THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Mirror Images as Naturally Competing Conformations in Protein Folding

Jeffrey K. Noel,[‡] Alexander Schug,[§] Abhinav Verma,[§] Wolfgang Wenzel,[⊥] Angel E. Garcia,*[,∥] and José N. Onuchic*[,‡]

[‡]Center for Theoretical Biological Physics and Department of Physics, Rice University, Houston, Texas 77005, United States

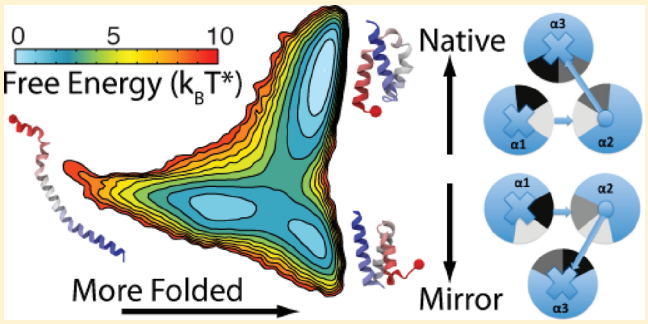
[§]Steinbuch Centre for Computing, Karlsruhe Institute of Technology, Karlsruhe, Germany

[∥]Center for Biotechnology and Interdisciplinary Studies and Department of Physics, Applied Physics, and Astronomy, Rensselaer Polytechnic Institute, Troy, New York 12180, United States

[⊥]Institute of Nanotechnology, Karlsruhe Institute of Technology, Karlsruhe, Germany

Supporting Information

ABSTRACT: Evolution has selected a protein's sequence to be consistent with the native state geometry, as this configuration must be both thermodynamically stable and kinetically accessible to prevent misfolding and loss of function. In simple protein geometries, such as coiled-coil helical bundles, symmetry produces a competing, globally different, near mirror image with identical secondary structure and similar native contact interactions. Experimental techniques such as circular dichroism, which rely on probing secondary structure content, cannot readily distinguish these folds. Here, we want to clarify whether the native fold and mirror image are energetically competitive by investigating the free energy landscape of three-helix bundles. To prevent a bias from a specific computational approach, the present study employs the structure prediction forcefield PFF01/02, explicit solvent replica exchange molecular dynamics (REMD) with the Amber94 forcefield, and structure-based simulations based on energy landscape theory. We observe that the native fold and its mirror image have a similar enthalpic stability and are thermodynamically competitive. There is evidence that the mirror fold has faster folding kinetics and could function as a kinetic trap. All together, our simulations suggest that mirror images might not just be a computational annoyance but are competing folds that might switch depending on environmental conditions or functional considerations.



1. INTRODUCTION

Much of protein folding theory works under the assumption that the energy landscape is biased toward a single attractive basin: the native state. Anfinsen's thermodynamic hypothesis states that a protein's native conformation lies in the global minimum of its free-energy landscape.[1] Evolution achieves this robustness by selecting for sequences in which the interactions present in the native state are mutually supportive and cooperatively lead to the functional structure.[2] This gives rise to protein sequences that are minimally frustrated, meaning sequences are only consistent with a single native structure. The resulting energy landscape is smooth and funnel-shaped. Any competing or "trapping" protein configurations have an energetic depth much smaller than the overall bias to the native structure. Under this framework, which has been called *energy landscape theory*,[3,4,2] protein dynamics is dominated by the geometry of the protein's native configuration. In this scenario, we pose the question: What are the consequences of native configurations with a high degree of symmetry?

Prior studies have investigated the detailed folding mechanisms for several proteins with native structures that have specific domains arranged in a symmetric pattern.[5] For example, protein L consists of an α-helix packed against two, symmetrically arranged, β-hairpins. Protein L, though, folds asymmetrically, through a transition state ensemble (TSE) consisting of an ordered N-terminal β-hairpin and largely unstructured C-terminal β-hairpin.[6] A homologous protein, protein G, instead folds by ordering the C-terminal β-hairpin in the TSE.[7] Simulations have suggested that the detailed side-chain packing determines one folding route over the other.[8,9] Another fold family consisting of symmetric domains are the β-trefoils.[10] A series of theoretical and experimental studies of the Interleukin-1 family of β-trefoils has shown that the folding degeneracy is broken by functional regions of the protein,

Special Issue: B: Harold A. Scheraga Festschrift

Received: December 31, 2011
Revised: April 10, 2012
Published: April 12, 2012

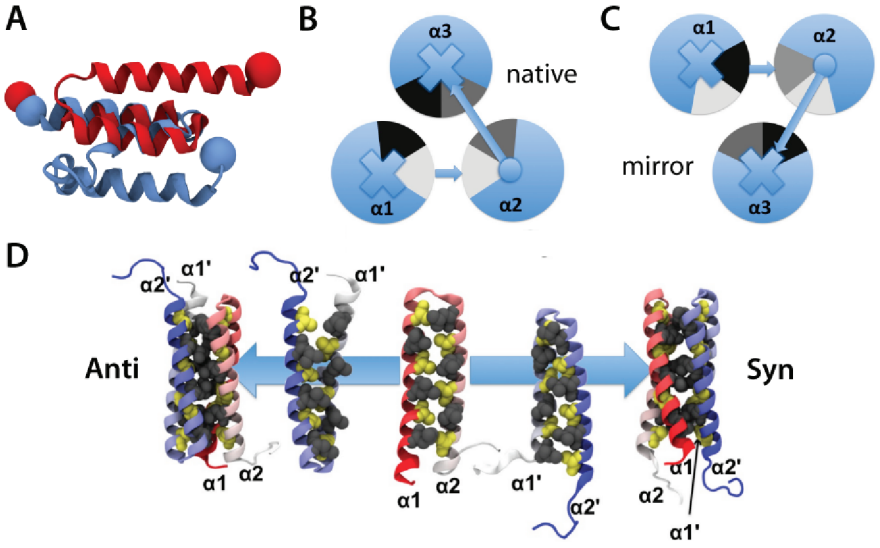


Figure 1. Protein symmetry gives rise to multiple consistent structures. (A) Mirror images of Bdpa, a three-helix bundle. The C-terminal helices are aligned, and the native N-terminal helix packs on the top, while the mirror N-terminal helix packs on the bottom. (B,C) The packing of amphipathic helices; the hydrophobic core residues are shown in gray. Compared to the native fold (B), a slight rotation of helices 1 and 2 and reorientation of the third helix facilitate a hydrophobic core composed of the same residues in the mirror fold (C). The packing of interacting residues, however is different. (D) The Rop-homodimer is composed by two all-helical monomers (chain A red to white, chain B white to blue). If one mutates core residues at the hydrophobic interface into optimally packed Ala (yellow) and Leu (gray), there is competition between two possible arrangements of the monomers (displayed mutant A2L2-6).[24] One possibility is the functional wild-type anti-packing in which six small Ala side chains perfectly fit into six larger Leu side chains. For anti, helix $\alpha1$ is packed against $\alpha1'$ and $\alpha2$ against $\alpha2'$. Symmetry, however, enables the competition of a dysfunctional syn packing, in which Ala and Leu form a similarly perfect packing. Here helix $\alpha1$ is packed against $\alpha2'$ and $\alpha2$ against $\alpha1'$. Both configurations have been measured to coexist in single-molecule fluorescence resonance energy transfer (FRET) experiments.[25]

which slightly alter the structure among each member.[11−13] The overall lesson from these studies is that, while the proteins are able to fold via multiple routes, they tend to choose one of the routes allowed by symmetry to dominate the TSE. The choice is based on residual frustration that can arise, for example, from the geometry of the side-chain packing or energetic heterogeneity. These differences are subtle, and not robust, as a couple kilocalories per mole is enough to bias the folding down a particular route.[14,2]

Let us now consider protein structures where symmetry allows, in addition to multiple folding routes, multiple structures within a nominally singly funnelled energy landscape. For example, for a three-helix bundle like the B-domain of protein A (Bdpa), the near mirror image of the native state is compatible with the native contact network (Figure 1). [Technically, the chirality of naturally occurring L-amino acids precludes folding of a "mathematically" exact mirror image. For the sake of simplicity, we call the mirror arrangement of helix axes a mirror image.] The three helices pack around a hydrophobic core and arrange (left-) right-handed for the (native) mirror fold. Nonspecific interactions, such as hydrophobic interactions, leave little to distinguish between the native and mirror helical packings. This missing specificity toward one particular conformation has been seen in structure prediction, where energy functions are unable, or only weakly able, to discriminate between the native fold and competing mirrored folds/decoys with low energies.[15−19] For all-helical coiled-coil proteins, the composing helices form a hydrophobic core with, and orient against, each other to form the native fold, but a simple reorientation in their mirror image forms a comparable hydrophobic interface.

The ability of proteins to explore multiple structures allowed by symmetry is seen in domain-swapped homodimers.[20] Domain swapping occurs when structural elements crucial for stabilizing the monomeric fold are replaced by the same structural elements from its dimeric partner. As expected in a funneled energy landscape, simulations have shown that the signals for domain swapping are encoded by the monomeric fold,[21] i.e., they arise solely from symmetry. The low specificity toward a singular fold is, perhaps, best exemplified by the Rop-homodimer and its various mutants. While the wild-type (un)folds slowly (k_F = 0.013 s^{-1}), specific mutants have an optimized Ala/Leu core-packing and speed up folding by up to 4 orders of magnitude.[22] At the same time, these mutations symmetrize the interface,[23] reduce the specificity to the native fold, and open what has been called a "trapdoor",[24] an energetically competitive structure with a different symmetrically related global fold (Figure 1). Further, this "trapdoor-fold" is stabilized by small concentrations of denaturant, which explains the unusual kinetic behavior.[25]

The present study investigates the free-energy landscape of three-helix bundle proteins that may have naturally occurring trapdoors built in by symmetry. Their simple coiled-coil structures allow the mirror image structures to be energetically competitive. The B/E-domains of staphylococcal protein A (Bdpa and Edpa) are two of five homologous IgG-binding domains.[26] They act as pathogenicity factors for the bacterium *Staphylococcus aureus*, by binding tightly to the Fc region of IgG or Fab region of IgM, and each consist of three helices packed against one another. α_3d is a *de novo* superstable designed three-helix bundle that expresses very quick kinetics with folding times around 4 μs.[27]

2. METHODS

This study utilizes several simulation methods, not only to guard against forcefield bias, but also in an attempt to gain a

dx.doi.org/10.1021/jp212623d | J. Phys. Chem. B 2012, 116, 6880−6888

complete perspective on the energetic, kinetic, and thermodynamic accessibilities of mirror-symmetrical protein structures. The structure prediction forcefield PFF01/02 shows that mirror configurations of Bdpa, Edpa, and α_3d are enthalpically competitive with the native configurations. This competition between mirrored helical bundles is explored in detail for Bdpa with replica exchange molecular dynamics (REMD) and structure-based simulations. REMD simulations reveal two thermodynamically competitive folded basins: an enthalpically favored native-like basin and an entropically favored mirror-like basin. The structure-based simulations are built from representative REMD structures and corroborate the REMD predictions for the kinetic accessibility of the two basins.

2.1. Structure Prediction. *2.1.1. Forcefield: PFF01/02.* The all-atom (with the exception of apolar CH_N groups) free-energy forcefield PFF01/02[28,19,29] models the internal free-energy of the protein along with an averaged implicit solvent interaction. Contributions from backbone entropy are not considered. It has found wide application in structure prediction.[30−32] The energies can be used to reconstruct folding kinetics. Its functional form is

$$V(\vec{r}) = \sum_{ij} V_{ij}\left[\left(\frac{R_{ij}}{r_{ij}}\right)^{12} - 2\left(\frac{R_{ij}}{r_{ij}}\right)^{6}\right] + \sum_{ij} \frac{q_i q_j}{\varepsilon_{g(i)g(j)} r_{ij}} + \sum_i \sigma_i A_i + \sum_{\text{hbonds}} V_{\text{hb}}$$

where r_{ij} denotes the distance between atoms i and j, and $g(i)$ is the type of amino acid containing the atom i. The Lennard-Jones parameters (V_{ij} for potential depth and R_{ij} for equilibrium distance) depend on the type of the atom pair and were adjusted to satisfy constraints from a set of 138 proteins out of the Protein Data Bank (PDB) database. The electrostatic interactions contain group-specific dielectric constants $\varepsilon_{g(i)g(j)}$, and partial charges q_i were derived in a potential of mean-force approach. The implicit solvent interaction constants are obtained by a minimal solvent accessible approach and parametrized by free energies per unit area σ_i to reproduce the solvation enthalpies of the Gly-X-Gly family of peptides. A_i corresponds to the area of atom i that is in contact with a fictitious solvent. Hydrogen bonding is modeled as dipole−dipole interactions in the electrostatic term and an additional short-range term for backbone hydrogen bonding (CO to NH). It depends on the OH distance, the angle between N, H, and O atoms along the bond, and the angle between the CO and NH axis.

2.1.2. Global Optimization. A variety of approaches have been tested to find the global minimum of PFF01/02.[28,33] For the present study we used an evolutionary algorithm for the proteins α_3d and Edpa.[34] Starting from random initial conformations, a population of structures is fully minimized by repeated rounds of section and minimization. Each round selects a subset of structures that balances energetic favorability with structural diversity. This approach can easily be implemented on heterogeneous computational resources and is very efficient.

2.2. REMD Simulations. We use REMD to study the folding/unfolding equilibrium of the *Staphylococcus aureus* Bdpa using an all-atom and explicit solvent, forcefield-based model. Bdpa consists of amino acids 1−57 (TADNKFNKEQ QNAFYEILHL PNLNEEQRNG FIQSLKDDPS QSANL-LAEAK KLNDAQA) and was modeled with charged C- and N-termini. His 19 was charged. All Glu, Asp, Lys, and Arg were modeled as charged. The initial configuration of the system was modeled as an extended PPII structure, without any bias toward the native state. The protein was modeled by the all-atom Amber ff94 forcefield, and the solvent was modeled by 6583 TIP3P water molecules, as in our previous calculations on this protein.[35] An extended, PPII conformation was generated with the Amber leap program. To remove any bias from this initial configuration, the system was heated to 800 K and simulated in vacuum for 100 ps. The resulting collapsed and unfolded configuration was immersed in a cubic box of water molecules and equilibrated at 300 K and 1 atm for 1 ns. The equilibrated protein−water system was a cube 5.97 nm on a side and has a density of 0.9699 g/L. Copies of this system were then simulated for 1 ns at various T ranging from 275 to 600 K at constant volume. These configurations were used as initial configurations in the REMD calculation. The REMD simulations were extended for 450 ns per replica. The total simulation time was 28.8 μs. REMD simulations were done over a wide range of temperatures, 287−643 K, chosen to get a uniform exchange rate among replicas sampling neighboring temperatures.[36] Exchange attempts were possible at every integration step with a 5% percent probability. The exchange rate was chosen to be close to 20%. On average, exchanges were attempted every 1.4 ps. The average time between successful exchanges for all replicas was 8 ps. The integration time step was 2 fs. Temperatures were maintained using the Nose−Hoover thermostat, with 2 ps coupling time. Hydrogen containing bonds were constrained by SHAKE and SETTLE. We used the particle mesh Ewald summation with a 0.1 nm grid size. Lennard-Jones interactions were truncated at 1.0 nm. Pair interaction lists were updated every 25 integration steps. Figure S1 (Supporting Information) shows the convergence of the REMD simulation. Thermodynamic averages were calculated over the last 250 ns/replica.

In our experience, Amber ff99SB, which has been shown to describe better dynamics[37] and thermodynamics,[38] does not fold Bdpa nor does it maintain the folded state in REMD simulations (40 μs total) started from the folded state.

2.3. Structure-Based Models (SBMs). SBMs define a target structure (with low free energy like protein native structures) as the explicit energetic minimum.[39] We used a standard coarse-grained, native-centric SBM, commonly referred to as a "C_α-model". We used the potential described in ref 40, where the usual Lennard-Jones contact potentials were replaced by Gaussian contact potentials. The width of the Gaussians scales linearly with the native contact distance and resulted in an average width of $\overline{\sigma} \sim 0.7$ Å. The all-atom SBM is described here,[41] except that we replace the Lennard-Jones contact potentials with Gaussian contact potentials as described in refs 42 and 43, and the excluded volume parameters were $\varepsilon_{NC} = 1$ and $\sigma_{NC} = 2.1$ Å, giving more realistic atomic sizes. The native contact maps were constructed using Shadow.[42,43] The SBMs were constructed using the SMOG webserver.[43] All SBMs were sampled with under-damped stochastic dynamics using Gromacs v4.5.[44]

In order to compare the kinetics of the symmetrical structures, we constructed an SBM with "dual-basins" that has equal energetic minima at two structures. This type of SBM has also been used to look at, for example, conformational transitions[45] and homodimeric folding.[24,46] The two "native" structures, used as input for the dual-basins, were taken as representative members from REMD clusters 1 and 3 from

dx.doi.org/10.1021/jp212623d | J. Phys. Chem. B 2012, 116, 6880−6888

REMD. The structures with the highest number of atomic contacts as determined by Shadow were chosen and minimized in Amber94. The native-like structure thus obtained, called S_N, has 444 atomic contacts, whereas the mirror-like structure S_M has 373 atomic contacts. They are shown in Figure 1A. In the dual-basin SBM potential, all contacts were included. If a contact exists in both structures, but at different distances, it was included using a double-basin Gaussian contact potential, which consists of two Gaussian wells of equal depth with minima at the respective native distances.[40,42] To make the basins equally energetically stable, the contact potentials in S_M were scaled by 444/373 ≈ 1.2. The torsional angles have a form such that they are minimized at both S_N and S_M. Angles and bond lengths, nearly identical between S_N and S_M, were taken from S_N.

2.4. Contact Map Definitions. In order to quantify a structure's similarity to S_N and S_M, we needed to construct two "native" contact maps C_N and C_M. A contact between two residues was considered "native" to its basin X if, in a subset of structures <3 Å away from S_X, a contact as defined by the Shadow algorithm[43,42] existed with a probability greater than 0.5. The Shadow algorithm defines residues in contact if they have any directly interacting atoms, i.e., two atoms within 6 Å and not occluded by other atoms. A tertiary contact is any contact between residues separated by more than 5 amino acids in sequence. Applying this definition of the contact map, the native basin contact map C_N has 112 total contacts and 65 tertiary contacts, and the mirror basin contact map C_M has 101 total contacts and 51 tertiary contacts. The reaction coordinate $Q_X = \sum_{ij}\theta(1.2r_{ij}^X/r_{ij})$ measures the similarity to S_X, where the sum goes over tertiary residue pairs ij in C_X, r_{ij} is the distance between C_α atoms ij, r_{ij}^X is the distance between C_α atoms ij in S_X, and θ is the unit step function. In order to separate the native basin from the mirror basin, we used the combined coordinate $Q_N - Q_M$. Q_X^{AA} is a finely grained reaction coordinate that sums over all tertiary atomic pairs from an all-atom contact map C_X^{AA} constructed by running Shadow on S_X. Q_X^{AA} is used with the all-atom SBM. C_N^{AA} has 444 atomic contacts (254 tertiary) and C_M^{AA} has 373 atomic contacts (194 tertiary).

3. RESULTS

3.1. Coarse-Grained Native-Centric Protein Model Populates the Mirror Basin. The near degeneracy of the native and mirror structures is seen through simulating a coarse-grained SBM (C_α-model)[47,40] of the three-helix bundles. The model includes short-range interactions between native contacts and torsional angles, and all interactions have their minima at the native structure. Although the contact interactions on their own are unable to discern between the native structure and its mathematical mirror image because they only depend on scalar distances, the torsional angles are vector quantities, and therefore bias the local chirality to that of the native structure. Since all the helices in these proteins are right-handed, this results in a large energetic penalty for the left-handed helices that would arise in a mathematical mirror image. In the simulations, the helices are maintained with right-handed chirality, but the overall packing of the helices switches between the native packing and a mirror image packing. It is this change in helical packing that defines our use of the term *mirror* (Figure 1).

Figure 2 shows the relative populations of the native and the mirror packings. The native packing is monitored by the root

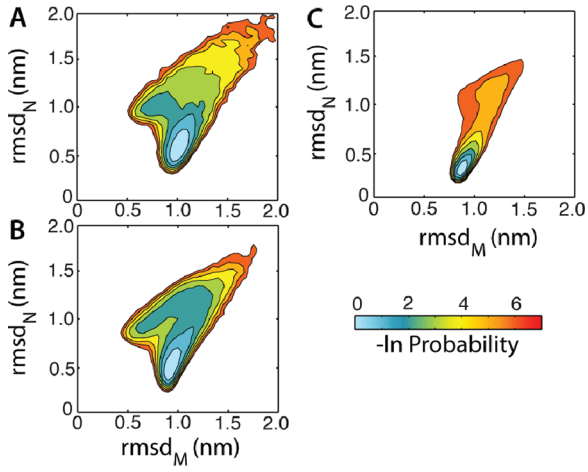


Figure 2. Native-centric, coarse-grained protein model shows the geometric similarity between the mirrored structures: rmsd to the mirror structure (rmsd$_M$) versus rmsd to the native structure (rmsd$_N$) for Bdpa (A), α_3D (B), and CI2 (C). Histograms are shown at $T = 0.93T_F$. In A and B, the mirror image shows up as an accessible conformation for both three-helix bundles. Roughly 5% of folded structures are in the mirror configuration. CI2 has an α/β-fold with no symmetrical analogue and therefore only populates a single basin.

mean squared distance (rmsd) of the C_α atoms from the native structure rmsd$_N$. The mirror packing is monitored by rmsd$_M$, the rmsd from a mirror packing of right-handed helices S_M taken from REMD (see Section 2.3). The three-helix bundles, Bdpa and α_3d, both show a significant population of the mirror conformation. Roughly 5% of folded structures are in the mirror configuration (mirror basin is destabilized by $\sim 2k_BT_F$ relative to native). For comparison, CI2, a well-studied α/β two-state protein, is shown in Figure 2C. CI2 only populates the native state.

The native and mirror helical packings are not completely degenerate because the helices maintain their right-handed chirality. The structures populating the mirror basin are balancing the energetic cost of straining the native torsional angles to the energetic benefit of forming native tertiary contacts whose distances are not optimized to the mirror helical packing. The ability of the protein to find such structures at a low enough energy is a testament to the plasticity of simple protein structures. While it can be said that the stability of the mirror is overemphasized because the changes in side-chain packing are not described in the coarse-grained model, the energetic heterogeneity is underrepresented by the native-centric energy function. The tertiary contact network is largely hydrophobic, constructed through the packing of amphipathic helices. The nonspecific nature of hydrophobic interactions should allow the mirror packing to utilize additional tertiary contacts beyond the native set. These non-native contacts are not included in the coarse-grained model. The hydrophobic side-chain packing is explored in the next section through a structure prediction forcefield with no knowledge of the native structure.

3.2. Mirrored Structures Are Energetically Competitive. To assess the stability of native and mirror configurations, we performed structure prediction simulations of Edpa in the all-atom, implicit water forcefield PFF01/02[29] using an evolutionary algorithm[34] to minimize the energy and sample the low energy landscape of the protein. Over 5000 protein conformations were visited during the simulations, and the

dx.doi.org/10.1021/jp212623d | J. Phys. Chem. B 2012, 116, 6880−6888

energy and $rmsd_N$ (from the PDB structure) of the final population of conformations are shown in Figure 3. Edpa

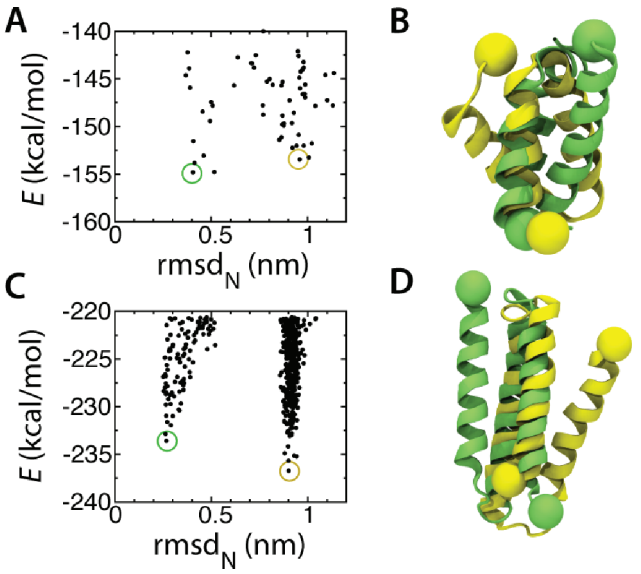


Figure 3. Distribution of structure prediction energies versus rmsd from the PDB structures (PDB codes: (A) 1edk and (B) 2a3d). Edpa (A and B) and α_3d (C and D). The lowest energy native-like structure is shown in green, and the lowest energy mirror-like structure is shown in yellow. The N-terminal helices are aligned. The energies of the structures shown in B and D are indicated by green and yellow circles in A and C. Two populations with low energy are seen for both proteins. Edpa has more structural diversity than α_3d, probably due to its smaller size.

shows two broad funnels at around 4 Å and 10 Å $rmsd_N$. Although the native-like conformation in PFF01/02 is lower in energy than the competing misfolded conformation (~1 kcal/mol), such a small difference is within the resolution attainable by empirical forcefields. The misfolded funnel ($rmsd_N$ ~ 10 Å) consists of mirrored configurations of Edpa with the N-terminal helix flipping to the other side of the structure formed by middle and C-terminal helix. The native-like and mirrored conformations are shown in Figure 3. In these simulations, we find no specific rearrangement of the side-chain packing of helices 2 and 3 to accommodate the change in orientation of helix 1 with respect to the native state.

Similarly, we performed simulations for the α_3d protein in PFF01/02. Here we also observe a double funnel in the energy landscape as shown by the distribution of all conformations visited during the simulation (Figure 3C). The native-like funnel found its minimum at ~2.5 Å, while the mirror-like funnel had a minimum at ~9 Å. The conformation at the minimum of the mirrored funnel had the same arrangement of helix 1 and 2 as native-like, while the C-terminal helix is on the other side with comparable side-chain packing. Note that the chirality of helical packing in α_3d is opposite that of Bdpa/Edpa.

The consistent observation of mirror images in three-helix bundles in PFF01/02 (Bdpa data with similar results as Edpa can be found in Figure 5 of ref 19) and by others[15−18,48,49] leads us to speculate that the mirror configurations may affect the folding and function of these proteins. In the next section, we examine the thermodynamic competitiveness of the mirror configuration on the folding landscape.

3.3. Atomistic Simulations Show a Mirror Basin That Is Thermodynamically Competitive and Kinetically Accessible. In order to predict the occupation of the mirror structure, for example, in an in vitro protein folding experiment, we need to know more than the relative energetic stabilities of the two competing structures. The mirror structure needs to both be thermodynamically competitive and kinetically accessible in order for the mirror packing to be visited. To investigate the thermodynamic behavior of symmetrical proteins, we performed REMD of Bdpa in an explicit solvent environment. Both helical packings are observed, at a native/mirror ratio of ~3. We now present results from both the REMD and various all-atom structure-based simulations to show the kinetic accessibility of the native and mirror helical packings.

3.3.1. REMD of Bdpa. A total of 28.8 microseconds of REMD was performed, 450 ns for each of 64 replicas of Bdpa. The simulation representation of Bdpa used the same sequence as in the experimental studies of Sato et al.[50] Blind cluster analysis of the resulting structures shows that, at T = 287 K, the most populated cluster (cluster 1) corresponds to the native structure, and the fifth most populated cluster (cluster 3) corresponds to a mirror helical packing (Figure 4). The best

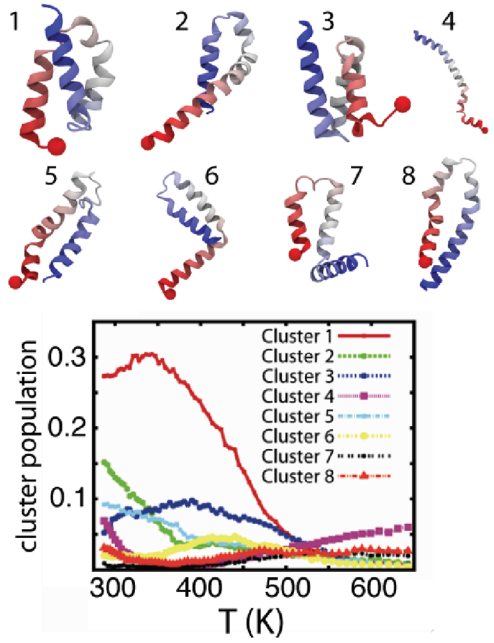


Figure 4. Top: Central structures from the largest eight structural clusters in the REMD Bdpa simulations at 287 K. Cluster 1 is closest to the NMR structure (PDB code: 1bdd[51]), and cluster 3 is the mirror image. Bottom: Cluster population as a function of temperature. The native-like cluster is the most populated until 500 K.

agreement between the deposited NMR structure for the protein[51] and the simulation is 1.2 Å rmsd of the backbone. From 350 K to 525 K, the native and mirror clusters become respectively the first and second most populated clusters. The relative thermodynamic stabilization of the mirror structure with temperature is reminiscent of the Rop dimer (Figure 1) and suggests that the mirror structure is more entropically favorable.

The REMD simulations provide details of the equilibrium populations of the symmetric folds, but they do not provide a direct description of the kinetics of the process. Analysis of

dx.doi.org/10.1021/jp212623d | J. Phys. Chem. B 2012, 116, 6880−6888

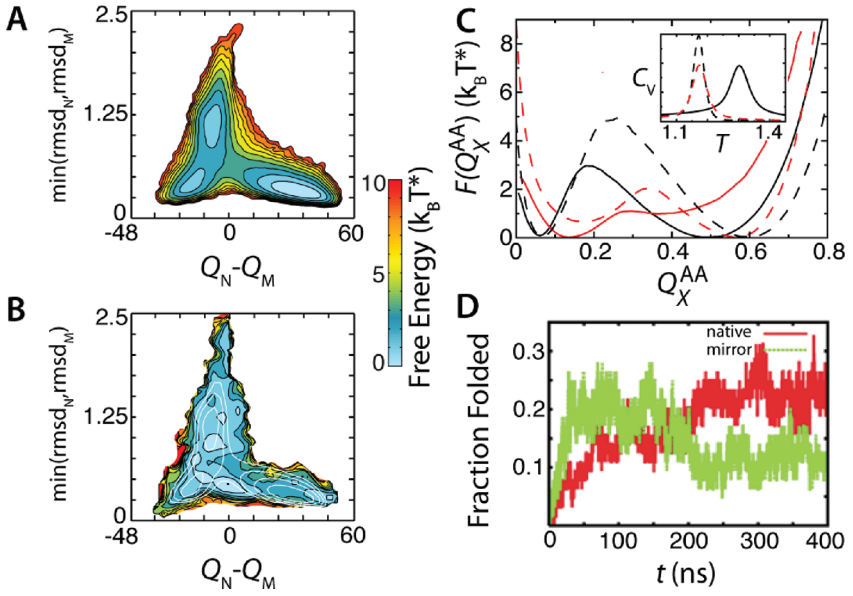


Figure 5. Kinetic accessibility and thermodynamic stability of the mirror image. Free energy contours are shown for two models: MD simulations of a dual-basin all-atom SBM at $\overline{T}$ = 1.29 (A), and REMD of Amber94 at $\overline{T}$ = 0.6 or T = 300K (B). The reduced temperature $\overline{T} = T/T^*$, where $k_BT^* = 1$. The minimum rmsd to either the native or mirror structure, min(rmsd$_N$, rmsd$_M$), is plotted versus the difference in native contacts from the mirror contacts, $Q_N - Q_M$. Therefore, the abscissa partitions structures between mirror-like and native-like and the ordinate partitions helical-bundle-like structures from extended ones. The SBM basins ($P > 10^{-3}$) are overlaid on the Amber94 as the white lines in B. (C) Comparison of barrier heights between three simulations: two control single-basin (SB) SBMs (dotted lines) at $\overline{T}$ = 1.17 and the dual-basin (DB) SBM (solid lines) at $\overline{T}$ = 1.29. Free energy barriers computed with similarity to the native structure Q_N^{AA} are shown in black, and those to the mirror structure Q_M^{AA} are shown in red. In both the SB and DB simulations, the mirror structure has a lower barrier compared to the native. The mirror structure is more stable than the native structure in SB, but it is less stable in DB. Both barriers in the DB simulations are lower than those in the SB simulations. The inset in C shows the specific heat for the three calculations. (D) Corroborating evidence from REMD that the mirror structure is more kinetically accessible than the native structure. Fraction of structures across all 64 replicas belonging to the folded clusters are plotted as a function of simulation time. The native (cluster 1) is shown in red and the mirror (cluster 3) is shown in green. If t = 400 ns corresponds to equilibrium, the occupation of the native relative to the mirror would be ~2 in the ensemble of all replicas. The free energies in A, B, and C were computed using WHAM.[59].

individual replica trajectories, even though it does not correspond to a single temperature, can give hints of the kinetic properties. The REMD was initialized from random collapsed, but unfolded, conformations. Monitoring the number of replicas that sample the native or mirror structural clusters as a function of time shows that the mirror cluster is initially populated roughly twice as fast as the native cluster (Figure 5D).

3.3.2. Dual-Funneled Energy Landscape Description of Bdpa. A funneled energy landscape has a single structure (or rather a small ensemble) that is both the enthalpy minimum and (below the folding temperature) the free energy minimum of the energy landscape. "Enthalpy" refers to the renormalized enthalpy obtained by averaging over solvent contributions. Here, for three-helix bundles, our evidence suggests an energy landscape with *two* structures near both the enthalpy minimum (as seen by PFF01/02 in Section 3.2) and the free energy minimum (as seen by REMD in the previous section), i.e., the native and the mirror helical packings. Therefore, a first-order approximation to the funneled energy landscape is a SBM that includes "dual-basins", where both the mirror and native structures are made explicit energetic minima.

A coarse-grained dual-basin SBM has been previously applied to a Rop dimer mutant containing homogeneous hydrophobic core packing.[24,40] Here, where the side-chain packing is likely of great importance, we use an all-atom SBM that explicitly represents all heavy atoms in the protein. The two "native" structures are representative members of the native and mirror clusters, called S_N and S_M, taken from the REMD. The dual-basin SBM is set up such that the two structures are equally energetically stable. Thus, the SBM only tests the relative entropy and kinetic accessibility of the two structures. Constant-temperature MD simulations were performed, and the results are presented in Figure 5. Comparing Figure 4A,B shows that, while the dual-basin captures the essence of the REMD landscape, even neglecting the spurious high rmsd structures of the REMD, the SBM misses some of the structural heterogeneity seen in REMD.

Figure 5C presents the thermodynamics of three SBMs: two single-basin SBMs to each S_N and S_M, and a dual-basin SBM. The two single-basin models indeed have nearly equal thermal stability as seen by the specific heat, and the dual-basin model is considerably more stable. This is expected, because the dual-basin SBM can form contacts from both S_N and S_M, and the dual-basin torsional angles are more forgiving. This also decreases the cooperativity of the dual-basin SBM (increases width of C_V) because of the additional structural heterogeneity. The free energy is plotted versus the number of native atomic contacts formed, Q_N^{AA} and Q_M^{AA} defined by S_N and S_M, respectively. The heights of the free energy barriers correspond to the kinetics of the transitions.[52] The single-basin SBMs predict much slower kinetics for the native S_N compared to the mirror S_M. The barriers to both the native basin and the mirror basin decrease in the dual-basin SBM, but the barrier to the native is still 2 k_BT^* larger than to the mirror. The relative thermodynamic stabilities between the native and mirror, favors the mirror structure for the single-basin SBM, but switches to favoring the native structure in the dual-basin SBM by 1 k_BT^*.

dx.doi.org/10.1021/jp212623d | J. Phys. Chem. B 2012, 116, 6880–6888

The dual-basin SBM predicts the following quantities for Bdpa: with the diffusion taken as constant, the folding rate to the mirror is $\exp(-\Delta F^{\ddagger}/k_B T) = \exp(1.8/\overline{T}) = 4.0$ times faster, and the native structure is more stable by a factor of $\exp(\Delta F^{\dagger}/k_B T^*) = \exp(1) = 2.7$, where $F^{\dagger}$ and $F^{\ddagger}$ are the free energies at the barrier and the folded basin, respectively, and Δ implies a subtraction of native from mirror.

The entropic favorability of the mirror basin is suggested by the REMD, since the population of the mirror cluster increases with temperature. This is corroborated by the SBM in two ways. First, in the single-basin SBMs, where S_N and S_M are given equal energetic stability, the mirror is slightly more stable than the mirror (Figure 5C). Second, as temperature is increased in the dual-basin SBM, the mirror basin becomes stabilized relative to the native basin. At $\overline{T} = 1.33$, the free energies of the native and mirror become equal, $F^{\dagger}_N(Q^{AA}_N = 0.5) = F^{\dagger}_M(Q^{AA}_M = 0.4)$ (data not shown).

3.3.3. Side-Chain Packing Differences May Lead to Native Preference in Bdpa. The native state of Bdpa consists of three amphipathic helices that turn their hydrophobic faces inward to form a common hydrophobic core. Packing the helices in the mirror configuration disrupts the hydrophobic core, which can be restored through a reorientation of the side-chain packing (Figure 1). This reorientation completely reorganizes the identities of the side chains that pack together. Analysis of REMD shows that the native and mirror configurations have few hydrophobic core contacts in common, and that the native configuration is more tightly packed than the mirror configuration.

Figure 6A compares the contact maps of the structures in the native basin to the structures in the mirror basin. (See Section 2.4 for a description of the contact maps.) This analysis showed that the native and mirror basins had only eight tertiary residue contacts in common. While the two basins share nearly all secondary structure α-helical contacts, turn 1 (N-terminal) diverges between the two structures. This difference in turn 1 is a consequence of helix 1 (N-terminal) packing against helices 2 and 3 with different registers between the two basins. [Amphipathic helices can be broken in heptad repeats with the first and fourth residues hydrophobic. These two residues are colored greyscale in Figure 6. If two helices have the same interacting heptad repeats as in the native structure, then the hydrophobic packing between them is considered "in register".] Notice that in Figure 1A, the N-terminus of the native is one full α-helical turn farther from the C-terminus compared to the mirror. The register shift combined with the helical rearrangement makes the hydrophobic core packing between the two basins very different.

The contact map calculation shows that the native basin and its representative structure S_N are more tightly packed than their mirror counterparts. The native basin contact map C_N has 65 tertiary residue contacts, and S_N's contact map C^{AA}_N has 254 tertiary atomic contacts (444 total), while C_M has 51 tertiary residue contacts and C^{AA}_M has 194 tertiary atomic contacts (373 total). This difference in contact number is a contributing factor in stabilizing the native basin relative to the mirror basin when going from a single-basin SBM to a dual-basin SBM (Figure 5C). In order to equalize the energetics between the native and mirror structures, each mirror contact must be strengthened by a factor of 444/373 = 1.2 (Section 2.3). This means that, even though the native basin and mirror basin overlap structurally in equal amounts, i.e., structures with $Q^{AA}_N = 0.5$ average 30 C^{AA}_M tertiary contacts and structures with $Q^{AA}_M = 0.4$ average 31 C^{AA}_N tertiary contacts, since a mirror contact formed in the native basin is more favorable than a native contact made in the mirror basin, the overlap of the two folded basins energetically favors the native basin. $\langle Q^{AA}_M(Q^{AA}_N = 0.5)\rangle = 0.15$, while $\langle Q^{AA}_N(Q^{AA}_M = 0.4)\rangle = 0.12$.

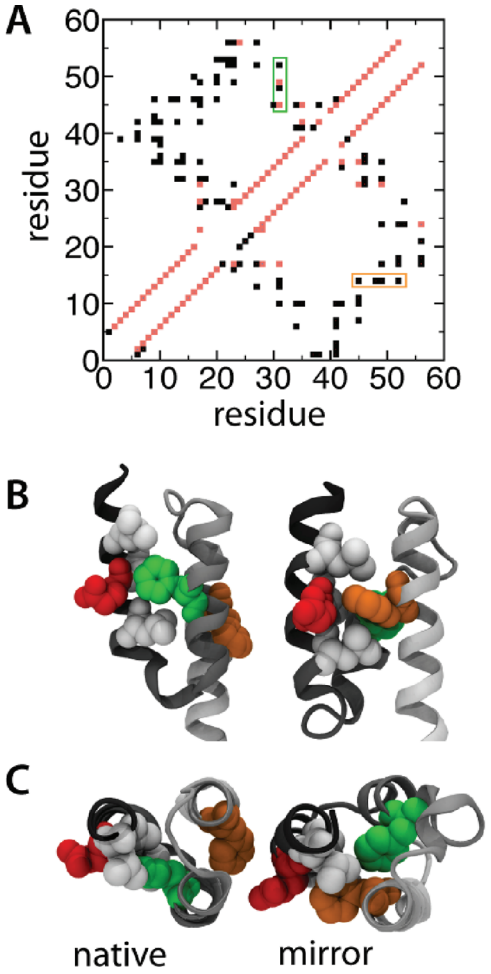


Figure 6. Hydrophobic core packing differs between the native and mirror in Bdpa. (A) Native contact map for the native (black) and mirror (red). The contacts shared between the two structures are colored blue. While the two configurations share nearly all $(i,i+4)$ α-helical contacts, they share only eight tertiary contacts. (B) Detailed view of the tertiary packing of Phe14 (orange) and Phe31 (green) with hydrophobic residues Leu45, Ala49, and Leu52 (white) and charged Glu48 (red). (C) Rotated view shows Phe31 inserted into the center of the core of the mirror. The rectangles in A highlight the similar packing of Phe14 and Phe31 with helix 3.

The poorly packed mirror conformation may be caused by the difficulty in packing Phe31. In S_N, Phe31 is tightly sandwiched between Leu45 and Leu52 (Figure 6B,C), reminiscent of interactions between coiled-coils.[51] As the hydrophobic core reorganizes to compensate for the mirror helical packing, Phe14 is similarly sandwiched between Leu45 and Leu52. Phe31 though cannot mimic the native packing of Phe14 in the mirror configuration, so it is buried into the middle of the hydrophobic core, disrupting the packing of the core along with distorting both turn 1 and the beginning of helix 2. The benefit of sandwiching Phe14 between Leu45 and Leu52 may be the reason why helix 1 undergoes a register shift in S_M.

dx.doi.org/10.1021/jp212623d | J. Phys. Chem. B 2012, 116, 6880−6888

4. DISCUSSION AND CONCLUSIONS

We have presented a thorough study of the occurrence of mirror images in simple, symmetrical proteins. A coarse-grained, single-funnel protein model highlights the near structural degeneracy between the native and mirror helical arrangements. It shows that the right-handed mirror helical arrangement is easily accessible, with roughly 5% occupation. The existence of suitable hydrophobic core packing in the mirror conformations is demonstrated by all-atom structure prediction with PFF01/02. It produces competitive folded enthalpies between native and mirror configurations of Bdpa,[19] Edpa and α_3d. Equilibrium folding simulations of Bdpa, using REMD in Amber94, result in the native cluster being only 3 times more occupied than the mirror cluster. Relaxation to equilibrium in the REMD suggests that the kinetics of folding to the mirror are faster, and thus that the mirror configuration can function as a kinetic trap during folding. Folding simulations of single- and dual-basin all-atom SBMs constructed from the REMD structures corroborate the REMD thermodynamic and kinetic findings.

These results are not only self-consistent, they also agree with previous studies on protein folding and structure prediction using empirical forcefields that also suggested the presence of mirror image folds.[15−19,48,49] Favrin et al.[18] studied a reduced model protein based on Bdpa using an empirical model focused on hydrogen bonding and hydrophobicity. The authors noted that it was difficult to differentiate between the native fold and its mirror image in their pariwise-additive potential. Scheraga and co-workers[16] studied the folding of Bdpa and apo calbindin D9K in a thermodynamic framework using the UNRES forcefield.[53] They were successful in locating the native state, but they also encountered mirror images for both proteins, and noted that the mirrors were difficult to discriminate based only on their energies (within a few kcal/mol). More recently, Scheraga and co-workers[49] showed using comprehensive sampling of Bdpa with the UNRES forcefield that, not only do folding trajectories at low temperature often visit the mirror configuration as kinetic trap, but also at folding temperature the native and mirror configurations are equally populated. While the occurrence of mirror images has often been regarded as a deficiency of empirical energy functions, their consistent observation among diverse forcefields, across multiple proteins, and from coarse-grained to all-atom representations, is convincing evidence that mirrored protein conformations are truly competitive.

Solution NMR has been performed on three homologous domains of protein A, Bdpa,[51] Edpa,[54] and the Z-domain (Zdpa),[55] which has two mutations relative to Bdpa: Ala1 → Val and Gly29 → Ala. There are also X-ray crystallography structures of Bdpa in complex with human IgG Fc fragment (Bdpa-Fc)[56] and Ddpa in complex with human IgM (Ddpa-Fab).[26] Bdpa-Fc lacks helix 3 and has no coordinates at all for Ala49-Lys59, while all three solution NMR structures and Ddpa-Fab show a tightly packed three-helix bundle with some N-terminal fraying of helix 1. HD-exchange experiments, though, do show protection of the helix 3 hydrogen bonds in Bdpa-Fc, which suggests that the lack of helix 3 in the crystal structure may be a crystal artifact. The solution NMR structures and Ddpa-Fab are largely consistent, but all show different orientations of helical packings. In Bdpa, helix 1 is tilted 30° with respect to helices 2 and 3, while in Zdpa and Ddpa-Fab, helix 1 is only tilted 15°, and in Bdpa-Fc helix 1 and helix 2 are nearly parallel. Although none of the experimental structures display the mirror image, the native structural heterogeneity can be taken as a sign of frustration in the domain. Likely, nonspecific hydrophobic interactions give rise to many adequate hydrophobic core packings. Protein A is a virulence factor; one of its goals is to bind strongly to immunoglobulins. Achieving strong binding may require the domain to have considerable flexibility.[14,57] Helix swapping may be part of the functional dynamics of the system, and these proteins may adopt one conformation or the other depending on the proteins to which they bind.

Regardless of any functional advantages of mirror images, energy landscape considerations predict that the two symmetric helical arrangements of three-helix bundles should only marginally differ in stability. The principle of minimal frustration[3,2] explains that evolution works to ensure a sequence is consistent with its structure. Geometry therefore becomes the prime determinant of the folding landscape. When symmetry leads to degeneracy between protein structures, the choice between symmetrical helical arrangements occurs at a finer energy scale. Our results show that this difference may be as small as a few k_BT. As conditions change, or new interaction partners are introduced, the energy landscape is altered, and the protein may fall through a trapdoor[24] to its symmetric structural neighbor. In the case of Bdpa REMD, temperature stabilizes the mirror relative to native, similar to the Rop-dimer. Therefore, like the Rop-dimer, the Bdpa mirror may also be stabilized by denaturant.[25] Another intriguing possibility is that a Phe31 → Ala mutant will open the trapdoor. If these mirror conformations indeed exist as metastable excitations from the native basin, sensitive NMR experiments should be able to capture their signatures.[58] More experimental studies are needed in order to quantify the energy landscapes of symmetric protein structures.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Figure S1 presents data showing the degree of equilibration achieved with REMD. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: angel@rpi.edu; phone: (518) 276-6310; fax: (518) 276-6680 (A.E.G.). E-mail: jonuchic@rice.edu; phone: (713) 348-4197; fax: (858) 534-7697 (J.N.O.).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Center for Theoretical Biological Physics sponsored by the National Science Foundation (NSF) (Grant PHY-0822283) and NSF Grants NSF-MCB-1214457 to J.N.O. and NSF-MCB-1050966 to A.E.G. J.N.O. is a CPRIT Scholar in Cancer Research sponsored by the Cancer Prevention and Research Institute of Texas. A.S. and A.V. acknowledge the Impuls- und Vernetzungsfond of the Helmholtz Association of German Research Centers. W.W. is supported by BW Stiftung Grant HPC-5.

REFERENCES

(1) Anfinsen, C. B. *Science* **1973**, *181* (4096), 223−230.
(2) Onuchic, J. N.; Wolynes, P. G. *Curr. Opin. Struct. Biol.* **2004**, *14* (1), 70−75.
(3) Bryngelson, J.; Wolynes, P. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 7524.
(4) Leopold, P. E.; Montal, M.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89* (18), 8721−8725.
(5) Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93* (25), 14249−14255.
(6) Kim, D.; Fisher, C.; Baker, D. *J. Mol. Biol.* **2000**, *298* (5), 971−984.
(7) McCallister, E. L.; Alm, E.; Baker, D. *Nat. Struct. Mol. Biol.* **2000**, *7* (8), 669−673.
(8) Clementi, C.; García, A. E.; Onuchic, J. N. *J. Mol. Biol.* **2003**, *326* (3), 933−954.
(9) Zhou, Y.; Linhananta, A. *Proteins* **2002**, *47* (2), 154−162.
(10) Broom, A.; Doxey, A. C.; Lobsanov, Y. D.; Berthin, L. G.; Rose, D. R.; Howell, P. L.; Mcconkey, B. J.; Meiering, E. M. *Structure* **2012**, *20* (1), 161−171.
(11) Gosavi, S.; Chavez, L. L.; Jennings, P. A.; Onuchic, J. N. *J. Mol. Biol.* **2006**, *357* (3), 986−996.
(12) Gosavi, S.; Whitford, P. C.; Jennings, P. A.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105* (30), 10384−10389.
(13) Capraro, D. T.; Roy, M.; Onuchic, J. N.; Gosavi, S.; Jennings, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109* (5), 1490−1493.
(14) Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (18), 6837−6838.
(15) Kolinski, A.; Skolnick, J. *Proteins* **1994**, *18* (4), 353−366.
(16) Lee, J.; Liwo, A.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96* (5), 2025−2030.
(17) Irbäck, A.; Sjunnesson, F.; Wallin, S. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97* (25), 13614−13618.
(18) Favrin, G.; Irbäck, A.; Wallin, S. *Proteins* **2002**, *47* (2), 99−105.
(19) Herges, T.; Wenzel, W. *Biophys. J.* **2004**, *87* (5), 3100−3109.
(20) Bennett, M. J.; Choe, S.; Eisenberg, D. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91* (8), 3127−3131.
(21) Yang, S.; Cho, S. S.; Levy, Y.; Cheung, M. S.; Levine, H.; Wolynes, P. G.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (38), 13786−13791.
(22) Munson, M.; Anderson, K. S.; Regan, L. *Folding Des.* **1997**, *2* (1), 77−87.
(23) Levy, Y.; Cho, S. S.; Shen, T.; Onuchic, J. N.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102* (7), 2373−2378.
(24) Schug, A.; Whitford, P. C.; Levy, Y.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104* (45), 17674−17679.
(25) Gambin, Y.; Schug, A.; Lemke, E. A.; Lavinder, J. J.; Ferreon, A. C. M.; Magliery, T. J.; Onuchic, J. N.; Deniz, A. A. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106* (25), 10153−10158.
(26) Graille, M.; Stura, E. A.; Corper, A. L.; Sutton, B. J.; Taussig, M. J.; Charbonnier, J. B.; Silverman, G. J. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97* (10), 5399−5404.
(27) Zhu, Y.; Alonso, D. O. V.; Maki, K.; Huang, C.-Y.; Lahr, S. J.; Daggett, V.; Roder, H.; DeGrado, W. F.; Gai, F. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (26), 15486−15491.
(28) Schug, A.; Herges, T.; Wenzel, W. *Phys. Rev. Lett.* **2003**, *91* (15), 158102.
(29) Verma, A.; Wenzel, W. *Biophys. J.* **2009**, *96* (9), 3483−3494.
(30) Schug, A.; Wenzel, W. *J. Am. Chem. Soc.* **2004**, *126* (51), 16736−16737.
(31) Schug, A.; Herges, T.; Wenzel, W. *Proteins* **2004**, *57* (4), 792−798.
(32) Strunk, T.; Hamacher, K.; Hoffgaard, F.; Engelhardt, H.; Zillig, M. D.; Faist, K.; Wenzel, W.; Pfeifer, F. *Mol. Microbiol.* **2011**, *81* (1), 56−68.
(33) Schug, A.; Herges, T.; Verma, A.; Lee, K. H.; Wenzel, W. *ChemPhysChem* **2005**, *6* (12), 2640−2646.
(34) Schug, A.; Wenzel, W. *Biophys. J.* **2006**, *90* (12), 4273−4280.
(35) García, A. E.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (24), 13898−13903.
(36) Garcia, A.; Herce, H.; Paschek, D. *Annu. Rep. Comput. Chem.* **2006**, *2*, 83−95.
(37) Showalter, S. A.; Bruschweiler, R. *J. Chem. Theory Comput.* **2007**, *3* (3), 961−975.
(38) Day, R.; Paschek, D.; Garcia, A. E. *Proteins* **2010**, *78* (8), 1889−1899.
(39) Noel, J. K.; Onuchic, J. N. In *Computational Modeling of Biological Systems*; Dokholyan, N., Ed.; Springer: New York, 2012; Chapter 2.
(40) Lammert, H.; Schug, A.; Onuchic, J. N. *Proteins* **2009**, *77* (4), 881−891.
(41) Whitford, P. C.; Noel, J. K.; Gosavi, S.; Schug, A.; Sanbonmatsu, K. Y.; Onuchic, J. N. *Proteins* **2009**, *75* (2), 430−441.
(42) Noel, J. K.; Whitford, P. C.; Onuchic, J. N. *J. Phys. Chem. B* **2012**, submitted for publication.
(43) Noel, J. K.; Whitford, P. C.; Sanbonmatsu, K. Y.; Onuchic, J. N. *Nucleic Acids Res.* **2010**, *38*, W657−61.
(44) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4* (3), 435−447.
(45) Whitford, P. C.; Miyashita, O.; Levy, Y.; Onuchic, J. N. *J. Mol. Biol.* **2007**, *366* (5), 1661−1671.
(46) Baxter, E. L.; Jennings, P. A.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109* (6), 1955−1960.
(47) Clementi, C.; Nymeyer, H.; Onuchic, J. N. *J. Mol. Biol.* **2000**, *298* (5), 937−953.
(48) St-Pierre, J.; Mousseau, N.; Derreumaux, P. *J. Chem. Phys.* **2008**, *128*, 045101.
(49) Maisuradze, G. G.; Liwo, A.; Ołziej, S.; Scheraga, H. A. *J. Am. Chem. Soc.* **2010**, *132* (27), 9444−9452.
(50) Sato, S.; Religa, T. L.; Daggett, V.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (18), 6952−6956.
(51) Gouda, H.; Torigoe, H.; Saito, A.; Sato, M.; Arata, Y.; Shimada, I. *Biochemistry* **1992**, *31* (40), 9665−9672.
(52) Cho, S.; Levy, Y.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103* (3), 586−591.
(53) Rojas, A. V.; Liwo, A.; Scheraga, H. A. *J. Phys. Chem. B* **2007**, *111* (1), 293−309.
(54) Starovasnik, M. A.; Skelton, N. J.; O'Connell, M. P.; Kelley, R. F.; Reilly, D.; Fairbrother, W. J. *Biochemistry* **1996**, *35* (48), 15558−15569.
(55) Tashiro, M.; Tejero, R.; Zimmerman, D. E.; Celda, B.; Nilsson, B.; Montelione, G. T. *J. Mol. Biol.* **1997**, *272* (4), 573−590.
(56) Deisenhofer, J. *Biochemistry* **1981**, *20* (9), 2361−2370.
(57) Wahlberg, E.; Lendel, C.; Helgstrand, M.; Allard, P.; Dincbas-Renqvist, V.; Hedqvist, A.; Berglund, H.; Nygren, P.-A.; Härd, T. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (6), 3185−3190.
(58) Bouvignies, G.; Vallurupalli, P.; Hansen, D. F.; Correia, B. E.; Lange, O.; Bah, A.; Vernon, R. M.; Dahlquist, F. W.; Baker, D.; Kay, L. E. *Nature* **2011**, *477* (7362), 111−114.
(59) Kumar, S.; Rosenberg, J.; Bouzida, D.; Swendsen.. *J. Comput. Chem.* **1992**, *13* (8), 1011.