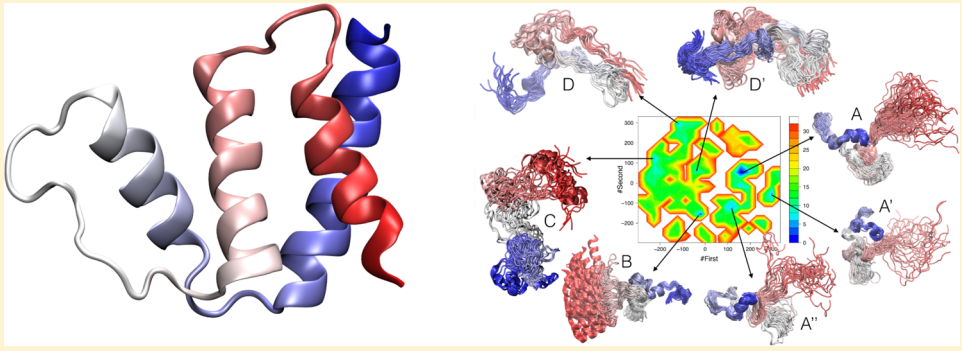
JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

Statistical Mechanics of the Denatured State of a Protein Using Replica-Averaged Metadynamics

Carlo Camilloni and Michele Vendruscolo*

Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, United Kingdom

Supporting Information



ABSTRACT: The characterization of denatured states of proteins is challenging because the lack of permanent structure in these states makes it difficult to apply to them standard methods of structural biology. In this work we use all-atom replica-averaged metadynamics (RAM) simulations with NMR chemical shift restraints to determine an ensemble of structures representing an acid-denatured state of the 86-residue protein ACBP. This approach has enabled us to reach convergence in the free energy landscape calculations, obtaining an ensemble of structures in relatively accurate agreement with independent experimental data used for validation. By observing at atomistic resolution the transient formation of native and non-native structures in this acid-denatured state of ACBP, we rationalize the effects of single-point mutations on the folding rate, stability, and transition-state structures of this protein, thus characterizing the role of the unfolded state in determining the folding process.

INTRODUCTION

Molecular dynamics simulations represent a powerful approach to study biomolecules at atomic resolution.[1,2] This method has a long history in structural biology, and the range of its applications is steadily increasing.[1−14] Most of the methodological efforts are being devoted to overcoming two major aspects of this method, the extension of sampling of the conformational space and the accuracy of the force fields used to integrate the equations of motion. Great advances have been recently made in the improvement of hardware and software,[15−18] in the parametrization of the force fields,[19−30] and in the implementation of advanced algorithms to enhance the sampling efficiency.[31−41]

In particular, a great variety of approaches have been proposed to improve the quality of the force fields by considering available experimental information.[19−30] Among these approaches, replica-averaged molecular dynamics simulations have been successfully employed with a number of different experimental observables, including specifically those measured by NMR spectroscopy, for a number of different systems.[42−54] Recently, it has been shown that the use of replica-averaged simulations represents an effective way to interpret experimental data in the sense of the maximum entropy principle,[55−58] a result that implies that the inclusion of experimental data in a force field in this way provides the approach with the least number of assumptions needed to modify the force field itself to make it consistent with the measured data.[55−58]

In order to combine the strengths of replica-averaged simulations with that of advanced sampling techniques we recently proposed the replica-averaged metadynamics (RAM) approach,[59] in which a force field is modified in a system-dependent manner through the use of replica-averaged experimental data, while the conformational sampling is enhanced by metadynamics,[37] a versatile advanced sampling technique that has been applied to a number of challenging problems.[36] In the RAM approach we used NMR chemical shifts, as these parameters have recently emerged as powerful reporters of the structure and dynamics of native,[60−69] intermediate[70−72] and disordered[73−75] states of proteins. In unfolded and denatured states and for intrinsically disordered proteins chemical shifts are routinely used as probe for secondary structure propensities[76] and populations[77] and have been also used to probe tertiary structure contact propensities.[78,79]

In this work we implemented the RAM approach[59] with backbone chemical shift restraints[64,80] and bias-exchange

Received: March 22, 2014
Published: June 2, 2014

   dx.doi.org/10.1021/ja5027584 | J. Am. Chem. Soc. 2014, 136, 8982−8991

metadynamics[38] to sample the conformational space of the acid-denatured state of ACBP at pH 2.3, a 86-residue protein (Figure 1) whose folding mechanism has been extensively characterized.[52,54,78,79,81−86] This approach enabled us to obtain a converged sampling and thus to characterize the major structural features of this state. From the converged sampling, we can derive an accurate free energy landscape, which is a fundamental quantity in a statistical mechanics description of a disordered protein. The ensemble of conformations that we determined provides insights into the relationship between the residual structure of the denatured state and the folding mechanism of this protein. We found in particular correlations between the fractions of native and of non-native contacts and the change in the folding rate upon mutation as well as the Φ-values for the transition state. We also found negative correlations between the fractions of native and of non-native contacts and the change in stability upon mutation. Taken together, these results suggest that the residual structure of the denatured state, both native and non-native, can provide quantitative information about the folding process of this protein.

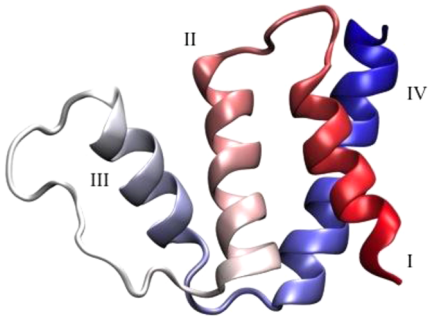


Figure 1. Native structure of ACBP. The four native α-helices (I−IV) are colored from the N-terminal in red to the C-terminal in blue. The structure is taken from the PDB file 1NTI.

METHODS

Molecular dynamics simulations of ACBP were performed using the Amber03W force field[87] with the TIP4P05 water model.[88] All the simulations were run in GROMACS[89] using PLUMED2.[90] A time step of 2 fs was used together with LINCS constraints.[91] The van der Waals interactions were implemented with a cutoff at 0.9 nm, and long-range electrostatic effects were treated with the particle mesh Ewald method.[92] All simulations were carried out in the canonical ensemble at constant volume and by thermosetting the system with the Nosè-Hoover thermostat.[93] The starting conformation was taken from an available NMR structure (PDB code 1NTI). This structure was fully protonated in order to mimic the acidic conditions and solvated with 18,000 water molecules in a dodecahedron box of 700 nm^3 of volume. A high-temperature (450 K) 30 ns preliminary unfolding simulation was used to select four starting conformations. Each conformation was then subsequently relaxed at 300 K for 10 ns.

RAM simulations[59] were performed using chemical shifts as replica-averaged restraints[64,94] and bias-exchange metadynamics.[38] The bias-exchange metadynamics approach combines replica exchange[95] with metadynamics,[36,37] in which several metadynamics simulations are performed in parallel on different replicas of the system, each replica biasing a different collective variable (CV). Exchanges between the replicas are attempted periodically according to a replica-exchange scheme.

The introduction of the bias-exchange metadynamics method in the RAM scheme offers the possibility of performing replica-averaged restrained simulations without the need of using further replicas at different conditions, as in the case of a previous RAM implementation[59] (Figure S1). It is thus possible to maximize the use of the available computational resources by simulating a limited number of replicas.

Four replicas of the system were simulated in parallel at 300 K with a restraint applied on the average value of the CamShift[80,96] back-calculated NMR chemical shifts[94]

$$E^{cs} = \alpha \sum_{k=1}^{N} \sum_{l=1}^{6} \left(\delta_{kl}^{exp} - \frac{1}{M} \sum_{m=1}^{M} \delta_{klm}^{calc} \right)^2 \quad (1)$$

where the α is the force constant, set to 24 kJ/(mol ppm^2), k runs over the amino acids of the protein, l runs over the six backbone atoms used in the simulations (Cα, Cβ, C′, Hα, H_N and N) and m runs over the $M = 4$ replicas (the chemical shifts of Pro, Gly, Asp, Glu and His residues were not used in the simulations). In this way the system evolves with a force field that is perturbed in such a way to increase the agreement with the experimental chemical shifts[84] as resulting by the application of the maximum entropy principle.[55−58] In principle the number of replicas can be increased at expense of an increasing computational cost, but, as previously shown,[59] four replicas are sufficient to recover with very good accuracy the dynamics from chemical shifts.

In the present case four CVs have been employed: the total α-helical content (the "αC" CV),[97] the total β-sheet content (the "βC" CV),[97] the radius of gyration (the 'R_g' CV), and the distance between the centers of mass of α-helix I (residues 2−13) and α-helix IV (residues 66−85) (the "dI-IV" CV), as described.[31,98] The choice of the secondary structures and radius of gyration as CVs was guided by the hypotheses that transient secondary structures formation, and the volume fluctuations of the polypeptide chain capture to a good extent the relevant dynamics of a disordered protein. The additional choice of dI-IV as a CV was suggested by the fact that the docking of α-helix I and IV has been identified as the rate-limiting step for the folding of ACBP. dI-IV and R_g are weakly correlated along the simulations, $r \sim 0.3$, suggesting that dynamic of the two α-helices is in part independent by the volume fluctuations. Gaussians deposition was performed with an initial rate of 0.125 kJ/mol/ps, where the σ values were set to 0.15, 0.11, 0.027, and 0.08 nm, for αC, βC, R_g and dI-IV, respectively.

It has been shown that if the collective variables capture the relevant slow dynamics of a system, a metadynamics run can converge to the opposite of the free energy landscape along the selected collective variables, but only if the heights of the Gaussians are slowly rescaled to zero[33,35] and if the borders of the collective variables are properly treated.[99] In order to keep under control the convergence of the simulations we rescaled the height of the Gaussians using the well-tempered scheme[100] with a bias-factor of 10. Furthermore, in order to limit the extent of accessible space along each collective variable and correctly treat the problem of the borders, we set the bias as constant outside a defined interval for each CV,[101] as it has been shown that this approach lead to a correct reconstruction of a one-dimensional free energy landscape inside the chosen range; intervals were set to 0−36, 0−11, 1.3−3.8, and 0−4.2 for the four CVs, respectively. Each replica have been evolved for 650 ns, with exchange trials every 50 ps.

The convergence of the sampling was assessed by monitoring the differences of the free energies in the range between 0 and 50 kJ/mol at increasing simulation length during the simulations. The average differences between two free energy landscapes calculated at 13 ns intervals are shown in Figure S2. After the first 450 ns the free energy landscapes are stable within <2.5 kJ/mol, suggesting that all the relevant minima in the landscape have been found, and the average changes in the free energy landscapes over the last 80 ns of simulations are below 1 kJ/mol (Figure 3). These results suggest that the free energies that we obtained from the RAM simulations are on average correct within <2 kJ/mol (Figure S2).

The sampling of the four replicas was used to generate a four-dimensional free energy landscape where a set of microstates is identified by dividing the four-dimensional CV-space into a homogeneous grid of small hypercubes whose free energies were

dx.doi.org/10.1021/ja5027584 | J. Am. Chem. Soc. 2014, 136, 8982−8991

obtained using a standard weighted histogram analysis.[31,101−103] The free energy of each microstate and its estimated error are shown in Figure S3. The free energy landscape as a function of three out of four CVs is shown in Figure 3 together with three representative microstates, which were selected as the microstates that corresponds to local minima separated by the highest barriers in the landscape.

RESULTS

Convergence of the Free Energy Landscape and Role of the Chemical Shifts. In order to determine the conformational properties of the acid-denatured state of ABCP at pH 2.3 that we studied here, we used replica-averaged metadynamics (RAM)[59] simulations. This technique combines the sampling efficiency of the metadynamics approach,[37] which we implemented here in its bias-exchange version,[38] with the on-the-fly modification of the force field using experimental data, in this case backbone NMR chemical shifts,[59] as ensemble-averaged structural restraints.[52−54] It has been recently shown that this ensemble-averaged approach provides ensembles of structures that, given a force field, are in the best possible agreement with the experimental data used as restraints.[55−58]

The application of the RAM simulations to the acid-denatured state of ACBP provides a description of this state at a fraction of the computational cost needed for unrestrained molecular dynamics simulations. Indeed, it has been recently shown that it is challenging to reach convergence in the free energy calculations of the denatured state of ACBP using standard molecular dynamics simulations.[104] It has also been reported that the quality of current force fields is still often relatively poor when they are used to describe unstructured states of proteins.[105,106] For example, the stability of secondary structures elements is strongly dependent on the force field employed, and the overall conformations of the proteins are in many cases too compact.[105,106]

Using the RAM approach, after a total simulation time of 2.5 μs we obtained a converged sampling resulting in a free energy landscape within a statistical uncertainty of <2 kJ/mol for free energies up to 10 kJ/mol (Figures 2, S2, and S3). The collective variables explore their accessible phase space in a relatively uniform manner (Figure S4) and the exchange rate among replicas is ∼10%. The structures in the ensemble representing the acid-denatured state of ACBP cover a broad range of values for the collective variables (CVs) used in the simulations: the radius of gyration (R_g) was between 1.6 and 3.8 nm, the α-helical content (αC) between 0% and 36%, the β-sheet content (βC) between 0% and 10%, and the dI-IV distance between 0.7 and 4.0 nm. The distribution of the RMSD values calculated among all pairs of sampled structures provides a measure of the size of the conformational space visited during the simulations; the sampled structures are on average at 1.7 ± 0.7 nm RMSD from each other (Figure S5).

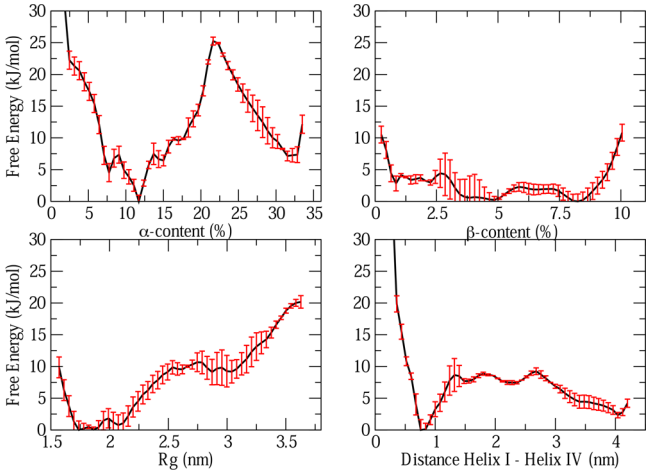


Figure 2. Convergence of the free energy calculations. The free energy landscapes (in kJ/mol) of the four different collective variables (see Methods) are averaged over the final segment (the last 80 ns) of the RAM simulations, and their standard deviations are reported as red error bars.

The effects of the chemical shift restraints on the force field can be analyzed from the point of view of their contributions to the energy. We estimated these effects to be small since their associated energy is on average 480 ± 40 kJ/mol per replica, a value to be compared with an average total energy of $-6.66 \times 10^5 \pm 2 \times 10^3$ kJ/mol per replica. We also carried out an analysis of the average force intensities per atom resulting from the addition of the chemical shift restraints to the force field. The average chemical shift contributions to the total forces on the protein is ∼15% (Figure S6). An analysis per amino acid suggest that in the case of Amber03W the chemical shifts are in particular acting on the Trp side chain atoms and Arg backbone and side chain atoms (Figure S6).

Analysis of the Microstates in the Acid-Denatured State of ACBP. The free energy landscape as a function of three of four CVs used in the simulations is characterized by the presence of three distinct free energy basins (Figure 3). The most populated basin (84% population, labeled as A in Figure 3) corresponds to a R_g of 2.0 nm (the folded protein has a R_g of 1.3 nm), a αC of 12% (which is localized in the region of α-helix IV), and a βC of the 8% (which is localized in the region or residues 22−27). The second most populated basin (12% population, labeled as B in Figure 3) corresponds to a R_g of 1.7 nm, an αC of 32% (which is localized in the region of α-helices I and IV), and a βC of 3%. The third basin (4% population, labeled as C in Figure 3) is the most extended and less structured, with a R_g of 2.7 nm, an αC of 6%, and a βC of 1%. A clear feature that emerges from this free energy landscape is that the acid-denatured state is characterized by two well-separated regions (Figure 3). The first region, which is formed by basins A and B is more compact and structured with an average R_g of about 2.0 nm, and the second, less populated region corresponding to basin C, is more extended and less structured, at higher free energy with an average R_g close to 3.0 nm. This expanded state resembles a random coil conformation, and the separation between the two basins resembles a coil−globule transition. The overall ensemble averaged over the entire free energy landscape R_g is of 1.96 ± 0.04 nm, a value significantly higher than that of 1.75 ± 0.05 nm found in a recent 200 μs long standard molecular dynamics simulation, which did not reach convergence.[104] We also note that the height of the barriers in the free energy landscapes that we presented does not necessarily correspond quantitatively to the height of the barriers along the reaction coordinates, which is consistent with the idea that the basins that we identified are in fast exchange on the chemical shifts time scale.

The visualization of the microstates presented above was based on the collective variables employed in the metadynamics simulations. In order to generalize the results to other parameters that can be employed to represent the free energy landscape we used the recently developed SketchMap

dx.doi.org/10.1021/ja5027584 | J. Am. Chem. Soc. 2014, 136, 8982−8991

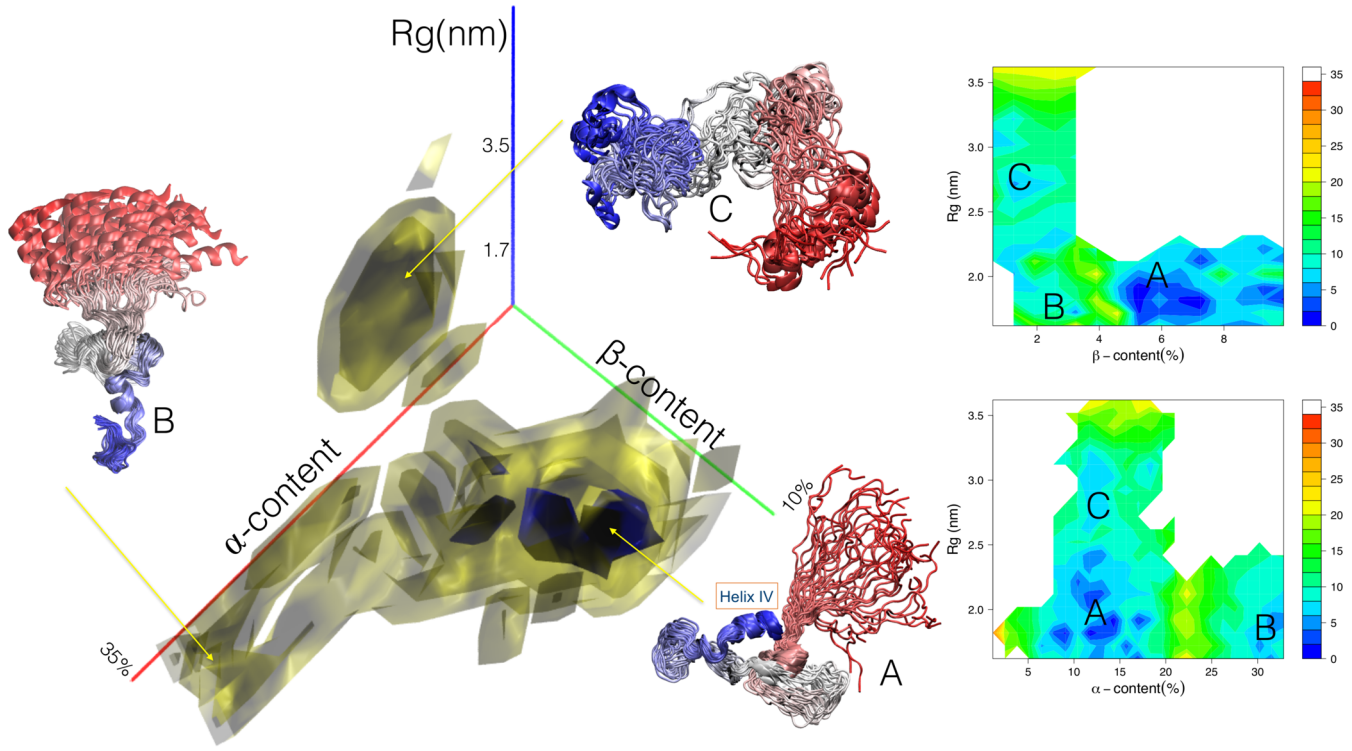


Figure 3. Identification of three major microstates in the acid-denatured state of ACBP. (a) Three-dimensional free energy landscape (in kJ/mol) as a function of the collective variables for the α-helical content, the β-sheet content, and the radius of gyration (αC, βC and R_g); representative structures (color code as in Figure 1) are shown for the three major microstates, which are labeled as A, B and C, and have statistical weights of, respectively, 84%, 12%, and 4%. (b,c) Two-dimensional projections of the three-dimensional free energy landscape shown in (a).

projection technique.[107−109] Sketchmap provides a way of reducing a multidimensional representation to two dimensions by means of Mercator-like approach in which the distance between two structures in the multidimensional space is qualitatively preserved after a rescaling through a step function.[107−109] In the present case we analyzed our trajectories in terms of the 172 backbone dihedral angles of ACBP, and we employed SketchMap to identify two parameters (or CVs) for the projection of the free energy. We then used these two CVs together with the four old ones to build a new bidimensional free energy landscape as a function of the SketchMap CVs only (Figure 4). The most populated microstate corresponds to the ground state (i.e., basin A) identified in Figure 3. In the lower right region of the free energy landscape in Figure 4 the microstates show a partially formed α-helix IV with the rest of the protein largely disordered. Going toward the left side of the landscape, α-helix IV becomes less and less formed while a partial structuring emerges in α-helix I. In the upper region of the free energy landscape α-helices I and IV are not formed, while there is a partial formation of α-helix II and the presence of some β-structure. The free energy landscape in Figure 4 is thus particularly suitable to represent the transient formation of secondary structure elements in the acid-denatured state of ACBP. An additional representation of the transient structure of the denatured state is provided in Figure 5 where the average contact maps of the various microstates of the denatured ensemble are compared with the native contact map. Consistent with the results presented in Table 1, microstate D exhibits the largest number of transient contacts, while microstate C the lowest one.

Validation with Experimental Data Not Used As Restraints in the Simulations. In order to validate the ensemble of structures of the acid-denatured state of ACBP determined in this work, we used it to back-calculate a series of experimental observables that were not used as restraints in the RAM simulations. Each observable was first averaged over the structures of each microstate and then was weighted for the free energy of the microstate.[31,101−103]

We first verified that the back-calculated backbone NMR chemical shifts (obtained using SPARTA+),[110] which were used as restraints in the RAM simulations (where they were calculated using CamShift)[96] are in agreement with the experimental ones (Figure S7). The quality of this agreement (0.49, 0.37, 0.09, 0.15, 0.48, and 1.26 ppm for Cα, Cβ, Hα, HN, C′, and N chemical shifts, respectively) can be assessed by considering that the differences between experimental and back-calculated chemical shifts are smaller, in particular for the Cα and Hα atoms, than the average deviations between the experimental chemical shifts and the random coil chemical shifts predicted using the CamCoil method[111] (0.83, 0.37, 0.17, 0.20, 0.56, and 1.30 ppm for Cα, Cβ, Hα, HN, C′, and N chemical shifts, respectively, Figure S7). We also compared the experimental and calculated secondary chemical shifts with respect to an intrinsic experimentally measured scale[78] (Figure S8). The agreement between the two approaches for estimating residual secondary structures is particularly good (coefficient of correlation of 0.64) if one consider the limitations of the current empirical chemical shifts predictors.[66]

Since the secondary structures populations also reflect the agreement with the secondary chemical shifts, we calculated secondary structures for the ensemble using STRIDE[112] and using a polyproline II (PPII) definition recently reported.[113] The total average α-helical content in the acid-denatured ensemble of ACBP that we determined was of 13 ± 1% to be

dx.doi.org/10.1021/ja5027584 | J. Am. Chem. Soc. 2014, 136, 8982−8991

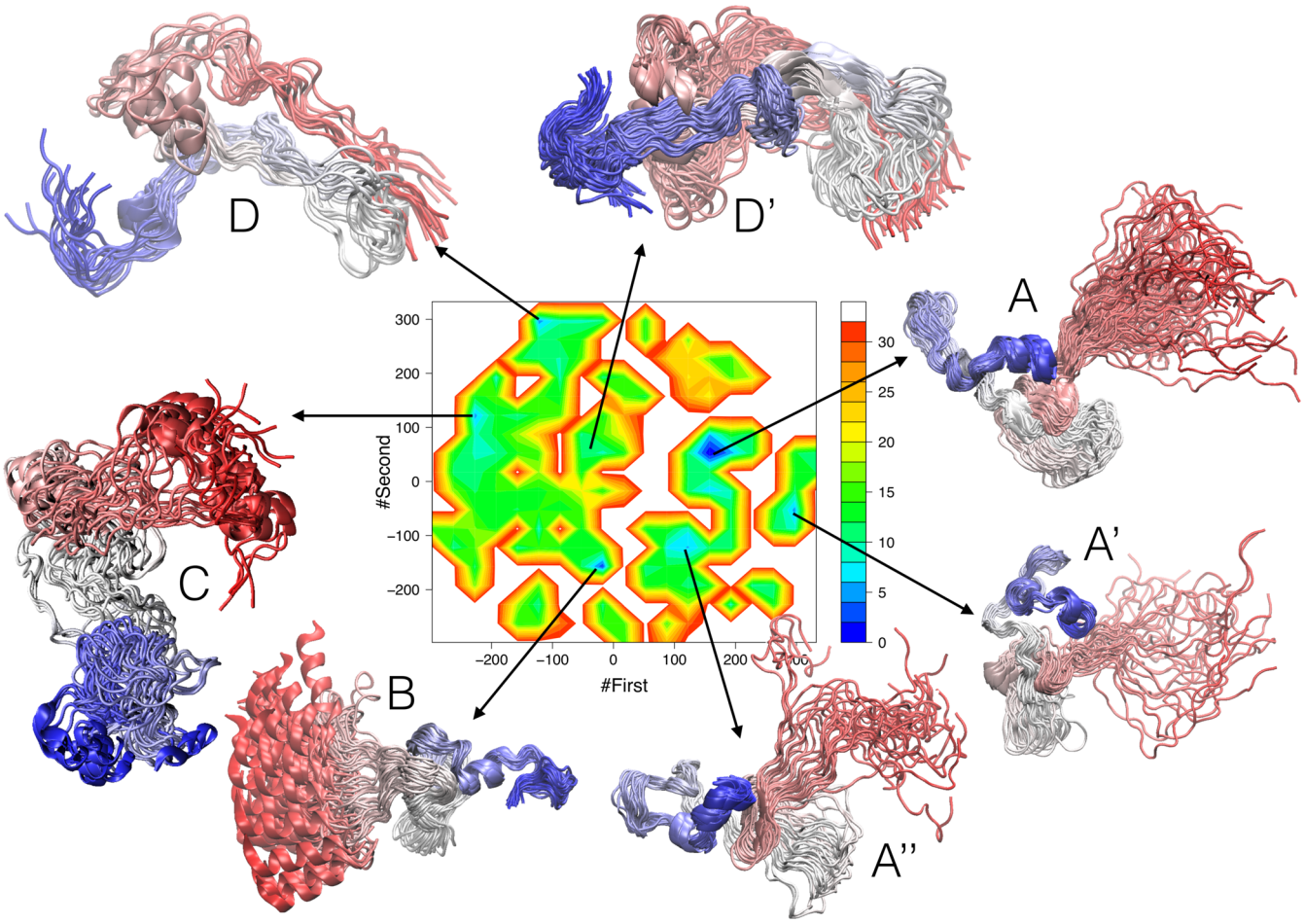


Figure 4. Characteristion of the microstates in the acid-denatured state of ACBP. Two-dimensional free energy landscape (in kJ/mol) as a function of the two SketchMap collective variables. A set of representative structures (color code as in Figure 1) is shown for all the microstates with a free energy lower than 5 kJ/mol. The statistical weights of the microstates are, respectively, A 57% (A 43%, A′ 5%, A″ 9%), B 11%, C 4%, and D 22% (D 11%, D′ 11%).

compared with the 16% estimate obtained for the same conditions using circular dichroism.[84] The total β-content and PPII-content account for 4 ± 1% and 10 ± 1%, respectively. Also, we used the δ2D method to predict secondary structures populations from chemical shifts,[77] with a total α-, β- and PPII-content of 8%, 2%, and 3%, respectively. These predictions are in good agreement with the results obtained by calculating directly the secondary structure populations from the structures in the ensemble considering that in the δ2D predictions the standard error is about 5%[77] and that the calculated and predicted secondary structure population profiles follows very closely each other (Figure S9). The largest difference is in the population of α-helix III. Overall, the differences in secondary structure populations in Figure S9 are within the uncertainties of the different methods used to calculate them. One could thus expect such differences will be reduced in the future upon further development of these methods. In particular, with the ever-increasing expansion of the BMRB the δ2D method will increase in accuracy, and with the inclusion of other types of restraints (e.g., residual dipolar couplings) the RAM method will also provide more accurate results. The secondary structure populations of the ensemble are also compared with the experimental Cα secondary chemical shifts, determined using an intrinsic random coil scale.[78] We observed the presence of transient β structure in regions near residues with negative secondary chemical shifts, consistently with previous reports about the presence of non-native β structure in the acid-denatured state.[78]

The overall average value of R_g of 1.96 ± 0.04 nm reported above for the acid-denatured state of ACBP corresponds to a hydrodynamic radius R_h, calculated using a standard model,[52,54] of 2.30 ± 0.05 nm, which is larger than the one derived recently by standard molecular dynamics simulations,[104] although still slightly smaller than the measured value of 2.5 nm.[81]

In addition to transient secondary structures[78,82,84] and the hydrodynamic radius,[81] the acid-denatured state of ACBP has been studied by means of chemical shift perturbations.[79] With this technique the effect of a single-point mutation is measured in terms of the small corresponding variations of the chemical shifts of other residues along the chain. By using this technique it has been suggested that the presence of transient long-range native and non-native contacts could provide possible initiation sites for the folding of ACBP.[79] In order to compare our results with these measurements we calculated the contact probability among residues, by considering a residue−residue contact to be formed every time at least two respective heavy atoms are within 5 Å. The resulting contact map is shown in Figure 5 ("ensemble" panel, above the diagonal) where it is compared with the native one (below the diagonal) and in Figure S10 where it is compared with the results from chemical shifts

dx.doi.org/10.1021/ja5027584 | J. Am. Chem. Soc. 2014, 136, 8982−8991

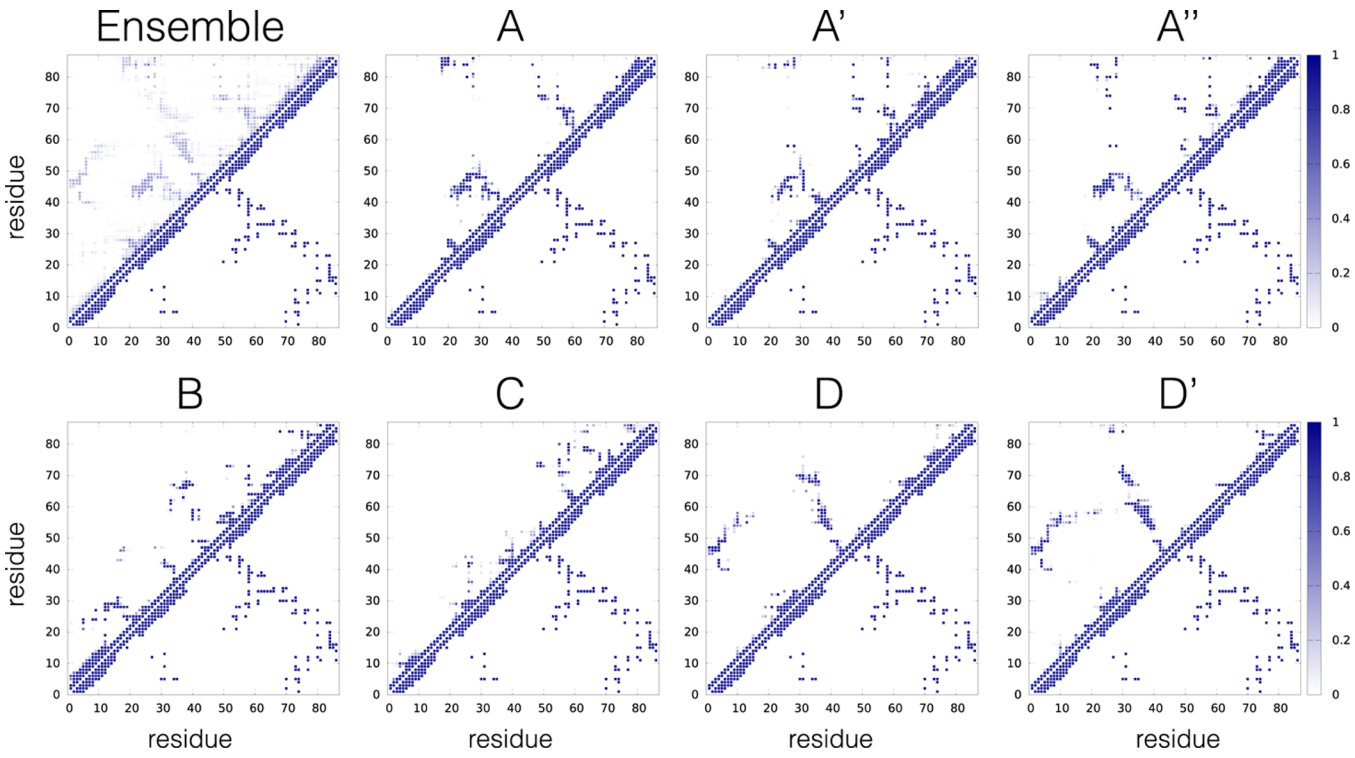


Figure 5. Residual structure in the acid-denatured state of ACBP. (a) Comparison of the average probability of contact formation in the acid-denatured state ensemble (above the diagonal) with the native state contact map (below the diagonal). (b–i) Comparison of the average probability of contact formation in the different microstates shown in Figure 4 making up the acid-denatured state ensemble (above the diagonal) with the native state contact map (below the diagonal).

Table 1. Distance of the Microstates from the Native State[a]

microstate	III–IV	II–IV	II–III	II–III–IV	I–II–III–IV
A	**0.57**	1.04	1.10	1.25	1.84
A′	0.69	0.88	0.82	1.09	1.62
A″	0.59	0.91	1.00	1.21	1.62
B	0.59	1.24	**0.66**	1.25	1.84
C	*0.95*	*1.42*	*1.52*	*1.78*	*2.18*
D	0.66	0.86	0.96	**0.85**	1.69
D′	0.83	**0.65**	0.85	1.05	**1.43**

[a]For each microstate in Figure 4 we calculated the average distance (in nm) from the native structure. The average distance is given in terms of the root mean square distance (RMSD) with different combinations of native α-helices (Figure 1): III–IV (α-helices III and IV), II–IV (α-helices II and IV), II–III (α-helices II and III), II–III–IV (α-helices II, III and IV), I–II–III–IV (α-helices I, II, III, and IV). For each combination of native α-helices the most native-like microstate is shown in bold, while the more unstructured one is shown in italics. The overall most unstructured microstate is C, while the overall most native-like microstate is D.

perturbation;[79] the rectangular boxes in Figure S10 represent the regions where the effect of the mutation was measurable.

With the same goal of probing transient long-range structure, the acid-denature state of ACBP has been studied by paramagnetic relaxation enhancement (PRE) methods, a technique that reports on the distances between residues.[52,54,83] Here, PRE intensities have been simulated using the standard r^{-6} averaging[114] and considering an uncertainty of 5 Å to take into account the absence of the spin labels in the structures of our ensemble, which were obtained for the wild-type protein. The comparison with the measured PRE intensities[52] is shown in Figure S11. While the chemical shift analysis suggested that α-helix IV is more stable and the other α-helices are weakly populated in the acid-denatured state, chemical shifts perturbation and PRE suggested that the region comprising α-helix I is not very much in contact with the rest of the polypeptide chain. This result is also in agreement with a Φ-value analysis that suggested the docking α-helices I and IV to be the rate-limiting step for the folding process.[115] Our ensemble, even if it is slightly more compact than expected, is in agreement with all these observations. Indeed, from the contact maps in Figure 5 as well as from the free energy landscapes in Figures 3 and 4, it is clear that α-helix IV is more populated than the other ones and also that the region comprising α-helix I form less contacts with the rest of the chain than any other regions (see also Figure S12 for the total number of contacts per residue). From these results we thus conclude that the ensemble that we determined here captures well both for the long and the short-range behavior of the protein.

Residual Structure in the Acid-Denatured State of ACBP and Its Relationship with the Folding Process. The ensemble of structures calculated in this work gives us the possibility to visualize at atomistic detail the structure and dynamics of a denatured state of a protein. An important question that we can address is whether there is a relationship between the transient structure in this denatured state and the folding process of this protein under physiological conditions. This question has been investigated for ACBP,[52,54,78,79,81–84,115,116] as well as for other proteins,[51,117–123] with indications that in many cases the transient structure is native-like. Using the ensemble of structures determined in this work we can investigate this relationship in a quantitative manner.

dx.doi.org/10.1021/ja5027584 | J. Am. Chem. Soc. 2014, 136, 8982–8991

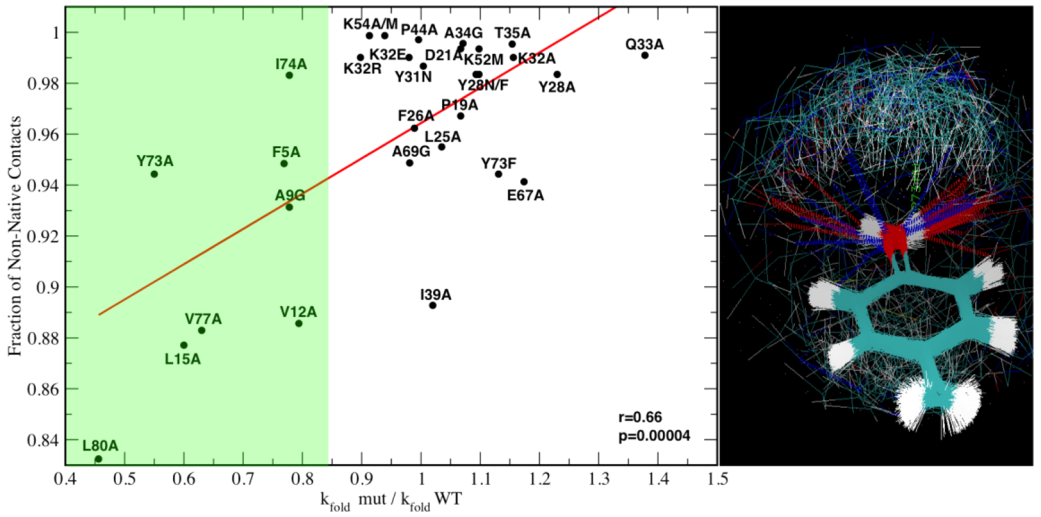


Figure 6. Relationship between the residual structure in the acid-denatured state and the folding process of ACBP. (a) Correlation between the fraction of non-native contacts formed in the acid-denatured state and the change in folding rates upon mutation. All the residues belonging to the rate-limiting native-like state (RLNLS) slow down folding when mutated (green-shaded region). (b) Representation of the interactions of Y73 in the denatured state ensemble.

To achieve this goal, we analyzed the fraction of native contact as well as the fraction of non-native contacts with respect to the total number of contacts formed by each residue in the acid-denatured state, and we compared these fractions with the effects of single point mutations on the folding kinetics and thermodynamic stability of the native state.[115] The standard interpretation of the changes in folding rates upon conservative mutations of the type employed in the Φ-value analysis is based on the observation that the mutations can change the stability of the folded state (i.e., the free energy difference between the folded and the unfolded states) and the free energy barriers for folding and for unfolding, i.e., the free energy differences between the unfolded and transition states and between the native and transition states, respectively.[124]

We found a relationship between the change in the folding rate upon single-point mutation and the fraction of non-native side-chain contacts with respect to all side-chain contacts (not only those directly affected by the mutation) at the position of the mutation in the acid-denatured state of wild-type ACBP (Figure 6). These results are statistically significant, with a coefficient of correlation of 0.66. In order to understand this relationship we consider the indications that the residual structure in the acid-denatured state resembles that in the unfolded state under physiological conditions,[52,54] and the observation that the Φ-value analysis is based on the rapid denaturation (to probe unfolding rates) and renaturation (to probe folding rates) of a protein, and thus the folding process that is observed is related to both the unfolded state and the denatured state of a protein.[124] Given these observations, the results shown in Figure 6 suggest that a mutation that affects a position along the polypeptide chain in which there is a significant residual native structure in the acid-denatured state of wild-type ABCP can slow down the folding process. These results also suggest that a mutation that affects a position in which there is a large fraction of non-native contacts in the unfolded state can speed up the folding process.

As a control we verified that the change in unfolding rate upon mutation is not correlated with the residual structure of the denatured state (Figure S13a); the change in folding rate upon mutation is instead correlated with the residual structure of the denatured state (Figure S13b). These correlations are weaker if one compares, instead of the ratio of the non-native over the total number of contacts formed in the acid-denatured state, the fraction of side-chain native contacts with respect to the folded state (Figure S14), suggesting that it is the balance between native and non-native structures that matters in the folding process of ACBP.

Complementary to these results we found that the change in the stability upon mutation ($\Delta\Delta G_{U-F}$) has a weak negative correlation with the residual structure of the acid-denatured state (Figure 7a) suggesting that single point mutations that

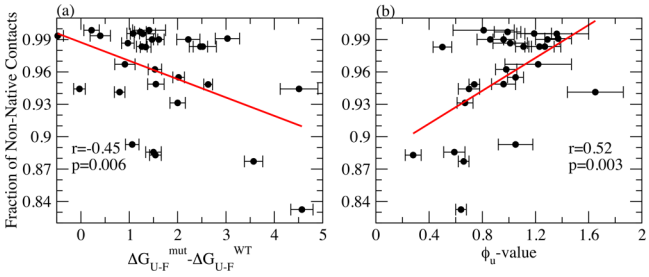


Figure 7. Relationship between the residual structure in the acid-denatured state and the free energy landscape of ACBP. (a) Correlation between the change in folding stability upon mutation (*x*-axis) and the residual structure of the acid-denatured state (*y*-axis). (b) Correlation between unfolding Φ-values (*x*-axis) and the residual structure of the acid-denatured state (*y*-axis). The mutants are the same as shown in Figure 6.

increase the fraction of non-native contacts at expense of the native ones can stabilize the denatured state instead of destabilize the native one. Furthermore, we observed a weak positive correlation with the unfolding Φ-values (Φ_u, Figure 7b). Again, by building on the observation that the residual structure in the acid-denatured state resembles that in the unfolded state under physiological conditions,[52,54] the results in Figure 7 are in agreement with the idea that residues that form native contacts in the unfolded state are those that guide the formation of the transition state, while residues that are strongly non-native in the unfolded state can conserve non-native

dx.doi.org/10.1021/ja5027584 | J. Am. Chem. Soc. 2014, 136, 8982−8991

contacts ($\Phi_u > 1$) also in the transition state. By contrast, we did not find a significant correlation between the changes in *m*-values upon mutation and the fraction of non-native contacts in the denatured state (Figure S15). This lack of correlation is mainly caused by a set of mutants with a very high ratio of non-native contacts (top left region in Figure S15), whose response to increasing concentration of denaturants (which is the effect reported by the *m*-values) does not involve a further increase in this ratio. Taken together, these results suggest the intriguing possibility that all the residual structure of the denatured state, both native and non-native, can play a role in the folding process.

Overall these findings confirm the notion that the residual structure present in the unfolded state, at least for fast-folding proteins, is a key feature in the folding process proteins.[51,117−123] Although the correlations that we reported (Figures 6, 7, S13, and S14) by themselves do not represent direct evidence of a causative effect, they are particularly interesting because the observables that we have analysed report on the whole folding process and so they are a function of both the denatured, the transition, and the folded states. These results are also consistent with the assumption underlying the Φ-value analysis that it is possible to introduce mutations that do not modify the unfolded state independently of the transition state.

Evolutionary Conservation of the Residues with Native-Like Residual Structure in the Acid-Denatured State. The eight residues (F5, A9, V12, L15, Y73, I74, V77, and L80) that were observed to slow down the folding process when mutated (Figure 6a, green shaded region) were suggested to be important in the formation of the rate-limiting native-like state (RLNLS).[115] Seven of these residues are evolutionary highly conserved, with I74 being the only non-conserved one. An analysis of the structures of the acid-denatured ensemble that we determined in this work revealed that these residues have a very high probability of forming native contacts, with I74 being the only one that departs from this trend (Figure 6). These results suggest that the native-like residual structure in the unfolded state of ACBP favors the folding process.

To investigate further this aspect we analyzed in more detail the interactions formed by Y73 in the acid-denatured state. It has been observed that while the Y73A mutation reduces the folding rate, the Y73F mutation increases it;[115] this latter mutation is also one of the two out of 30 mutations that were reported to increase the stability of the folded state.[115] These results suggested that while the phenyl group of Y73 is important to promote folding by forming a number of interactions in the RLNLS, the phenolic hydroxyl group of Y73 is instead inhibitory in the folding process.[115] To verify this suggestion we analyzed the presence of a hydrogen bond formed by the phenolic hydroxyl group of Y73 in the acid-denatured state finding a remarkable 10% population for this interaction, which is non-native. For comparison, the removal of the phenolic hydroxyl group of Y73 results in the value $k_f(\text{Y73F})/k_f(\text{WT}) = 1.13$,[115] i.e., a speed up in the folding of about 13%. Despite the negative effects of the phenolic hydroxyl group of Y73 on the folding process, the evolutionary conservation of this group could be explained by its role in protein function, as it forms a hydrogen bond with palmitoyl-coenzyme A in all the known structures of the complex (1ACA, 1NVL and 2CB8).

DISCUSSION

Denatured and unfolded states of proteins exhibit in many cases more transient structure than random coil chains.[51,117−123] This transient structure, in particular if native-like, is considered to be a fundamental feature in the folding process.[51,117−123] It has been shown that transient native-like contacts in the acid-denatured state of ACBP govern its compactness,[52,54,78,79,81−84,115,116] suggesting a minimally frustrated unfolded state, i.e., a state with a minimal amount of non-native interactions.[83] In particular it was observed that, by mutating residues that form transient native interactions in the denatured state, the overall size of the denatured protein increases.[83] Furthermore, computational studies suggest that, at least for fast folding proteins, only native contacts are relevant in determining the search for the native state.[118−120,125]

The results that we have presented provide a converged sampling of a denatured state of a globular protein determined using NMR chemical shift restraints (Table 1, Figures 2−4, S2, and S3), which gives us the possibility of analyzing quantitatively and extending earlier conclusions obtained from experimental studies. The acid-denatured state of ACBP is characterized by the presence of significant residual α-helical structure, in particular in the C-terminal α-helix IV, with some non-native β structure also being transiently present in the central region of the polypeptide chain (Table 1 and Figure S9). The ensemble of structures that we have determined exhibits a relatively compact and structured state of low-free energy and an unstructured state of high-free energy (Figures 3 and 4). The presence of both native and non-native tertiary contacts is in agreement with chemical shifts perturbation and PRE measurements (Figures S7−S11). In particular the region comprising α-helix I is on average more detached from the rest of the polypeptide chain than any other region (Figures 5 and S12). By studying in detail the role of the transient structure of the denatured state we have found a significant correlation between the fraction of non-native contacts and the variation of the folding rates upon single-point mutations (Figure 6). This relationship is also reflected in a weak correlation with the unfolding Φ-value and a weak negative correlation with the effect of mutation on the stability (Figure 7). Overall these observations suggest that the structure in the denatured state has a functional role in driving the folding process as perturbing non-native interactions in the denatured state can speed up the folding and increase its stability. Furthermore residues that show more transient structure in the denatured state show a higher degree of native or non-native structure also in the transition state.

CONCLUSIONS

We have presented an approach that enables converged free energy landscapes to be calculated for unstructured states of proteins. This approach is based on the combination of bias-exchange metadynamics to enhance the sampling of the conformational space with the use of ensemble-averaged chemical shift restraints to modify the force field to increase the consistency of the sampling with experimental measurements. Application of this approach to the acid-denatured state of ACBP has enabled us to characterize the relationship between the transient native and non-native structures in the unfolded state and the folding process.

dx.doi.org/10.1021/ja5027584 | J. Am. Chem. Soc. 2014, 136, 8982−8991

■ ASSOCIATED CONTENT

Supporting Information

Figures S1−S15. This material is available free of charge via the Internet at http://pubs.acs.org/.

■ AUTHOR INFORMATION

Corresponding Author

mv245@cam.ac.uk

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This study was supported by a Marie Curie IEF Fellowship (C.C.) and the BBSRC (M.V.).

■ REFERENCES

(1) Karplus, M.; Kuriyan, J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6679.

(2) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y. *Science* **2010**, *330*, 341.

(3) Best, R. B. *Curr. Opin. Struct. Biol.* **2012**, *22*, 52.

(4) Endres, N. F.; Das, R.; Smith, A. W.; Arkhipov, A.; Kovacs, E.; Huang, Y.; Pelton, J. G.; Shan, Y.; Shaw, D. E.; Wemmer, D. E.; Groves, J. T.; Kuriyan, J. *Cell* **2013**, *152*, 543.

(5) Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Shaw, D. E. *Science* **2011**, *334*, 517.

(6) Nygaard, R.; Zou, Y.; Dror, R. O.; Mildorf, T. J.; Arlow, D. H.; Manglik, A.; Pan, A. C.; Liu, C. W.; Fung, J. J.; Bokoch, M. P.; Thian, F. S.; Kobilka, T. S.; Shaw, D. E.; Mueller, L.; Prosser, R. S.; Kobilka, B. K. *Cell* **2013**, *152*, 532.

(7) Zhang, C.; Srinivasan, Y.; Arlow, D. H.; Fung, J. J.; Palmer, D.; Zheng, Y.; Green, H. F.; Pandey, A.; Dror, R. O.; Shaw, D. E.; Weis, W. I.; Coughlin, S. R.; Kobilka, B. K. *Nature* **2012**, *492*, 387.

(8) Becker, T.; Bhushan, S.; Jarasch, A.; Armache, J.-P.; Funes, S.; Jossinet, F.; Gumbart, J.; Mielke, T.; Berninghausen, O.; Schulten, K.; Westhof, E.; Gilmore, R.; Mandon, E. C.; Beckmann, R. *Science* **2009**, *326*, 1369.

(9) Zhao, G.; Perilla, J. R.; Yufenyuy, E. L.; Meng, X.; Chen, B.; Ning, J.; Ahn, J.; Gronenborn, A. M.; Schulten, K.; Aiken, C.; Zhang, P. *Nature* **2013**, *497*, 643.

(10) Lane, T. J.; Shukla, D.; Beauchamp, K. A.; Pande, V. S. *Curr. Opin. Struct. Biol.* **2013**, *23*, 58.

(11) Schlick, T.; Collepardo-Guevara, R.; Halvorsen, L. A.; Jung, S.; Xiao, X. *Q. Rev. Biophys.* **2011**, *44*, 191.

(12) Sutto, L.; Gervasio, F. L. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 10616.

(13) Sanbonmatsu, K. Y. *Curr. Opin. Struct. Biol.* **2012**, *22*, 168.

(14) Prigozhin, M.; Gruebele, M. *Phys. Chem. Chem. Phys.* **2013**, *15*, 3372.

(15) Dror, R. O.; Dirks, R. M.; Grossman, J.; Xu, H.; Shaw, D. E. *Annu. Rev. Biophys.* **2012**, *41*, 429.

(16) Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. *J. Chem. Theory Comput.* **2013**, *9*, 3878.

(17) Wang, K.; Chodera, J. D.; Yang, Y.; Shirts, M. R. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 989.

(18) Lindert, S.; Bucher, D.; Eastman, P.; Pande, V.; McCammon, J. A. *J. Chem. Theory Comput.* **2013**, *9*, 4684.

(19) Bottaro, S.; Lindorff-Larsen, K.; Best, R. B. *J. Chem. Theory Comput.* **2013**, *9*, 5641.

(20) Best, R. B.; Mittal, J.; Feig, M.; MacKerell, A. D., Jr. *Biophys. J.* **2012**, *103*, 1045.

(21) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D., Jr. *J. Chem. Theory Comput.* **2012**, *8*, 3257.

(22) Lindorff-Larsen, K.; Maragakis, P.; Piana, S.; Eastwood, M. P.; Dror, R. O.; Shaw, D. E. *PLoS One* **2012**, *7*, e32131.

(23) Best, R. B.; Hummer, G. *J. Phys. Chem. B* **2009**, *113*, 9004.

(24) Freddolino, P. L.; Harrison, C. B.; Liu, Y.; Schulten, K. *Nat. Phys.* **2010**, *6*, 751.

(25) Beauchamp, K. A.; Lin, Y.-S.; Das, R.; Pande, V. S. *J. Chem. Theory Comput.* **2012**, *8*, 1409.

(26) Cerutti, D. S.; Rice, J. E.; Swope, W. C.; Case, D. A. *J. Phys. Chem. B* **2013**, *117*, 2328.

(27) Vymětal, J. í.; Vondrášek, J. í. *J. Chem. Theory Comput.* **2012**, *9*, 441.

(28) Piana, S.; Lindorff-Larsen, K.; Shaw, D. E. *Biophys. J.* **2011**, *100*, L47.

(29) Lange, O. F.; Van der Spoel, D.; De Groot, B. L. *Biophys. J.* **2010**, *99*, 647.

(30) Li, D.-W.; Brüschweiler, R. *J. Chem. Theory Comput.* **2011**, *7*, 1773.

(31) Granata, D.; Camilloni, C.; Vendruscolo, M.; Laio, A. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 6817.

(32) Noé, F.; Schütte, C.; Vanden-Eijnden, E.; Reich, L.; Weikl, T. R. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 19011.

(33) Abrams, C.; Bussi, G. *Entropy* **2013**, *16*, 163.

(34) Maragliano, L.; Vanden-Eijnden, E. *Chem. Phys. Lett.* **2006**, *426*, 168.

(35) Bussi, G.; Laio, A.; Parrinello, M. *Phys. Rev. Lett.* **2006**, *96*, 090601.

(36) Laio, A.; Gervasio, F. L. *Rep. Prog. Phys.* **2008**, *71*, 126601.

(37) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562.

(38) Piana, S.; Laio, A. *J. Phys. Chem. B* **2007**, *111*, 4553.

(39) Sinko, W.; Miao, Y.; de Oliveira, C. A. F.; McCammon, J. A. *J. Phys. Chem. B* **2013**, *117*, 12759.

(40) Mitsutake, A.; Mori, Y.; Okamoto, Y. *Methods Mol. Biol.* **2013**, *924*, 153.

(41) Deighan, M.; Bonomi, M.; Pfaendtner, J. *J. Chem. Theory Comput.* **2012**, *8*, 2189.

(42) Best, R. B.; Vendruscolo, M. *J. Am. Chem. Soc.* **2004**, *126*, 8090.

(43) Bonvin, A.; Boelens, R.; Kaptein, R. *J. Biomol. NMR* **1994**, *4*, 143.

(44) Clore, G. M.; Schwieters, C. D. *Biochemistry* **2004**, *43*, 10678.

(45) Guerry, P.; Salmon, L.; Mollica, L.; Roldan, J. L. O.; Markwick, P.; van Nuland, N. A. J.; McCammon, J. A.; Blackledge, M. *Angew. Chem., Int. Ed.* **2013**, *52*, 3181.

(46) Huang, J. R.; Grzesiek, S. *J. Am. Chem. Soc.* **2010**, *132*, 694.

(47) Kessler, H.; Griesinger, C.; Lautz, J.; Muller, A.; van Gunsteren, W. F.; Berendsen, H. J. C. *J. Am. Chem. Soc.* **1988**, *110*, 3393.

(48) Krzeminski, M.; Fuentes, G.; Boelens, R.; Bonvin, A. *Proteins* **2009**, *74*, 895.

(49) Torda, A. E.; Scheek, R. M.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1989**, *157*, 289.

(50) Varadi, M.; Kosol, S.; Lebrun, P.; Valentini, E.; Blackledge, M.; Dunker, A. K.; Felli, I. C.; Forman-Kay, J. D.; Kriwacki, R. W.; Pierattelli, R. *Nucleic Acids Res.* **2014**, *42*, D326.

(51) Vendruscolo, M. *Curr. Opin. Struct. Biol.* **2007**, *17*, 15.

(52) Kristjansdottir, S.; Lindorff-Larsen, K.; Fieber, W.; Dobson, C. M.; Vendruscolo, M.; Poulsen, F. M. *J. Mol. Biol.* **2005**, *347*, 1053.

(53) Lindorff-Larsen, K.; Best, R. B.; DePristo, M. A.; Dobson, C. M.; Vendruscolo, M. *Nature* **2005**, *433*, 128.

(54) Lindorff-Larsen, K.; Kristjansdottir, S.; Teilum, K.; Fieber, W.; Dobson, C. M.; Poulsen, F. M.; Vendruscolo, M. *J. Am. Chem. Soc.* **2004**, *126*, 3291.

(55) Cavalli, A.; Camilloni, C.; Vendruscolo, M. *J. Chem. Phys.* **2013**, *138*, 094112.

(56) Pitera, J. W.; Chodera, J. D. *J. Chem. Theory Comput.* **2012**, *8*, 3445.

(57) Roux, B.; Weare, J. *J. Chem. Phys.* **2013**, *138*, 084107.

(58) Boomsma, W.; Ferkinghoff-Borg, J.; Lindorff-Larsen, K. *PLoS Comp. Biol.* **2014**, *10*, e1003406.

(59) Camilloni, C.; Cavalli, A.; Vendruscolo, M. *J. Chem. Theory Comput.* **2013**, *9*, 5610.

(60) Cavalli, A.; Salvatella, X.; Dobson, C. M.; Vendruscolo, M. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 9615.

(61) Shen, Y.; Lange, O.; Delaglio, F.; Rossi, P.; Aramini, J. M.; Liu, G.; Eletsky, A.; Wu, Y.; Singarapu, K. K.; Lemak, A.; Ignatchenko, A.; Arrowsmith, C. H.; Szyperski, T.; Montelione, G. T.; Baker, D.; Bax, A. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 4685.
(62) Wishart, D. S.; Arndt, D.; Berjanskii, M.; Tang, P.; Zhou, J.; Lin, G. *Nucleic Acids Res.* **2008**, *36*, W496.
(63) Rosato, A.; Aramini, J. M.; Arrowsmith, C.; Bagaria, A.; Baker, D.; Cavalli, A.; Doreleijers, J. F.; Eletsky, A.; Giachetti, A.; Guerry, P. *Structure* **2012**, *20*, 227.
(64) Camilloni, C.; Robustelli, P.; Simone, A. D.; Cavalli, A.; Vendruscolo, M. *J. Am. Chem. Soc.* **2012**, *134*, 3968.
(65) Robustelli, P.; Stafford, K. A.; Palmer, A. G., III *J. Am. Chem. Soc.* **2012**, *134*, 6365.
(66) Case, D. A. *Curr. Opin. Struct. Biol.* **2013**, *23*, 172.
(67) Robustelli, P.; Trbovic, N.; Friesner, R. A.; Palmer, A. G., III *J. Chem. Theory Comput.* **2013**, *9*, 5190.
(68) Kannan, A.; Camilloni, C.; Sahakyan, A. B.; Cavalli, A.; Vendruscolo, M. *J. Am. Chem. Soc.* **2014**, *136*, 2204.
(69) Pietrucci, F.; Mollica, L.; Blackledge, M. *J. Phys. Chem. Lett.* **2013**, *4*, 1943.
(70) 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*, 111.
(71) Korzhnev, D. M.; Religa, T. L.; Banachewicz, W.; Fersht, A. R.; Kay, L. E. *Science* **2010**, *329*, 1312.
(72) Neudecker, P.; Robustelli, P.; Cavalli, A.; Walsh, P.; Lundstroem, P.; Zarrine-Afsar, A.; Sharpe, S.; Vendruscolo, M.; Kay, L. E. *Science* **2012**, *336*, 362.
(73) Kjaergaard, M.; Poulsen, F. M. *Prog. Nucl. Magn. Reson. Spectrosc.* **2012**, *60*, 42.
(74) Marsh, J. A.; Forman-Kay, J. D. *Proteins* **2012**, *80*, 556.
(75) Jensen, M. R.; Ruigrok, R. W.; Blackledge, M. *Curr. Opin. Struct. Biol.* **2013**, *23*, 426.
(76) Marsh, J. A.; Singh, V. K.; Jia, Z.; Forman-Kay, J. D. *Protein Sci.* **2006**, *15*, 2795.
(77) Camilloni, C.; De Simone, A.; Vranken, W. F.; Vendruscolo, M. *Biochemistry* **2012**, *51*, 2224.
(78) Modig, K.; Jürgensen, V. W.; Lindorff-Larsen, K.; Fieber, W.; Bohr, H. G.; Poulsen, F. M. *FEBS Lett.* **2007**, *581*, 4965.
(79) Bruun, S. W.; Iešmantavičius, V.; Danielsson, J.; Poulsen, F. M. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 13306.
(80) Robustelli, P.; Kohlhoff, K.; Cavalli, A.; Vendruscolo, M. *Structure* **2010**, *18*, 923.
(81) Fieber, W.; Kragelund, B. B.; Meldal, M.; Poulsen, F. M. *Biochemistry* **2005**, *44*, 1375.
(82) Fieber, W.; Kristjansdottir, S.; Poulsen, F. M. *J. Mol. Biol.* **2004**, *339*, 1191.
(83) Ozenne, V.; Noel, J. K.; Heidarsson, P. O.; Brander, S.; Poulsen, F. M.; Jensen, M. R.; Kragelund, B. B.; Blackledge, M.; Danielsson, J. *J. Mol. Biol.* **2014**, *426*, 722.
(84) Thomsen, J. K.; Kragelund, B. B.; Teilum, K.; Knudsen, J.; Poulsen, F. M. *J. Mol. Biol.* **2002**, *318*, 805.
(85) Voelz, V. A.; Jäger, M.; Yao, S.; Chen, Y.; Zhu, L.; Waldauer, S. A.; Bowman, G. R.; Friedrichs, M.; Bakajin, O.; Lapidus, L. J. *J. Am. Chem. Soc.* **2012**, *134*, 12565.
(86) Heidarsson, P. O.; Valpapuram, I.; Camilloni, C.; Imparato, A.; Tiana, G.; Poulsen, F. M.; Kragelund, B. B.; Cecconi, C. *J. Am. Chem. Soc.* **2012**, *134*, 17068.
(87) Best, R. B.; Mittal, J. *J. Phys. Chem. B* **2010**, *114*, 14916.
(88) Abascal, J. L.; Vega, C. *J. Chem. Phys.* **2005**, *123*, 234505.
(89) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D. *Bioinformatics* **2013**, *29*, 845.
(90) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. *Comput. Phys. Commun.* **2014**, *185*, 604.
(91) Hess, B. *J. Chem. Theory Comput.* **2008**, *4*, 116.
(92) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
(93) Evans, D. J.; Holian, B. L. *J. Chem. Phys.* **1985**, *83*, 4069.
(94) Camilloni, C.; Cavalli, A.; Vendruscolo, M. *J. Phys. Chem. B* **2013**, *117*, 1838.
(95) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141.
(96) Kohlhoff, K. J.; Robustelli, P.; Cavalli, A.; Salvatella, X.; Vendruscolo, M. *J. Am. Chem. Soc.* **2009**, *131*, 13894.
(97) Pietrucci, F.; Laio, A. *J. Chem. Theory Comput.* **2009**, *5*, 2197.
(98) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A. *Comput. Phys. Commun.* **2009**, *180*, 1961.
(99) Crespo, Y.; Marinelli, F.; Pietrucci, F.; Laio, A. *Phys. Rev. E* **2010**, *81*, 055701.
(100) Barducci, A.; Bussi, G.; Parrinello, M. *Phys. Rev. Lett.* **2008**, *100*, 020603.
(101) Baftizadeh, F.; Cossio, P.; Pietrucci, F.; Laio, A. *Curr. Phys. Chem.* **2012**, *2*, 79.
(102) Marinelli, F.; Pietrucci, F.; Laio, A.; Piana, S. *PLoS Comp. Biol.* **2009**, *5*, e1000452.
(103) Biarnés, X.; Pietrucci, F.; Marinelli, F.; Laio, A. *Comput. Phys. Commun.* **2012**, *183*, 203.
(104) Lindorff-Larsen, K.; Trbovic, N.; Maragakis, P.; Piana, S.; Shaw, D. E. *J. Am. Chem. Soc.* **2012**, *134*, 3787.
(105) Freddolino, P. L.; Park, S.; Roux, B.; Schulten, K. *Biophys. J.* **2009**, *96*, 3772.
(106) Piana, S.; Klepeis, J. L.; Shaw, D. E. *Curr. Opin. Struct. Biol.* **2014**, *24*, 98.
(107) Ceriotti, M.; Tribello, G. A.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 13023.
(108) Tribello, G. A.; Ceriotti, M.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 5196.
(109) Ceriotti, M.; Tribello, G. A.; Parrinello, M. *J. Chem. Theory Comput.* **2013**, *9*, 1521.
(110) Shen, Y.; Bax, A. *J. Biomol. NMR* **2010**, *48*, 13.
(111) De Simone, A.; Cavalli, A.; Hsu, S.-T. D.; Vranken, W.; Vendruscolo, M. *J. Am. Chem. Soc.* **2009**, *131*, 16332.
(112) Frishman, D.; Argos, P. *Proteins* **1995**, *23*, 566.
(113) Berisio, R.; Loguercio, S.; De Simone, A.; Zagari, A.; Vitagliano, L. *Protein Pept. Lett.* **2006**, *13*, 847.
(114) Iwahara, J.; Schwieters, C. D.; Clore, G. M. *J. Am. Chem. Soc.* **2004**, *126*, 5879.
(115) Kragelund, B. B.; Osmark, P.; Neergaard, T. B.; Schiødt, J.; Kristiansen, K.; Knudsen, J.; Poulsen, F. M. *Nat. Struct. Mol. Biol.* **1999**, *6*, 594.
(116) Teilum, K.; Kragelund, B. B.; Poulsen, F. M. *J. Mol. Biol.* **2002**, *324*, 349.
(117) Dyson, H. J.; Wright, P. E. *Chem. Rev.* **2004**, *104*, 3607.
(118) Zagrovic, B.; Snow, C. D.; Khaliq, S.; Shirts, M. R.; Pande, V. S. *J. Mol. Biol.* **2002**, *323*, 153.
(119) Makarov, D. E.; Plaxco, K. W. *Protein Sci.* **2003**, *12*, 17.
(120) Voelz, V. A.; Singh, V. R.; Wedemeyer, W. J.; Lapidus, L. J.; Pande, V. S. *J. Am. Chem. Soc.* **2010**, *132*, 4702.
(121) Choy, W.-Y.; Forman-Kay, J. D. *J. Mol. Biol.* **2001**, *308*, 1011.
(122) Bernado, P.; Blanchard, L.; Timmins, P.; Marion, D.; Ruigrok, R. W.; Blackledge, M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17002.
(123) Marsh, J. A.; Forman-Kay, J. D. *J. Mol. Biol.* **2009**, *391*, 359.
(124) Fersht, A. R. *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*; W. H. Freeman: New York, 1999.
(125) Best, R. B.; Hummer, G.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 17874.