins with coded amino acid substitutions were expressed and purified as described previously (*45*). For the purification of the fusion protein of F47L and F47L/R55M, ethylenediaminetetraacetic acid and phenylmethanesulfonyl fluoride were added to final concentrations of 1 mM to prevent degradation by proteases during ion exchange chromatography. Cleavage of wild-type (WT) HP36, D44N, F47L, R55M, D44N/R55M, and F47L/R55M from fusion proteins was carried out by incubating the sample with Factor Xa in buffer [50 mM Tris-HCl, 100 mM NaCl, and 5 mM $CaCl_2$ (pH 8.0)] for 16 h. The conditions were as follows: WT HP36, D44N, R55M, and D44N/R55M, 2 units of Factor Xa at 23 °C; F47L, 1 unit of Factor Xa at 4 °C; and F47L/R55M, 0.5 unit of Factor Xa at 4 °C. The solutions were purified by RP-HPLC as described. All proteins were more than 95% pure. The identities of the proteins were analyzed by electrospray ionization mass spectrometry (ESI-MS). The expected and observed molecular weights were as follows: WT HP36, expected 4190.9, observed 4190.1; D44N, expected 4188.9, observed 4189.0; F47L, expected 4155.8, observed 4157.0; R55M, expected 4164.9, observed 4166.4; D44N/R55M, expected 4163.9, observed 4164.3; and F47L/R55M, expected 4130.8, observed 4130.2.

Circular Dichroism (CD) Spectroscopy. All circular dichroism (CD) experiments were carried out on an Aviv 62A DS and a 202SF circular dichroism spectrophotometer. All samples for CD experiments were prepared in 10 mM sodium acetate and 150 mM sodium chloride buffer solution at pH 5.0. The protein concentrations were determined from absorbance measurements using the method of Pace and coworkers (*47*). The absorbance was measured at 280 nm in 6 M guanidine hydrochloride and 20 mM sodium phosphate (pH 6.5), with an extinction coefficient of 5690 M^{-1} cm^{-1}. Thermal unfolding and urea denaturation experiments were carried out in a 1 cm quartz cuvette by monitoring the signal at 222 nm. Thermal unfolding experiments were performed from 2 to 98 °C with a 2 °C interval. The concentrations of urea were measured by measuring the refractive index. The concentration of urea was increased from 0 to 10 M by ∼0.25 M each step. Urea denaturation experiments were performed at 25 °C. Urea denaturation experiments in the osmolyte

trimethylamine *N*-oxide (TMAO) for F47L and F47L/R55M were also performed at 25 °C. The concentrations of TMAO were measured by measuring the refractive index (*48*).

All thermal and urea denaturations were analyzed by a nonlinear least-squares curve fitting equation using SigmaPlot (SPSS Inc.):

$$y(T) = \frac{a_n + b_n T + (a_d + b_d T)e^{-\Delta G^\circ_U(T)/RT}}{1 + e^{-\Delta G^\circ_U(T)/RT}} \quad (2)$$

$$y([\text{urea}]) = \frac{a_n + b_n[\text{urea}] + (a_d + b_d[\text{urea}])e^{-\Delta G^\circ_U([\text{urea}])/RT}}{1 + e^{-\Delta G^\circ_U([\text{urea}])/RT}} \quad (3)$$

The free energy of unfolding is assumed to be linearly proportional to the concentration of denaturant:

$$\Delta G^\circ_U([\text{urea}]) = \Delta G^\circ_U(H_2O) - m[\text{urea}] \quad (4)$$

where $\Delta G^\circ_U([\text{urea}])$ is the apparent free energy for the native to denatured transition, $\Delta G^\circ_U(H_2O)$ is the free energy of unfolding in the absence of denaturant, m is the slope [eq 3 shows the linear dependence of $\Delta G^\circ_U([\text{urea}])$ on the denaturant concentration, and the m value reveals quantitative information about the buried surface area change upon unfolding], T is the temperature, R is the gas constant, a_n and b_n represent the intercept and slope of the pretransition, respectively, and a_d and b_d represent intercept and slope of the post-transition, respectively. The fraction of unfolded state is:

$$P_d = \frac{y(T) - (a_n + b_n T)}{a_d + b_d T - (a_n + b_n T)} \quad (5)$$

$$P_d = \frac{y([\text{urea}]) - (a_n + b_n[\text{urea}])}{a_d + b_d[\text{urea}] - (a_n + b_n[\text{urea}])} \quad (6)$$

RESULTS

Simulations of the NMR Structure. Figure 3 shows the backbone rmsd versus time and rmsd distributions calculated during the last 5 ns for selected regions of HP36 during the simulation. The rmsd is shown relative to both the NMR and X-ray structures. At the end of the equilibration period, the backbone rmsd (residues L42−L75) from each experimental structure was ∼2.0 Å (Figure 3A). At 8 ns, a structural transition occurred causing the overall backbone rmsd (X-ray) to drop 1.0 Å below the rmsd (NMR). This greater similarity to the X-ray structure persisted throughout the remainder of the simulation.

In Figure 3B, the rmsd relative to the X-ray structure of the region containing helix 3 (residues P62−F76) demonstrates even more clearly a switch during the simulation from similarity to the initial NMR structure to a greater similarity to the X-ray structure, as indicated by a reduction in the rmsd to the X-ray structure from 3−4 to 0.5−1.0 Å. Clearly, the simulation shows the inclination of HP36 to sample structures with a backbone similar to the X-ray structure despite being initiated with the NMR solution structure. The rmsd values for the two other helices remained stable and also showed a clear preference for the X-ray structure (Figure S1A,B).

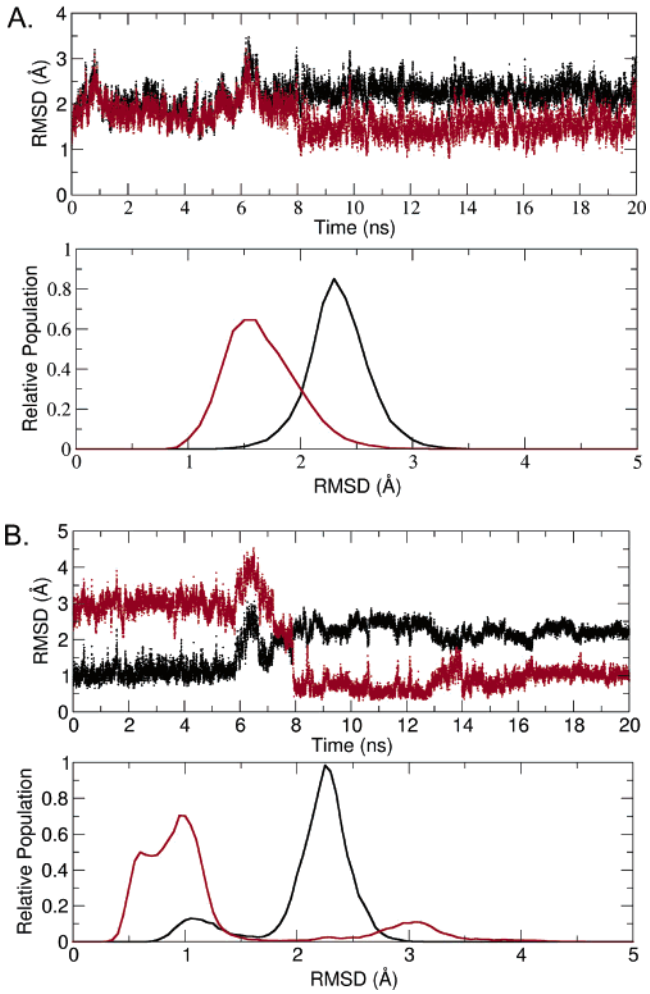

FIGURE 3: Time evolution and histogram distributions of the heavy atom backbone rmsd of (A) residues L42−F76 and (B) residues P62−F76 during the simulation of the NMR structure. Each calculation was performed using both the NMR (black) and X-ray (red) structures as the reference. A transition occurs near 8 ns, resulting in rmsd values lower when compared to the X-ray structure and higher when compared to the NMR structure. The C-terminal region (B) shows a particularly dramatic change from the initial NMR structure to one that much more closely matches the X-ray structure.

To investigate the source of the large reduction in rmsd relative to the X-ray structure, a best fit alignment was performed on residues 61−74 to compare the differences before and after the structural transition. In Figure 4A, the NMR, X-ray, and simulation structures are shown. The conformations of the C-terminus differ significantly between the X-ray and NMR structure. The simulation structure spontaneously converts from the conformation in the NMR structure to that in the X-ray structure, concomitant with formation of three hydrogen bonds that stabilize the observed conformation. G74 forms a C-capping interaction with K70 and K71 at the end of helix 3, along with an additional hydrogen bond formed between K70 and L75. Figure 4B shows the time evolution of these hydrogen bond distances. In the beginning of the simulation, all three distances are 4−9 Å. At 8 ns, the distances are reduced to 2−3 Å, indicating formation of the hydrogen bonds may play an important role in stabilizing the C-terminal helix. Importantly, all three hydrogen bonds are present in the X-ray structure but absent in the NMR structure (Figure 4A).

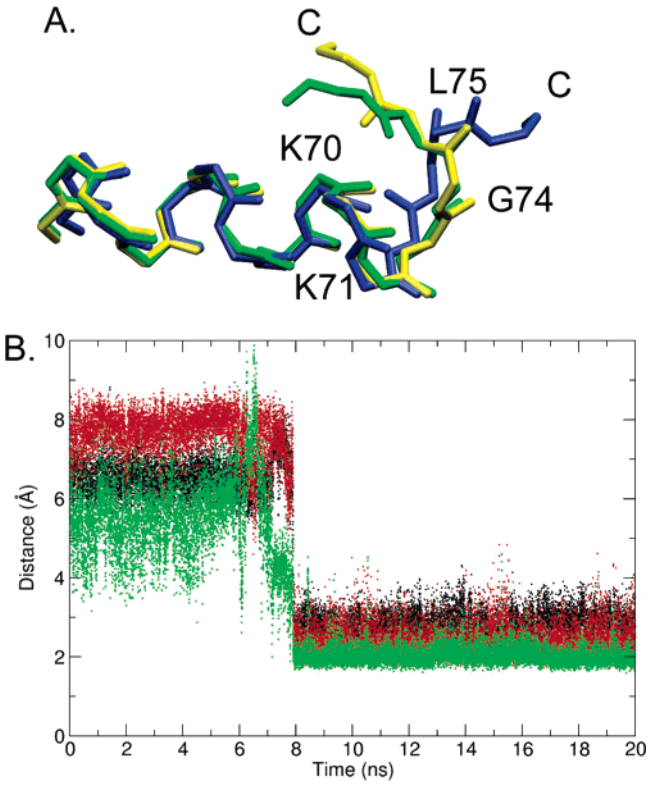

FIGURE 4: (A) Comparison of the C-terminal region (P62−F76) in the X-ray (yellow), NMR (blue), and simulation (green) structures. A key difference between the NMR and X-ray structures is the absence in the NMR structure of a C-capping motif on helix 3 observed in the X-ray structure. This motif is spontaneously adopted in simulations initiated from the NMR structure. (B) The C-capping motif involves three backbone hydrogen bonds (K70−G74 in black, K71−G74 in red, and K70−L75 in green) that are formed at ~8 ns and stable throughout the remainder of the simulation.

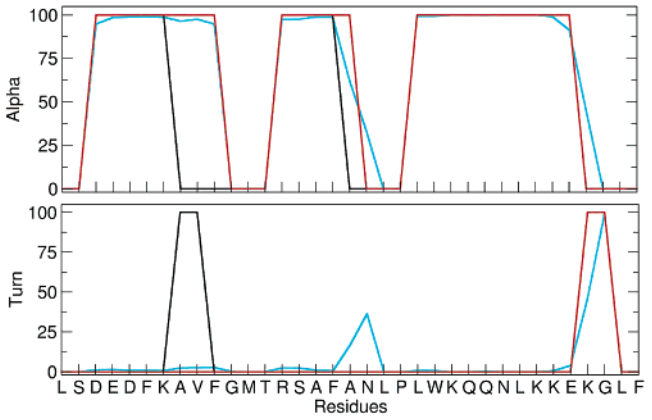

FIGURE 5: DSSP analysis of the NMR (black), X-ray (red), and simulation (cyan) structures of HP36: (A) α-helical content per residue and (B) turn content per residue. Overall, helix 2 and helix 3 are nearly the same length in the X-ray and NMR structures, but helix 1 is three residues longer in the X-ray structure than in the NMR structure. The α-helical content of the MD simulation is in very good agreement with the X-ray structure, though it was initiated from the NMR structure.

The dictionary of secondary structural prediction (DSSP) (*49*) analysis was employed to characterize the secondary structure in the simulation to facilitate comparisons with the X-ray and NMR structures (Figure 5). In the simulations, helix 1 spans the same eight residues as found in the X-ray structure (D44−F51), while the NMR structure contained only a five-residue helix from D44 to K48. Thus, the simulation significantly extends the length of the first helix, in agreement with the X-ray structure. Overall, the locations of the sequences of helices 2 and 3 are similar in the NMR and X-ray structures, although helix 2 is one residue shorter in the NMR structure (residues R55−F58 for the NMR structure vs R55−A59 for the X-ray structure). In the simulation, helix 2 appears to be consistent with both experimental structures; full α-helical content is sampled for residues 55−58, with a partial helical content (~50%) observed for A59. This may indicate that the C-terminus of the longer helix in the X-ray structure frays at the temperature of the NMR experiment. In both the NMR and X-ray structures, the α-helical content is the same for helix 3 (L63−K72). The simulations sample the same helix, with residue K73 sampling a partial population of helical structures. As noted above, the simulation spontaneously adopts a C-capping motif for this helix that is present in the X-ray structure. Overall, the α-helical structural content of the structures in the simulation is in much better agreement with the X-ray structure, particularly in helix 1.

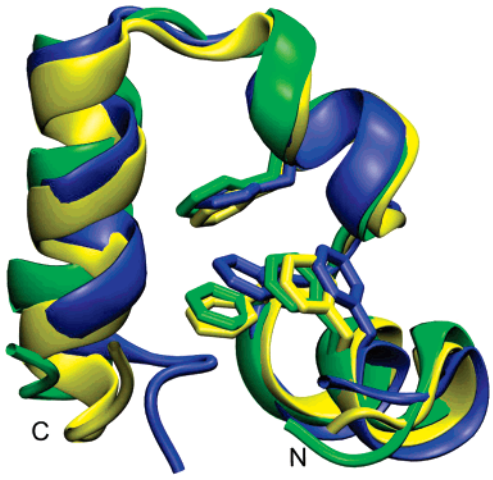

FIGURE 6: Comparison of backbone and core packing in the simulation (green), NMR (blue), and X-ray (yellow) structures, highlighting the differences in the core packing. A best fit alignment was performed on residues L42−P62. The packing of the phenylalanine core in the structure from the simulation structure is in much better agreement with the X-ray structure than with the NMR structure.

All-atom cluster analysis was used to generate a representative simulation structure using the last 5 ns of the trajectory. This structure has backbone and all-atom rmsd values relative to the X-ray structure of 1.5 and 2.7 Å (residues 42−75), respectively, while the rmsd values relative to the initial NMR structure were higher (2.3 Å for the backbone structure and 3.3 Å for the all-atom structure). Figure 6 shows all three structures after a best fit of the backbone from residue 42 to 62 (helices 1 and 2). Notably, the X-ray and simulation structures have a very similar spatial arrangement of their phenylalanine cores. In contrast to the X-ray and simulation structures, the NMR structure has F51 shifted more into the core. Thus, the backbone and core of the protein in the solution simulation possess structural features that are much more similar to the X-ray structure despite being initiated from the NMR structure.

Structural Similarities to the NMR Family. Given the diversity among the family of structures determined using

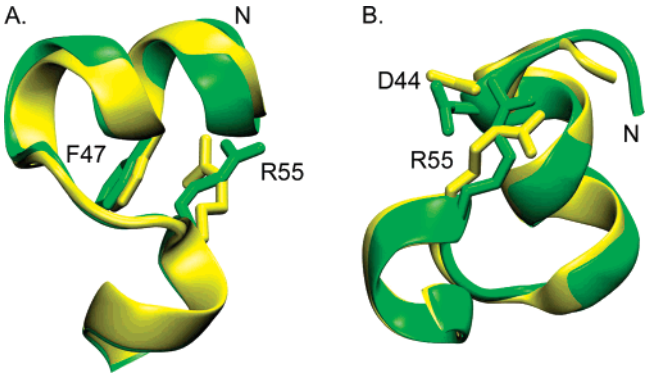

FIGURE 7: Comparison of selected side chain interactions in the simulation structure (green) and the X-ray structure (yellow). A best fit alignment was performed on residues L42−P62. In the simulation structure, R55 is 4.7 Å from the base of the phenylalanine ring (A) and 2.7 Å from Oδ1 of D44 (B). This suggests that both contacts may play a role in the stability of the protein.

the NMR data, it is reasonable to expect that some of them may be more similar than others to the X-ray structure. Figure S2 shows the backbone rmsd as compared to the X-ray, simulation, and NMR average structures for each structure in the NMR family. Overall, the individual NMR structures are all more similar to the NMR average than to the X-ray structure (average rmsd values of 1.7 and 2.4 Å, respectively). The rmsd values of the three individual helices demonstrate similar differences. However, some of the individual members of the NMR family are similar to the X-ray and simulation structures, especially in helix 1. According to the DSSP, seven of the 29 members of the NMR family sample α-helical conformations at V50 (data not shown) which is outside of the helical region in the average NMR structure. This suggests that extension of helix 1 beyond the range seen in the average structure remains consistent with the NMR family. However, the overall backbone of the X-ray and the simulation structure differs from all of the structures in the NMR family (Figure S2).

Specific Side Chain Interactions. There are several specific side chain interactions which differ in the NMR and X-ray structures. In the X-ray structure, R55 forms a van der Waals interaction with F47 and an interhelical side chain−side chain hydrogen bond with D44 (D44 Oδ1 and R55 Nϵ); both interactions are absent in the NMR structure. In panels A and B of Figure 7, the simulation structure was aligned with the X-ray structure to highlight the similarities in the interaction of those particular side chains. Since the simulation structure is a single snapshot, we also investigated the behavior of these contacts as a function of time during the MD run, observing fluctuations in both cases (Figure 8). In both the X-ray and simulation structures, the H-bond distance between D44 and R55 is 2.7 Å, in contrast with the much longer distance of 7.9 Å in the NMR structure. This specific contact also samples a range of distances from 6.7 to 11.6 Å in the family of NMR structures (Figure S3). During the simulation, this hydrogen bond is broken and re-formed multiple times, suggesting that a reasonable description of the equilibrium distance distribution has been sampled (Figure 8A). We used histogram analysis to calculate the potential mean force (PMF) for the pair to quantify the stability of the contact in the native state. While two free energy minima are located within hydrogen bonding distance, two other local minima at 5.0 and 7.0 Å have relative energies of <0.6 kcal/mol compared to the contact pair (Figure 8B). Thus, breaking this contact is expected to be a readily accessible thermal fluctuation. The stability of the contact between F47 and R55 was evaluated by measuring the distance from Cγ of F47 to Nϵ of R55 (Figure 8C). This distance had comparable values in the simulation and X-ray structures (4.7 and 4.3 Å, respectively), while a much longer distance of 6.3 Å is observed in the average NMR structure. Only two structures in the entire NMR family sample a contact distance of less than 5.5 Å (Figure S3). In contrast with the D44−R55 pair, the PMF for formation of the F47−R55 contact shows only a single minimum at 5.5 Å (Figure 8D). Overall, this suggests that R55 has a much more stable interaction with F47 than the salt bridge that it forms with D44.

Simulations of the X-ray Structure. Figure S4 shows the backbone rmsd versus time and rmsd distributions calculated during the simulation starting from the X-ray structure. The rmsd is shown relative to the X-ray, NMR, and simulation (from NMR) structures. After equilibration, the simulation samples backbone conformations (S43−L75) with an average rmsd relative to the X-ray structure of 1.5 Å and remains quite stable through the 30 ns duration. Overall, there is a preference to adopt structures comparable to the simulation structure discussed above rather than the NMR structure (the rmsd compared to the simulation-equilibrated NMR structure is 1.5 Å less than the rmsd to the original NMR structure). Individual helices demonstrate comparable preferences for the X-ray and simulation structures (data not shown). Hence, the simulations starting from the NMR and X-ray structures both converge to a common simulation structure that is much closer to the X-ray structure than the NMR structure.

Experimental Investigation of the Putative Side Chain Interactions. While simulations can provide a detailed view of molecule structure and dynamics, many approximations are involved, necessitating validation through experimentation. A set of single mutants and double mutants were prepared to probe the putative side chain interactions involving D44−R55 and F47−R55 pairs. D44 was mutated to Asn, F47 to Leu, and R55 to Met. Thermal unfolding experiments were performed for wild-type HP36 (WT HP36) and for each of the mutants at pH 5.0 (Figure 9A and Table 1). WT HP36 has a transition midpoint (T_m) of 73.0 °C, while all the variants exhibit a lower melting temperature. The T_m values of D44N, F47L, R55M, D44N/R55M, and F47L/R55M are 57.8, 45.6, 67.3, 55.4, and 35.3 °C, respectively. From the thermal unfolding curves, at 25 °C, 22% of the population of F47L and 40% of the population of F47L/R55M are unfolded.

Urea denaturation experiments were also carried out in 10 mM sodium acetate and 150 mM sodium chloride at 25 °C to determine the free energy of unfolding. The estimated free energy for unfolding ($\Delta G°_U$)was 3.22 kcal/mol for WT HP36, 2.48 kcal/mol for D44N, 2.19 kcal/mol for R55M, and 1.74 kcal/mol for D44N/R55M (Figure 9B and Table 1). The F47L and F47L/R55M mutants were so unstable that the native baseline was not observed (Figure 9B) and the unfolding free energy could not be accurately measured by urea denaturation. Thermal and urea denaturation experiments showed that F47L and F47L/R55M are partially

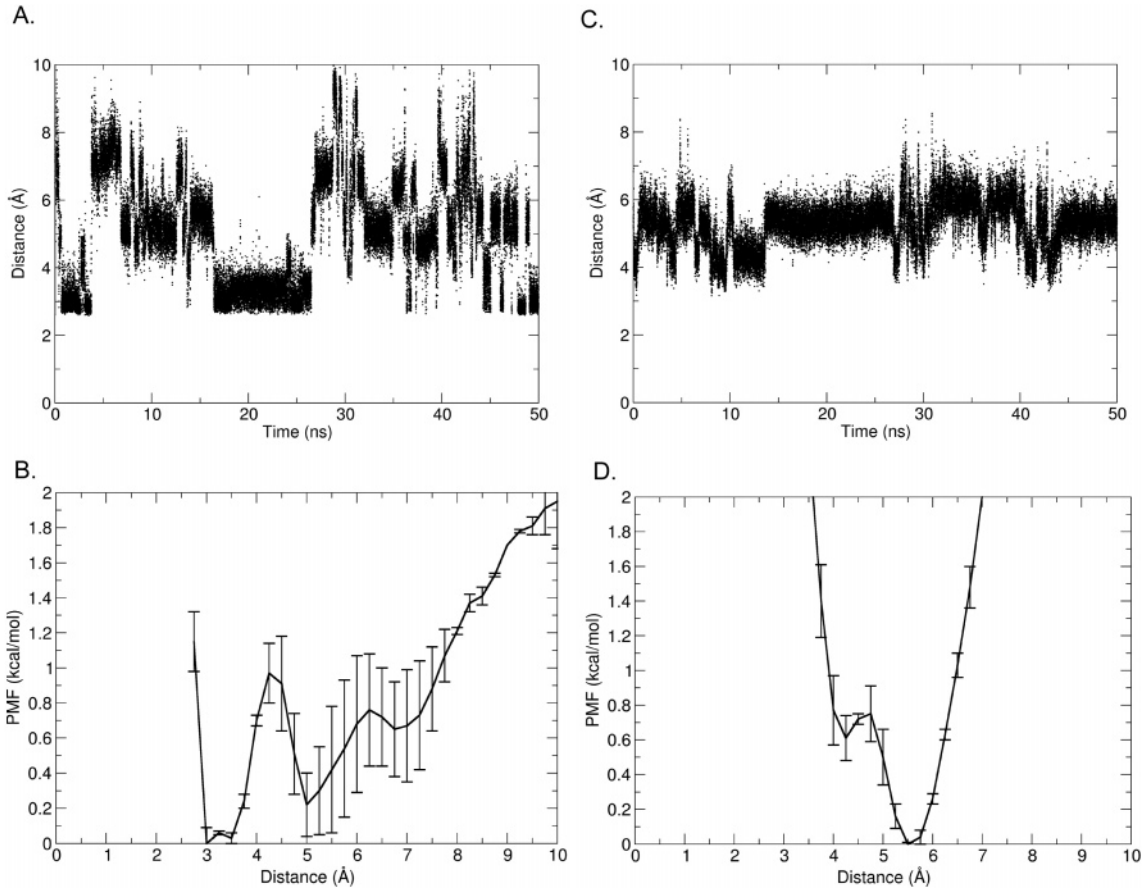

FIGURE 8: Time evolution and PMFs of specific contact distances involving (A and C) R55 and D44 and (B and D) R55 and F47. The distance between R55 and D44 fluctuates throughout the trajectory and shows multiple shallow free energy minima. In contrast, the distance measuring the contact between R55 and F47 is stable during the entire trajectory, with a single free energy minimum at 5.5 Å. The results indicate that the R55−F47 contact is the more stable of these two residue pairs. Error bars are described in Methods.

unfolded in the absence of urea at 25 °C. Previous studies have shown that TMAO can stabilize partially or completely unfolded proteins (*50*). Therefore, the combination of urea denaturation and TMAO stabilization can be utilized to estimate the stability of marginally stable proteins. To determine the unfolding free energy of F47L and F47L/R55M, we performed urea denaturation experiments with increasing TMAO concentrations. For F47L, the titration curves show good pre- and post-transitions at different TMAO concentrations (Figure 10A). With increasing TMAO concentrations, the urea denaturation curves shifted to higher urea concentrations. The free energy of unfolding at each TMAO concentration was measured. $\Delta G°_U$ ranges from 1.27 kcal/mol in 1.62 M TMAO to 1.67 kcal/mol in 2.50 M TMAO (Table S1). Mello and co-workers (*50*) have shown that the free energy of unfolding depends linearly on TMAO concentration. The extrapolated $\Delta G°_U$ of F47L at 0 M TMAO was estimated to be 0.52 kcal/mol at 25 °C (Figure 10B), which is in reasonable agreement with the value estimated from the thermal unfolding curve.

Unfortunately, the same strategy could not be applied to the F47L/R55M double mutant. High TMAO concentrations are necessary to stabilize the protein for detection of the pre-transition, but comparatively high urea concentrations are needed to observe the post-transition. Therefore, it is very difficult to find conditions under which full unfolding curves could be measured. Thus, we extrapolated from the thermal unfolding data using the Gibbs−Helmholtz equation:

$$\Delta G°_U(T) = \Delta H°(T_m)\left(1 - \frac{T}{T_m}\right) - \Delta C_p°\left[T_m - T + T\ln\left(\frac{T}{T_m}\right)\right] \quad (7)$$

This calculation requires knowledge of the heat capacity change, $\Delta C_p°$. HP36 is small, resulting in a very broad differential scanning calorimetry (DSC) transition, which makes it very difficult to calculate the heat capacity accurately by DSC. From the literature, the $\Delta C_p°$ of unfolding is expected to be ~0.012 kcal mol^{-1} K^{-1} per residue of protein (*51*). To a first approximation, the $\Delta C_p°$ for HP36 can be calculated to be 0.43 kcal mol^{-1} K^{-1}. Another small 41-residue helical protein, the peripheral subunit-binding domain, has a $\Delta C_p°$ value of 0.43 kcal mol^{-1} K^{-1} (*52*), suggesting that the estimate for HP36 is reasonable. To check whether the value of $\Delta C_p°$ significantly affects the results, we use heat capacities ranging from 0.30 to 0.70 kcal mol^{-1} K^{-1} to calculate the $\Delta G°_U$. F47L/R55M has a measured T_m of 35.3 °C and a $\Delta H°(T_m)$ of 9.5 kcal/mol, and the resulting calculated $\Delta G°_U$ of F47L/R55M at 25 °C ranged from 0.19 to 0.28 kcal/mol depending on the value of $\Delta C_p°$ that is used (Table S2). The value of $\Delta G°_U$ estimated from the Gibbs−Helmholtz equation is in good agreement with the fraction unfolded determined directly from the fit to the thermal melt.

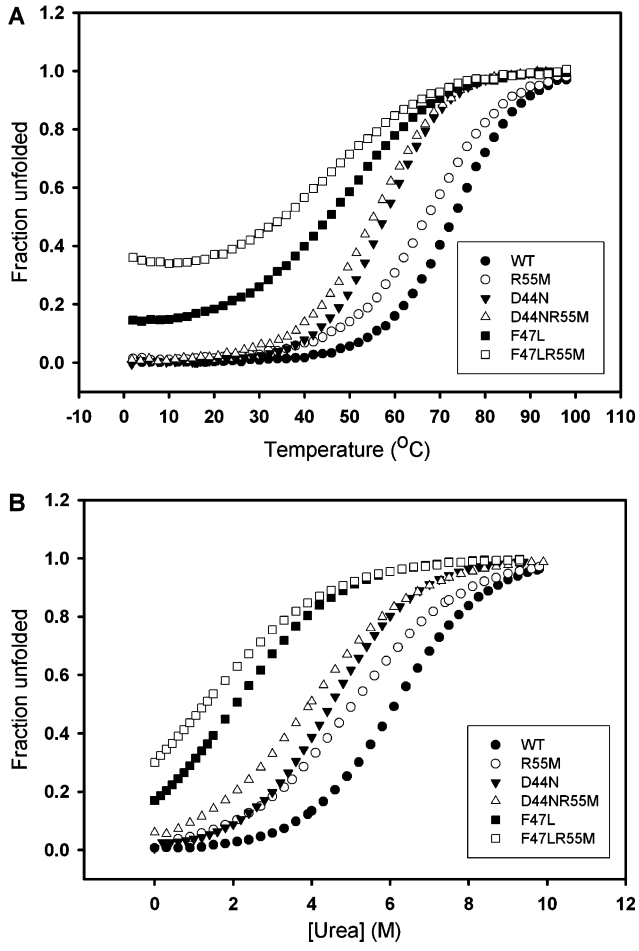

FIGURE 9: (A) Thermal unfolding curves for WT HP36 and its mutants and (B) urea unfolding curves for WT HP36 and its mutants: WT HP36 (●), R55M (○), D44N (▼), D44N/R55M (△), F47L (■), and F47L/R55M (□).

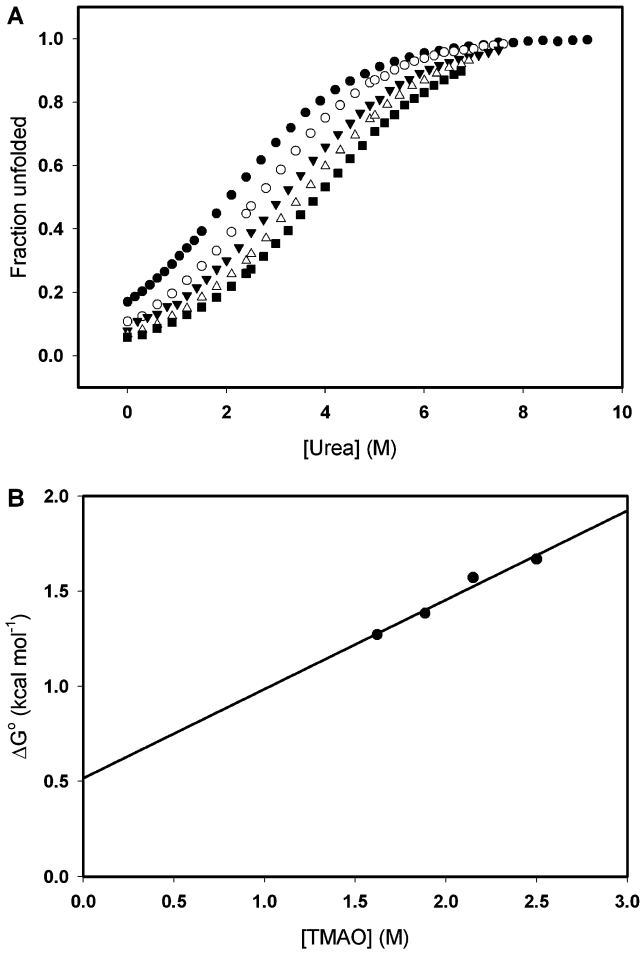

FIGURE 10: (A) Unfolding transitions of the F47L mutant in a mixed urea/TMAO cosolvent monitored by circular dichroism at 222 nm. Urea denaturation in increasing TMAO concentrations (from left to right, 0, 1.60, 1.88, 2.15, and 2.50 M TMAO). (B) Dependence of the unfolding free energy on TMAO concentration for the F47L mutant. Parameters are obtained by fitting urea unfolding curves at different TMAO concentrations. The straight line is the result of linear regression to each parameter.

The coupling free energy between the D44 or F47 side chain and the R55 side chain, $\Delta\Delta G^\circ_{coupling}$, was calculated using eq 8

$$\Delta\Delta G^\circ_{coupling} = (\Delta G^\circ_{WT} - \Delta G^\circ_{Double}) - [(\Delta G^\circ_{Single} - \Delta G^\circ_{Double}) + (\Delta G^\circ_{R55M} - \Delta G^\circ_{Double})] \quad (8a)$$

where ΔG°_{WT} and ΔG°_{R55M} are the free energies of unfolding for the wild-type protein and the R55M single mutant, respectively, ΔG°_{Single} represents that for a D44N or F47L single mutants, and ΔG°_{Double} represents that for the D44N/R55M or F47L/R55M double mutant.

The relationship can be rearranged to a simpler form:

$$\Delta\Delta G^\circ_{coupling} = \Delta G^\circ_{WT} - \Delta G^\circ_{Single} - \Delta G^\circ_{R55M} + \Delta G^\circ_{Double} \quad (8b)$$

Using the ΔG°_U values (Table 1) measured from experiments, the coupling free energy between the D44 side chain and the R55 side chain was close to zero (0.29 ± 0.20 kcal/mol). In contrast, the coupling free energy between the F47 side chain and R55 side chain ranged from 0.70 ± 0.20 to 0.79 ± 0.20 kcal/mol. The different estimates arise from using different ΔC_p° values to calculate ΔG°_{Double}. The analysis shows that there is a non-zero coupling between F47 and R55.

Table 1: Summary of Equilibrium Stability Measurements for WT HP36 and Its Mutants in 10 mM Sodium Acetate and 150 mM Sodium Chloride at pH 5.0 and 25 °C

protein	T_m (°C)	$\Delta H^\circ(T_m)$ (kcal/mol)	$\Delta G^\circ_U(H_2O)$ (kcal/mol)	m (kcal mol^{-1} M^{-1})
WT HP36	73.0	31.8	3.22	−0.52
D44N	57.8	32.1	2.48	−0.55
F47L	45.6	15.8	0.52[a]	−0.45[b]
R55M	67.3	26.3	2.19	−0.43
D44N/R55M	55.4	27.4	1.74	−0.44
F47L/R55M	35.3	9.8	0.19−0.28[c]	not available

[a] The $\Delta G^\circ_U(H_2O)$ values of F47L is extrapolated from urea denaturation at different TMAO concentrations. [b] m is the average value of m from urea denaturation at different TMAO concentrations. [c] The $\Delta G^\circ_U(H_2O)$ value of F47L/R55M is calculated from the Gibbs−Helmholtz equation using ΔC_p° values ranging from 0.30 to 0.70 kcal mol^{-1} K^{-1}.

DISCUSSION

The explicit water MD simulation starting from the NMR structure showed a clear preference to sample structures that are much more similar to the X-ray structure, as indicated by rmsd values, DSSP analysis, packing of the phenyalanine core, formation of a C-capping motif on helix 3, and adoption

of specific contacts between side chains. Double mutant cycle experiments were performed and demonstrated clear coupling between F47 and R55. It is apparent that these residues are not interacting in the NMR structure but appear to do so in the X-ray structure. On the basis of free energies calculated from MD simulations and obtained experimentally through double mutant cycles, the F47−R55 contact appears to be a stronger interaction than the proposed salt bridge between D44 and R55. Hence, the van der Waals interaction seen in the X-ray structure appears to play an important role in stabilizing the solution structure of HP36. The coupling free energy between the D44 side chain and the R55 side chain is small, only 0.29 ± 0.20 kcal/mol. The F47 side chain−R55 side chain coupling free energy is 0.70 ± 0.20 to 0.79 ± 0.20 kcal/mol. These results are consistent with the simulation results showing that the F47−R55 interaction is strong and that the stability of the D44−R55 pair is lower than the thermal energy.

Previous studies by Frank et al. (*8*) have shown the importance of each phenylalanine in stabilizing the core of the protein. Interestingly, the F47L/R55M double mutant is even less stable than these single Phe mutants, which suggests that the side chain of R55 also plays a key role in stabilizing the structure. It is likely that the optimal packing of the three phenylalanines in the core is enhanced by R55 because it helps to shield the core with its long side chain and also forms a cation−π interaction. Cation−π interactions can be important for folding and thermostability of various proteins and protein ligand systems (*53*−*55*). For the single mutant, R55M, the T_m dropped 6 °C in thermal stability, showing that more than just a bulky side chain is required at position 55. In the majority of villin sequences, Lys is found as a conservative mutation in place of R55 (*56*). This suggests that the charge is important for stabilizing the structure, but as the simulations and double mutant experiments indicate, the importance of this charge at position 55 does not arise from formation of an ion pair with D44 as observed in the crystal structure. It is worth noting, however, that in the X-ray structure D44 appears to be involved in a network of interactions, including a hydrogen bond to the backbone carbonyl L42. Backbone side chain interactions cannot be probed by double mutant cycle analysis.

These simulations are models, and as with any model, there are limitations, especially in the interpretation of results. Realistic, detailed simulations come at a high computational cost that must often be balanced against the need for obtaining extensive conformational sampling. Computational models continue to improve; the Amber and CHARMm force fields have been used extensively enough to identify weaknesses (*32*, *57*) such as overstabilization of secondary structure elements. In the simulations that we report here, this type of systematic error might contribute to the extension of α-helices that we observed, although we specifically addressed secondary structure bias in our development of the parameter set that was used for all of the simulations presented here (*32*). It has also been noted that the use of PME to calculate long-range electrostatics imposes long-range periodicity that that may result in artifacts from a crystal-like environment (*58*−*60*). In the case presented here, simulations with two alternate treatments of long-range interactions (including undesirable direct truncation) provided essentially the same conclusion, that the simulations adopt a structure in better agreement with the crystal structure than with the NMR structure. Thus, there is no evidence that these results are an artifact of PME.

Previous work by van der Spoel and Lindahl (*15*) reported a series of simulations of the NMR structure of HP36. These authors noted a modest degree of sensitivity to force field, water models, and protonation states. In their simulations, they noted larger structure fluctuations in the region connecting helices 1 and 2 as compared to the rest of the molecule. This observation is consistent with our results, which indicate this linker as one region in which the simulations spontaneously adopt a conformation more consistent with the crystal structure. At that time, there was no way for van der Spoel and Lindahl to determine whether this larger fluctuation resulted from a conversion toward the crystal structure which was reported 2 years later. Importantly, van der Spoel and Lindahl also noted the importance of taking into account the pH of the experiment when running simulations of HP36. Upon protonation of the glutamic acids side chains in the starting structure, the resulting simulation displayed a greater correlation to the chemical shift and *J*-coupling results which were originally measured at pH 3.7. This further suggests that one must be cautious in the quantitative comparison of simulations at neutral pH to experimental data obtained at low pH.

In summary, the results from our simulations and experiments show that the recently published X-ray structure is a more accurate representation of the structure in solution at neutral pH than the NMR structure at low pH. Importantly, the simulations also indicated that a salt bridge between R55 and D44 observed in the low-temperature crystal structure was thermally unstable, in contrast to the stable interaction between R55 and F47 in the simulation. Experimental double mutant analysis confirmed that the interaction free energy of the salt bridge was small and that the F47−R55 pair likely plays an important role in stabilizing the protein via a cation−π interaction. The analysis presented here shows how the combination of molecular dynamics simulations and experimental measurements can be used to develop a better understanding of the structural properties of proteins in solution.

ACKNOWLEDGMENT

We gratefully acknowledge computer time at NCSA through Grant MCA02N028 (C.S.). Additional computer time and assistance was provided by Roberto Gomperts and the SGI Engineering group. We thank C. James McKnight for helpful discussions.

SUPPORTING INFORMATION AVAILABLE

Summary of the results of the urea denaturation of the F47L mutant in various concentrations of TMAO (Table S1), the effect of varying ΔC_p° values on the calculated thermodynamic parameters of the F47L/R55M double mutant (Table S2), time evolution and distributions of the heavy atom backbone rmsds of the region containing helix 1 (residues S43−A49) (Figure S1A), time evolution and distribution of the heavy atom backbone rmsd of the region containing helix 2 (residues T54−A59) (Figure S1B), heavy atom backbone rmsd versus structure in the NMR family for the regions containing the whole backbone of the protein (residues L42−

L75), helix 1 (residues S43−A49), helix 2 (residues T54−A59), and helix 3 (residues P62−F76) (Figure S2), specific contact distances involving R55 and D44 and R55 and F47 for each structure in the NMR family (Figure S3), a time evolution and distributions of the heavy atom backbone rmsd for whole backbone of the protein (residues 43−75) during the simulation of the X-ray structure (Figure S4), time evolution and distributions of the heavy atom backbone rmsd for the whole backbone of the protein (residues 42−75) during the simulation of the NMR structure using a non-bonded cutoff of 12 Å (Figure S5A), and the time evolution and distributions of the heavy atom backbone rmsd for the whole backbone of the protein (residues 42−75) during the simulation of the NMR structure using an IPS electrostatic treatment (Figure S5B). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Brewer, S. H., Vu, D. M., Tang, Y., Li, Y., Franzen, S., Raleigh, D. P., and Dyer, R. B. (2005) Effect of modulating unfolded state structure on the folding kinetics of the villin headpiece subdomain, *Proc. Natl. Acad. Sci. U.S.A. 102*, 16662−16667.
2. Kubelka, J., Eaton, W. A., and Hofrichter, J. (2003) Experimental tests of villin subdomain folding simulations, *J. Mol. Biol. 329*, 625−630.
3. Buscaglia, M., Kubelka, J., Eaton, W. A., and Hofrichter, J. (2005) Determination of ultrafast protein folding rates from loop formation dynamics, *J. Mol. Biol. 347*, 657−664.
4. Wang, M., Tang, Y., Sato, S., Vugmeyster, L., McKnight, C. J., and Raleigh, D. P. (2003) Dynamic NMR line-shape analysis demonstrates that the villin headpiece subdomain folds on the microsecond time scale, *J. Am. Chem. Soc. 125*, 6032−6033.
5. McKnight, C. J., Doering, D. S., Matsudaira, P. T., and Kim, P. S. (1996) A thermostable 35-residue subdomain within villin headpiece, *J. Mol. Biol. 260*, 126−134.
6. McKnight, C. J., Matsudaira, P. T., and Kim, P. S. (1997) NMR structure of the 35-residue villin headpiece subdomain, *Nat. Struct. Biol. 4*, 180−184.
7. Chiu, T. K., Kubelka, J., Herbst-Irmer, R., Eaton, W. A., Hofrichter, J., and Davies, D. R. (2005) High-resolution X-ray crystal structures of the villin headpiece subdomain, an ultrafast folding protein, *Proc. Natl. Acad. Sci. U.S.A. 102*, 7517−7522.
8. Frank, B. S., Vardar, D., Buckley, D. A., and McKnight, C. J. (2002) The role of aromatic residues in the hydrophobic core of the villin headpiece subdomain, *Protein Sci. 11*, 680−687.
9. Vugmeyster, L., Trott, O., McKnight, C. J., Raleigh, D. P., and Palmer, A. G., III (2002) Temperature-dependent dynamics of the villin headpiece helical subdomain, an unusually small thermostable protein, *J. Mol. Biol. 320*, 841−854.
10. Tang, Y., Rigotti, D. J., Fairman, R., and Raleigh, D. P. (2004) Peptide models provide evidence for significant structure in the denatured state of a rapidly folding protein: The villin headpiece subdomain, *Biochemistry 43*, 3264−3272.
11. Havlin, R. H., and Tycko, R. (2005) Probing site-specific conformational distributions in protein folding with solid-state NMR, *Proc. Natl. Acad. Sci. U.S.A. 102*, 3284−3289.
12. Kubelka, J., Chiu, T. K., Davies, D. R., Eaton, W. A., and Hofrichter, J. (2006) Sub-microsecond protein folding, *J. Mol. Biol. 359*, 546−553.
13. Tang, Y., Goger, M. J., and Raleigh, D. P. (2006) NMR characterization of a peptide model provides evidence for significant structure in the unfolded state of the villin headpiece helical subdomain, *Biochemistry 45*, 6940−6946.
14. De Mori, G. M., Colombo, G., and Micheletti, C. (2005) Study of the villin headpiece folding dynamics by combining coarse-grained Monte Carlo evolution and all-atom molecular dynamics, *Proteins 58*, 459−471.
15. van der Spoel, D., and Lindahl, E. (2003) Brute-Force Molecular Dynamics Simulations of Villin Headpiece: Comparison with NMR Parameters, *J. Phys. Chem. B 107*, 11178−11187.
16. Duan, Y., and Kollman, P. A. (1998) Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution, *Science 282*, 740−744.
17. Duan, Y., Wang, L., and Kollman, P. A. (1998) The early stage of folding of villin headpiece subdomain observed in a 200-nanosecond fully solvated molecular dynamics simulation, *Proc. Natl. Acad. Sci. U.S.A. 95*, 9897−9902.
18. Shen, M. Y., and Freed, K. F. (2002) All-atom fast protein folding simulations: The villin headpiece, *Proteins 49*, 439−445.
19. Fernandez, A., Shen, M. Y., Colubri, A., Sosnick, T. R., Berry, R. S., and Freed, K. F. (2003) Large-scale context in protein folding: Villin headpiece, *Biochemistry 42*, 664−671.
20. Zagrovic, B., Snow, C. D., Shirts, M. R., and Pande, V. S. (2002) Simulation of folding of a small α-helical protein in atomistic detail using worldwide-distributed computing, *J. Mol. Biol. 323*, 927−937.
21. Sullivan, D. C., and Kuntz, I. D. (2001) Conformation spaces of proteins, *Proteins 42*, 495−511.
22. Sullivan, D. C., and Kuntz, I. D. (2002) Protein Folding as Biased Conformational Diffusion, *J. Phys. Chem. B 106*, 3255−3262.
23. Islam, S. A., Karplus, M., and Weaver, D. L. (2002) Application of the diffusion-collision model to the folding of three-helix bundle proteins, *J. Mol. Biol. 318*, 199−215.
24. Ripoll, D. R., Vila, J. A., and Scheraga, H. A. (2004) Folding of the villin headpiece subdomain from random structures. Analysis of the charge distribution as a function of pH, *J. Mol. Biol. 339*, 915−925.
25. Bandyopadhyay, S., Chakraborty, S., Balasubramanian, S., and Bagchi, B. (2005) Sensitivity of polar solvation dynamics to the secondary structures of aqueous proteins and the role of surface exposure of the probe, *J. Am. Chem. Soc. 127*, 4071−4075.
26. Herges, T., and Wenzel, W. (2005) Free-energy landscape of the villin headpiece in an all-atom force field, *Structure 13*, 661−668.
27. Jayachandran, G., Vishal, V., and Pande, V. S. (2006) Using massively parallel simulation and Markovian models to study protein folding: Examining the dynamics of the villin headpiece, *J. Chem. Phys. 124*, 164902.
28. Trebst, S., Troyer, M., and Hansmann, U. H. E. (2006) Optimized parallel tempering simulations of proteins, *J. Chem. Phys. 124*, 174903.
29. Zagrovic, B., and van Gunsteren, W. F. (2006) Comparing atomistic simulation data with the NMR experiment: How much can NOEs actually tell us? *Proteins 63*, 210−218.
30. Wickstrom, L., Okur, A., Song, K., Hornak, V., Raleigh, D. P., and Simmerling, C. L. (2006) The unfolded state of the villin headpiece helical subdomain: Computational studies of the role of locally stabilized structure, *J. Mol. Biol. 360*, 1094−1107.
31. Case, D., Cheatham, T. I. I. I., Darden, T., G., G., Luo, R., Merz, K. M., Jr., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) The Amber biomolecular simulation programs, *J. Comput. Chem. 26*, 1668−1688.
32. Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters, *Proteins 65*, 712−725.
33. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Jr., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules, *J. Am. Chem. Soc. 117*, 5179−5197.
34. Wang, J., Cieplak, P., and Kollman, P. A. (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem. 21*, 1049−1074.
35. Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes, *J. Comput. Phys. 23*, 327−341.
36. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984) Molecular dynamics with coupling to an external bath, *J. Chem. Phys. 81*, 3684−3690.
37. Shirts, M. R., and Pande, V. S. (2005) Solvation free energies of amino acid side chain analogs for common molecular mechanics water models, *J. Chem. Phys. 122*, 134508.
38. Still, W. C., Tempczyk, A., Hawley, R. C., and Hendrickson, T. (1990) Semianalytical treatment of solvation for molecular mechanics and dynamics, *J. Am. Chem. Soc. 112*, 6127−6129.
39. Zhou, R. (2003) Free energy landscape of protein folding in water: Explicit vs. implicit solvent, *Proteins: Struct., Funct., Genet. 53*, 148−161.

40. Pitera, J. W., and Swope, W. (2003) Understanding folding and design: Replica-exchange simulations of "Trp-cage'' miniproteins, *Proc. Natl. Acad. Sci. U.S.A. 100*, 7587−7592.
41. Zhou, R., and Berne, B. J. (2002) Can a continuum solvent model reproduce the free energy landscape of a β-hairpin folding in water? *Proc. Natl. Acad. Sci. U.S.A. 99*, 12777−12782.
42. Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An N.log(N) method for Ewald sums in large systems, *J. Chem. Phys. 98*, 10089−10092.
43. Wu, X. W., and Brooks, B. R. (2005) Isotropic periodic sum: A method for the calculation of long-range interactions, *J. Chem. Phys. 122* 044107.
44. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys. 79*, 926−935.
45. Bi, Y., Tang, Y., Raleigh, D. P., and Cho, J. H. (2006) Efficient high level expression of peptides and proteins as fusion proteins with the N-terminal domain of L9: Application to the villin headpiece helical subdomain, *Protein Expression Purif. 47*, 234−240.
46. Simmerling, C., Elber, R., and Zhang, J. (1995) MOIL-View: A Program for Visualization of Structure and Dynamics of Biomolecules and STO: A Program for Computing Stochastic Paths, in *Modeling of Biomolecular Structure and Mechanism* (Pullman, A., Ed.) pp 241−265, Kluwer, Dordrecht, The Netherlands.
47. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein, *Protein Sci. 4*, 2411−2423.
48. Wang, A., and Bolen, D. W. (1997) A naturally occurring protective system in urea-rich cells: Mechanism of osmolyte protection of proteins against urea denaturation, *Biochemistry 36*, 9101−9108.
49. Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features, *Biopolymers 22*, 2577−2637.
50. Mello, C. C., and Barrick, D. (2003) Measuring the stability of partly folded proteins using TMAO, *Protein Sci. 12*, 1522−1529.
51. Fersht, A. (1999) *Structure and mechanism in protein science: A guide to enzyme catalysis and protein folding*, W. H. Freeman and Co., New York.
52. Spector, S., Young, P., and Raleigh, D. P. (1999) Nativelike structure and stability in a truncation mutant of a protein minidomain: The peripheral subunit-binding domain, *Biochemistry 38*, 4128−4136.
53. Gallivan, J. P., and Dougherty, D. A. (1999) Cation−π interactions in structural biology, *Proc. Natl. Acad. Sci. U.S.A. 96*, 9459−9464.
54. Shi, Z., Olson, C. A., Bell, A. J. J., and Kallenbach, N. R. (2001) Stabilization of α-helix structure by polar side-chain interactions: Complex salt bridges, cation−π interactions, and C-H···O H-bonds, *Biopolymers 60*, 366−380.
55. Dougherty, D. A. (1996) Cation−π interactions in chemistry and biology: A new view of benzene, Phe, Tyr and Trp, *Science 271*, 163−168.
56. Vermeulen, W., Vanhaesebrouck, P., Van Troys, M., Verschueren, M., Fant, F., Goethals, M., Ampe, C., Martins, J. C., and Borremans, F. A. M. (2004) Solution structures of the C-terminal headpiece subdomains of human villin and advillin, evaluation of headpiece F-actin-binding requirements, *Protein Sci. 13*, 1276−1287.
57. Okur, A., Strockbine, B., Hornak, V., and Simmerling, C. (2003) Using PC clusters to evaluate the transferability of molecular mechanics force fields for proteins, *J. Comput. Chem. 24*, 21−31.
58. Hunenberger, P. H., and McCammon, J. A. (1999) Effect of artificial periodicity in simulations of biomolecules under Ewald boundary conditions: A continuum electrostatics study, *Biophys. Chem. 78*, 69−88.
59. Hunenberger, P. H., and McCammon, J. A. (1999) Ewald artifacts in computer simulations of ionic solvation and ion-ion interaction: A continuum electrostatics study, *J. Chem. Phys. 110*, 1856−1872.
60. Weber, W., Hunenberger, P. H., and McCammon, J. A. (2000) Molecular dynamics simulations of a polyalanine octapeptide under Ewald boundary conditions: Influence of artificial periodicity on peptide conformation, *J. Phys. Chem. B 104*, 3668−3675.

BI061785+