Supporting Information

Kachlishvili et al. 10.1073/pnas.1407837111

SI Native and Mirror-Image Structures of Protein A

The mirror-image and native conformations have identical secondary structures, the only difference between them being the location of the third (C-terminal) α-helix. Fig. S1 illustrates the native structure of the 10- to 55-residue fragment of protein A from the Protein Data Bank (PDB) (Fig. S1*A*) and the mirror image of the native structure generated with the UNRES force field (Fig. S1*B*). It should be noted that the residues of the fragment of protein A studied here (main text and *SI Text*) are numbered starting from 1, which corresponds to residue no. 10 in the 1BDD PDB structure.

SI Materials and Methods

United Residue Model. The united residue (UNRES) model of polypeptide chains (1–10) is illustrated in Fig. S2. A polypeptide chain is represented as a sequence of α-carbon (C^α) atoms linked by virtual $C^\alpha\cdots C^\alpha$ bonds with united peptide groups halfway between the neighboring C^αs and united side chains (SC), whose sizes depend on the nature of the amino acid residues, attached to the respective C^αs by virtual $C^\alpha\cdots$SC bonds. The effective energy is expressed by Eq. **S1** (9),

$$\begin{aligned}U = {} & w_{SC}\sum_{i<j}U_{SC_iSC_j} + w_{SCp}\sum_{i\neq j}U_{SC_ip_j} + w_{pp}f_2(T)\sum_{i<j-1}U_{p_ip_j}\\ & + w_{tor}f_2(T)\sum_i U_{tor}(\gamma_i) + w_{tord}f_3(T)\sum_i U_{tord}(\gamma_i,\gamma_{i+1})\\ & + w_b\sum_i U_b(\theta_i) + w_{rot}\sum_i U_{rot}\left(\alpha_{SC_i},\beta_{SC_i},\theta_i\right)\\ & + w_{bond}\sum_i U_{bond}(d_i) + \sum_{m=3}^{6} w_{corr}^{(m)} f_m(T) U_{corr}^{(m)} + w_{SS}\sum_i U_{SS;i}\end{aligned} \quad \textbf{[S1]}$$

with (9)

$$f_m(T) = \frac{\ln\left(e+e^{-1}\right)}{\ln\left\{\exp\left[(T/T_0)^{m-1}\right] + \exp\left[-(T/T_0)^{m-1}\right]\right\}};\quad T_0 = 300\ \text{K}, \quad \textbf{[S2]}$$

where the successive terms represent side chain–side chain, side chain–peptide, peptide–peptide, torsional, double-torsional, bond-angle bending (3), side-chain local (dependent on the angles α and β of Fig. S2), distortion of virtual bonds, and multibody (correlation) interactions and formation of disulfide bonds, respectively. The w's are the relative weights of each term. The side chain–side chain potentials include interactions with the solvent (2). The correlation terms arise from a cumulant expansion (4, 11) of the restricted free-energy function of the simplified chain obtained from the all-atom energy surface by integrating out the secondary degrees of freedom. The temperature-dependent factors of Eq. **S2**, introduced in our later work (9) and discussed further in ref. 12, reflect the fact that the UNRES effective energy is an approximate cumulant expansion of the restricted free energy. The virtual-bond vectors are the variables used in molecular dynamics.

$^{13}C^\alpha$ Chemical Shift. Although the details of this section have already been published (13), a brief description is provided for the convenience of the reader. For each amino acid μ, it is possible to define the differences Δ_μ^α between observed and predicted $^{13}C^\alpha$ chemical-shift values as

$$\Delta_\mu^\alpha = {}^{13}C^\alpha_{observed,\mu} - \frac{1}{\Omega}\sum_{i=1}^{\Omega} {}^{13}C^\alpha_{\mu,i}, \quad \textbf{[S3]}$$

where $^{13}C^\alpha_{\mu,i}$ is the chemical shift of residue μ in conformation i of Ω conformations. The average of the predicted chemical shifts over the Ω conformations is taken because proteins in solution exist as an ensemble of conformations.

The procedure for mapping the Δ_μ^α values onto a 3D protein model was formulated as follows. First, the Δ_μ^α value computed for each residue μ is smoothed by averaging it over the values of the two nearest-neighbor residues. Second, the resulting averaged $\langle\Delta_\mu^\alpha\rangle$ value is discretized according to the following rule:

$$\left\langle\Delta_\mu^\alpha\right\rangle_{\text{integer}} = \begin{cases} 1, & \left\langle\Delta_\mu^\alpha\right\rangle \le \sigma \ \text{(i.e., 1.70 ppm)} \\ 0, & \sigma < \left\langle\Delta_\mu^\alpha\right\rangle \le 2\sigma \ \text{(i.e., 3.40 ppm)} \\ -1, & \left\langle\Delta_\mu^\alpha\right\rangle > 2\sigma. \end{cases} \quad \textbf{[S4]}$$

The cutoff value of 1.7 ppm was obtained as explained in ref. 13. Third, the $\langle\Delta_\mu^\alpha\rangle_{\text{integer}}$ values, 1, 0, and −1 are mapped onto a 3D protein model and associated with a color, blue, white, and red, respectively. Blue-colored residues are those with small differences between the observed and computed chemical shifts, and white and red indicate medium and large differences, respectively.

Principal Component Analysis. A detailed description of the principal component analysis (PCA) method is available in our previous papers (10, 14–17) and in an earlier reference (18); therefore, only a brief outline of the approach is presented here. PCA, a covariance-matrix–based mathematical technique, is an effective method for extracting important motions from molecular dynamics (MD) simulations. In PCA, the Cartesian or internal coordinate space is rotated to a new space with new coordinates, principal components (PCs), a few of which are sufficient to describe a large part of the fluctuations of a protein. Here, structural fluctuations of the UNRES θ and γ angles [mean-square fluctuations (MSF)] can be decomposed into collective modes by PCA. The modes have "frequencies" and directions corresponding to the eigenvalues and eigenvectors of the covariance matrix. The modes k with the largest eigenvalues (λ^k) correspond to the modes that contribute the most to the structural fluctuations of the protein. The contribution of each angle (θ_i and γ_i) to mode k is called the influence, ν_i^k (17, 19, 20).

SI Energetic Competitiveness of Native and Mirror-Image Structures

To determine whether the mirror-image structure is energetically competitive with the native conformation, we have selected two MD trajectories at 270 K, as an example. Both of them start from the fully extended unfolded conformation; in one of them, protein A folds to its native conformation almost instantly from a fully unfolded conformation, and in the second trajectory, before jumping to the native state, protein A becomes trapped in a metastable state formed by a mirror-image conformation. Fig. S3, in which the C^α root-mean-square deviation (rmsd) from the native structure vs. energy of each snapshot for these two trajectories is plotted, shows that a difference between the lowest energies of the mirror-image and native conformations (indicated by two blue arrows in Fig. S3*B*) is only ∼2.5 kcal/mol; this

indicates that the mirror-image structure is energetically competitive with the native structure. This result is in almost quantitative agreement with the energy difference (1.5 kcal/mol), between the lowest energies of the native and mirror-image conformations, obtained from an atomically detailed folding simulation of protein A (21). Fig. S3*A* corresponds to the MD trajectory, in which protein A folds to its native conformation almost instantly from a fully unfolded conformation, and hence only one cluster is seen. There are two clusters in Fig. S3*B*, representing the MD trajectory in which protein A folds through the kinetic trap; the solid circles with 8.0 Å ≤ rmsd ≤ 10.0 Å correspond to the mirror-image topology, and the solid circles with rmsd ≤ 4.0 Å correspond to the native state.

SI Time Intervals for Free-Energy Profiles

Fig. S4 shows how the periods of time, over which the free-energy profiles (FEPs) were calculated, were selected for one of the pairs of MD trajectories (the same principle was used for the selection of time intervals for the other pairs of trajectories). Fig. S4*A* illustrates the rmsd as a function of time of the folding trajectory without a kinetic trap (black) and the folding trajectory with a kinetic trap (red). The first ∼29 ns (Fig. S4*B*) and then the first ∼700 ps (Fig. S4*C*) of both trajectories were selected and expanded. Moreover, the structures of protein A corresponding to significant changes in both trajectories are shown in Fig. S4 *B* and *C*. During the first ∼250 ps protein A remains unfolded in both trajectories (brown rectangle in Fig. S4*C*), although one of them (the trajectory without a kinetic trap) makes an attempt to fold to either its native or mirror-image topology within the first ∼100 ps (pink rectangle in Fig. S4*C*). Between ∼250 ps and ∼300 ps (brown and blue rectangles, respectively, in Fig. S4*C*), the system in both trajectories collapses, in one of them (the trajectory with a kinetic trap) to the mirror-image conformation, and remains in a metastable state for the next ∼27 ns (green rectangle in Fig. S4*B*), and, as shown in the representative structures, the mirror-image conformation converts to the native conformation at ∼28 ns; and in the second one (the trajectory without a kinetic trap) it first forms a molten-globule structure, as shown in the representative structure (red rectangle in Fig. S4*C*), and then (∼650 ps) jumps into the native state (illustrated by the representative structure in Fig. S4*C*).

SI FEPs Along θ_i and γ_i Angles of the Folding Trajectories with and Without Kinetic Traps

Fig. S5 shows the FEPs along all of the θ_i and γ_i angles computed at different times from the whole folding trajectory without a kinetic trap (Fig. S5 *A* and *B*, respectively) and the folding trajectory with a kinetic trap (Fig. S5 *C* and *D*, respectively). As in Fig. S4, the blue, red, green, and black curves correspond to the FEPs calculated over 300 ps, 650 ps, and 27 ns and over the entire duration of the trajectories, respectively. Because protein A jumps almost instantly (Fig. S4) into either its metastable state formed by the mirror-image conformation or its native state, the FEPs calculated over 100 ps and 250 ps (the pink and brown rectangles, respectively) are very shallow and overlap each other, and for most of the angles they do not contain extra information, which was not revealed by FEPs over 300 ps; therefore, the FEPs over these time intervals are omitted in Fig. S5. However, taking into account the importance of the initial time intervals in formation of the mirror-image topology, the important differences observed between the FEPs over 100 ps and 250 ps along a few angles are discussed in the main text.

By comparing the FEPs along the θ and γ angles of the folding trajectory without a kinetic trap (Fig. S5 *A* and *B*) to the FEPs along the θ and γ angles of the folding trajectory with a kinetic trap (Fig. S5 *C* and *D*), we found that 17 θ angles and 17 γ angles differ noticeably from each other and consequently play a crucial role in a folding pathway. In particular, of these 17 θ angles, 4 (θ_i, i = 13, 14, 30, 31) are from loops (blue numbers), 7 (θ_i, i = 8, 17, 34, 35, 41–43) represent helices (red numbers), and 6 (θ_i, i = 10, 11, 15, 16, 32, 33) represent edges, which include residues from both loop and α-helix (green numbers). Of these 17 γ angles, 3 (γ_i, i = 12, 13, 30) are from loops, 7 (γ_i, i = 7, 17, 20, 21, 26, 37, 40) are from helices, and 7 (γ_i, i = 9, 10, 14–16, 28, 31) are from edges. The plots in Fig. S5 show that the discrepancies between the FEPs with and without a mirror image are caused by different behaviors of the angles in the different time intervals. In particular, the θ angles (Fig. S5*A*) of the first loop (θ_i, i = 13, 14) and the edge (θ_i, i = 15), for the trajectory in which protein A folds without a kinetic trap, completely explore the shallow local minimum or "shoulder" in the region of ∼110°–140° before the protein jumps to the native state, whereas the same θ angles of the trajectory in which protein A folds through a kinetic trap jump back and forth between the local and global minima during the entire trajectory (Fig. S5*C*). The θ angles pertaining to the edge (θ_i, i = 10, 11) of both trajectories, plotted in Fig. S5 *A* and *C*, have similar FEPs for the entire trajectory. However, the θ_{10} angle in the trajectory with a kinetic trap explores mainly the local minimum or a shoulder at ∼120° while the whole system is in a kinetic trap, and after the whole system jumps to the native state, the θ_{10} angle starts exploring the global minimum. The same θ_{10} angle of the trajectory without a kinetic trap explores the region of 100°–130° at the beginning of the trajectory, while the whole system is in the nonnative state, and after the whole system jumps to the native state, the θ_{10} angle starts jumping back and forth between the local (shoulder) and global minimum. The differences between the FEPs of the θ_{11} angle of these trajectories are that the θ_{11} angle in the trajectory with a kinetic trap explores mainly the region around 120° while the whole system is in the kinetic trap; after the whole system jumps to the native state, the θ_{11} angle starts exploring global (∼110°) and local (∼90°) minima. The θ_{11} angle of the trajectory without a kinetic trap gradually explores its own global minimum (∼110°) during the entire trajectory.

The differences between the FEPs along the θ angles pertaining to the second loop (θ_i, i = 30, 31), the edge between the second loop and third α-helix (θ_i, i = 32, 33), and the third α-helix (θ_i, i = 34, 35, 41–43) of the trajectories, in which the protein folds with and without the kinetic trap, are that the θ_{30}, θ_{31}, θ_{32}, θ_{33}, θ_{34}, θ_{35}, θ_{41}, θ_{42}, and θ_{43} are more flexible at the beginning of the trajectory with a kinetic trap (Fig. S5*C*) either exploring only the region of 105°–140° (θ_{30}, θ_{31}) or jumping back and forth between the global minimum (90°) and the region of 100°–130° (θ_{32}, θ_{33}, θ_{34}, θ_{35}, θ_{41}, θ_{42}, θ_{43}), whereas the corresponding angles of the trajectory without a kinetic trap (Fig. S5*A*) gradually explore only their own global minima (90°) during the entire trajectory [except θ_{30}, which starts jumping back and forth between the global minimum (90°) and the region of 105°–140° after the protein reaches the native state]. The differences between the FEPs along the θ angles pertaining to the helices and edge (θ_i, i = 8, 16, 17) of these trajectories are quite similar. In particular, the θ_8, θ_{16} and θ_{17} angles of the trajectory with a kinetic trap are flexible only at the beginning of the trajectory (650 ps), exploring a large range of FEPs between 90° and 140°, whereas the same angles of the trajectory without a kinetic trap are flexible for ∼27 ns (θ_8) and even longer (∼45 ns) with (θ_{16}, θ_{17}) exploring the same regions of the FEPs. After these time intervals, all these θ angles of both trajectories jump to the global minimum (90°).

Based on the FEPs along the γ angles of the selected time periods and the entire trajectory plotted in Fig. S5 *B* and *D*, it can be observed that the γ angles of the first and the second loops and edges (γ_i, i = 10, 13–15, 28, 30, 31) for the trajectory with the kinetic trap (Fig. S5*D*) are more flexible and explore the regions of the FEPs (−70° to 30° for γ_{10}, −160° to −50° for γ_{13}, −180° to −50° for γ_{14}, −150° to −30° for γ_{15}, 60°–160° for

γ_{28}, −120° to −40° for γ_{30}, and −180° to −50° for γ_{31}) that were never explored by the same γ angles (in most of the cases) of the trajectory without a kinetic trap (Fig. S5*B*). A similar situation, but vice versa, was observed for γ_{16} pertaining to the edge. The differences between the rest of the FEPs along the γ angles, pertaining to the first loop and edge (γ_i, i = 9, 12), are less prominent and appear before the protein jumps to the native state. The differences between the FEPs along the γ angles, pertaining to the helices of these trajectories, are noticeable and appear either before the collapse (γ_i, i = 26, 37, 40) or after the collapse (γ_i, i = 7, 17, 20, 21). It is important to mention that, as for the θ angles, the main discrepancies between the FEPs along the γ angles (γ_{30}, γ_{31}) pertaining to the second loop and edge of the second loop and the third α-helix appear at the beginning of the trajectory before the protein jumps into either the native or the mirror-image state (for γ_{30}, the difference can also be observed after the protein remains in the mirror-image metastable state), whereas the differences between the FEPs pertaining to the first loop and edges are related to the time intervals after the collapse.

SI $^{13}C^{\alpha}$ Chemical Shift Analysis

Fig. S6 illustrates the differences between experimental $^{13}C^{\alpha}$ chemical shifts and those calculated from the MD trajectories. Fig. S6 *A*, *C*, and *E* and *B*, *D*, and *F* illustrate the $^{13}C^{\alpha}$ chemical shifts of the trajectories that fold with and without a mirror image, respectively. The time intervals used in the calculations of the $^{13}C^{\alpha}$ chemical shifts correspond to those used in the analysis of the FEPs. In particular, the results illustrated in Fig. S6 *A* and *B* are calculated before the collapse of the protein [brown rectangle (~250 ps) in Fig. S4*C*]. The results illustrated in Fig. S6 *D* and *E* are calculated when the protein remains in the metastable mirror-image state [green rectangle (~28 ns) in Fig. S4*B*], and the results illustrated in Fig. S6 *G* and *H* are calculated for the full trajectory (Fig. S4*A*).

Noticeable differences between the $^{13}C^{\alpha}$ chemical shifts calculated during the time interval when the protein remains in the metastable mirror-image state (Fig. S6 *D* and *E*) and the full trajectory (Fig. S6 *G* and *H*) are in the region of the first loop (residues 11–15). For example, in the full trajectory, the percentage of blue bars for the trajectory that folds without mirror image is >45% (Fig. S6*H*), whereas the percentage of blue bars for the trajectory that folds with a mirror image is ~35% (Fig. S6*G*). The reason for these differences is that the first loop does not reach the native geometry in the metastable mirror-image state and, to emerge from the mirror-image to the native conformation, the structure of the first loop undergoes drastic changes during this time; consequently, the differences between the $^{13}C^{\alpha}$ chemical shifts of the residues of the first loop and those from the experimental chemical shifts are larger.

The graphical representations of these differences are shown in Fig. S6 *C*, *F*, and *I* for 250-ps, 28-ns time intervals, and the full trajectory, respectively. In particular, each color bar of each residue in Fig. S6 *C*, *F*, and *I* is a difference between the corresponding color bars of the same residue in Fig. S6 *A* and *B*, *D* and *E*, and *G* and *H*, respectively. If the heights of the bars in Fig. S6 *A*, *D*, and *G* (trajectory with mirror image) are higher than the heights of the corresponding bars in Fig. S6 *B*, *E*, and *H* (trajectory without mirror image), the differences between them, illustrated in Fig. S6 *C*, *F*, and *I*, are represented by the bars with positive value and vice versa for the negative values. As was expected, the negative tall blue bars in Fig. S6*C* appear in the regions of the second loop and part of the third helix and in Fig. S6 *F* and *I* in the region of the first loop.

We also calculated the $^{13}C^{\alpha}$ chemical shifts within the exact time interval when these changes occur and, as expected, the heights of the blue bars in the region of the first loop were lower (~20%) for the mirror-image trajectory.

1. Liwo A, Pincus MR, Wawak RJ, Rackovsky S, Scheraga HA (1993) Prediction of protein conformation on the basis of a search for compact structures: Test on avian pancreatic polypeptide. *Protein Sci* 2(10):1715–1731.
2. Liwo A, et al. (1997) A united-residue force field for off-lattice protein-structure simulations. I. Functional forms and parameters of long-range side-chain interaction potentials from protein crystal data. *J Comput Chem* 18:849–873.
3. Liwo A, et al. (1997) A united-residue force field for off-lattice protein-structure simulations. II: Parameterization of local interactions and determination of the weights of energy terms by Z-score optimization. *J Comput Chem* 18:874–887.
4. Liwo A, Czaplewski C, Pillardy J, Scheraga HA (2001) Cumulant-based expressions for the multibody terms for the correlation between local and electrostatic interactions in the united-residue force field. *J Chem Phys* 115:2323–2347.
5. Liwo A, Ołdziej S, Czaplewski C, Kozlowska U, Scheraga HA (2004) Parametrization of backbone-electrostatic and multibody contributions to the UNRES force field for protein-structure prediction from ab initio energy surfaces of model systems. *J Phys Chem B* 108:9421–9438.
6. Ołdziej S, Liwo A, Czaplewski C, Pillardy J, Scheraga HA (2004) Optimization of the UNRES force field by hierarchical design of the potential-energy landscape. 2. Off-lattice tests of the method with single proteins. *J Phys Chem B* 108:16934–16949.
7. Ołdziej S, et al. (2004) Optimization of the UNRES force field by hierarchical design of the potential-energy landscape. 3. Use of many proteins in optimization. *J Phys Chem B* 108:16950–16959.
8. Scheraga HA, et al. (2004) The protein folding problem: Global optimization of the force fields. *Front Biosci* 9:3296–3323.
9. Liwo A, et al. (2007) Modification and optimization of the united-residue (UNRES) potential energy function for canonical simulations. I. Temperature dependence of the effective energy function and tests of the optimization method with single training proteins. *J Phys Chem B* 111(1):260–285.
10. Maisuradze GG, Senet P, Czaplewski C, Liwo A, Scheraga HA (2010) Investigation of protein folding by coarse-grained molecular dynamics with the UNRES force field. *J Phys Chem A* 114(13):4471–4485.
11. Kubo RJ (1962) Generalized cumulant expansion method. *Phys Soc Japan* 17:1100–1120.
12. Shen H, Liwo A, Scheraga HA (2009) An improved functional form for the temperature scaling factors of the components of the mesoscopic UNRES force field for simulations of protein structure and dynamics. *J Phys Chem B* 113(25):8738–8744.
13. Martin OA, Vila JA, Scheraga HA (2012) *Che*Shift-2: Graphic validation of protein structures. *Bioinformatics* 28(11):1538–1539.
14. Maisuradze GG, Liwo A, Scheraga HA (2009) Principal component analysis for protein folding dynamics. *J Mol Biol* 385(1):312–329.
15. Maisuradze GG, Liwo A, Scheraga HA (2009) How adequate are one- and two-dimensional free energy landscapes for protein folding dynamics? *Phys Rev Lett* 102(23): 238102.
16. Maisuradze GG, Liwo A, Scheraga HA (2010) Relation between free energy landscapes of proteins and dynamics. *J Chem Theory Comput* 6(2):583–595.
17. Maisuradze GG, Liwo A, Senet P, Scheraga HA (2013) Local vs global motions in protein folding. *J Chem Theory Comput* 9(7):2907–2921.
18. Jolliffe IT (2002) *Principal Component Analysis* (Springer, New York).
19. Altis A, Nguyen PH, Hegger R, Stock G (2007) Dihedral angle principal component analysis of molecular dynamics simulations. *J Chem Phys* 126(24):244111.
20. Cote Y, Senet P, Delarue P, Maisuradze GG, Scheraga HA (2012) Anomalous diffusion and dynamical correlation between the side chains and the main chain of proteins in their native state. *Proc Natl Acad Sci USA* 109(26):10346–10351.
21. Vila JA, Ripoll DR, Scheraga HA (2003) Atomically detailed folding simulation of the B domain of staphylococcal protein A from random structures. *Proc Natl Acad Sci USA* 100(25):14812–14816.

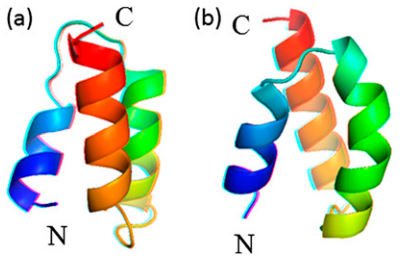

Fig. S1. The native (*A*) and mirror-image (*B*) conformations of protein A.

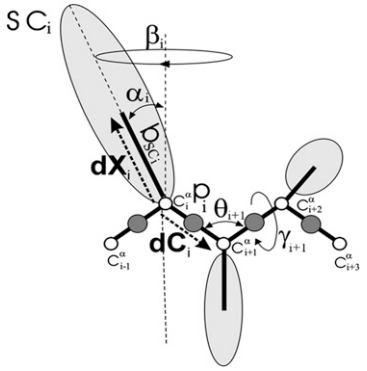

Fig. S2. The UNRES model of polypeptide chains. The interaction sites are peptide-bond centers (p), and side-chain ellipsoids of different sizes (SC) are attached to the corresponding α-carbons with different "bond lengths," b_{SC}. The α-carbon atoms are represented by small open circles. The equilibrium distance of the $C^\alpha \cdots C^\alpha$ virtual bonds is taken as 3.8 Å, which corresponds to planar *trans*-peptide groups. The geometry of the chain can be described either by the virtual-bond vectors $\mathbf{dC}_i$ ($C^\alpha_i \ldots C^\alpha_{i+1}$), $i = 1, 2, \ldots, N-1$ and $\mathbf{dX}_i$ ($C^\alpha_i \ldots SC_i$), $i = 2, 3, \ldots, N-1$ represented by thick dashed arrows, where N is the number of residues, or in terms of virtual-bond lengths, backbone virtual-bond angles θ_i, $i = 2, 3, \ldots, N-1$, backbone virtual-bond-dihedral angles γ_i, $i = 2, 3, \ldots, N-2$, and the angles α_i and β_i, $i = 2, 3, \ldots, N-1$ that describe the location of a side chain with respect to the coordinate frame defined by C^α_{i-1}, C^α_i, and C^α_{i+1}.

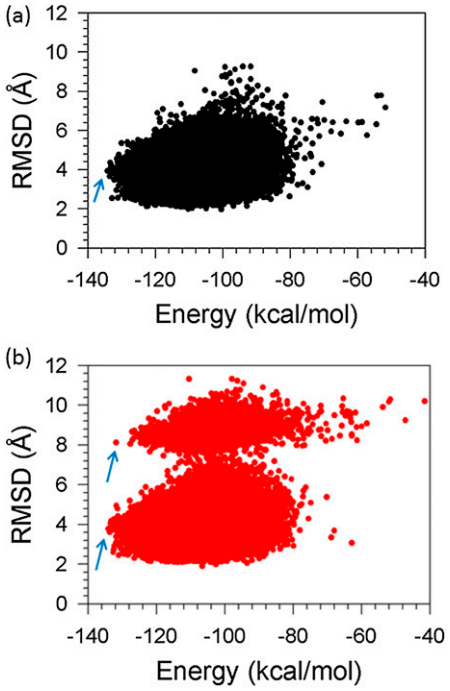

Fig. S3. C^α rmsd vs. energy for the two MD trajectories of protein A at 270 K. The black (*A*) and red (*B*) solid circles correspond to the trajectories in which protein A folds without and with a kinetic trap, respectively. Blue arrows indicate the conformations with the lowest energy in each cluster. In *B*, the top cluster corresponds to mirror-image conformations, and the bottom cluster corresponds to native conformations.

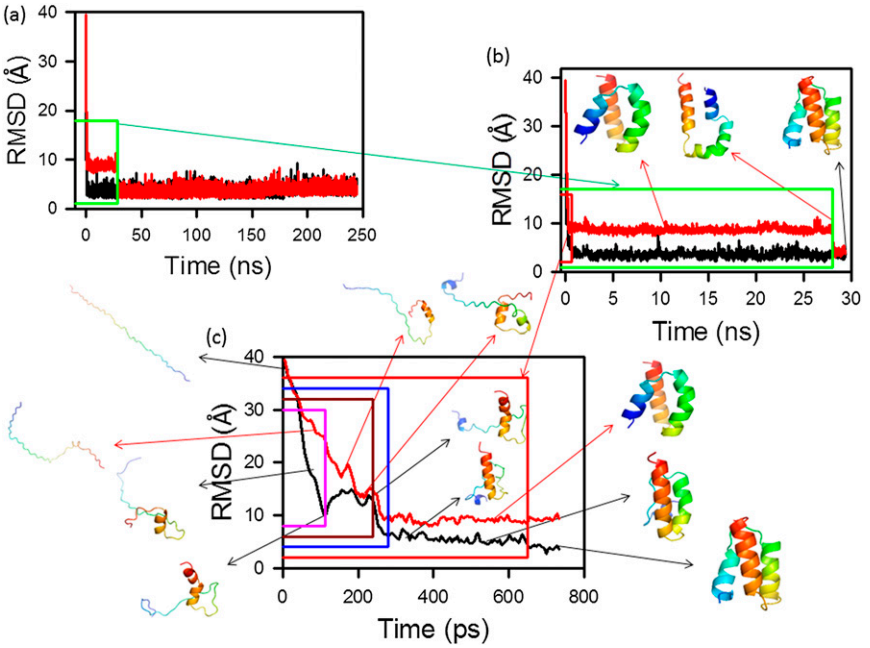

Fig. S4. (*A*) C^{α} rmsds vs. time for the MD folding trajectories without a kinetic trap (black) and with a kinetic trap (red). (*B* and *C*) C^{α} rmsds of the first 29-ns period (*B*) and the first 700-ps period (*C*) of both trajectories. The pink, brown, blue, and red rectangles in *C* correspond to ∼100-ps, ∼250-ps, ∼300-ps, and ∼650-ps time periods, respectively. The green rectangle in *B* corresponds to a 28-ns time interval. The structures of protein A in *B* and *C* correspond to significant changes in both trajectories.

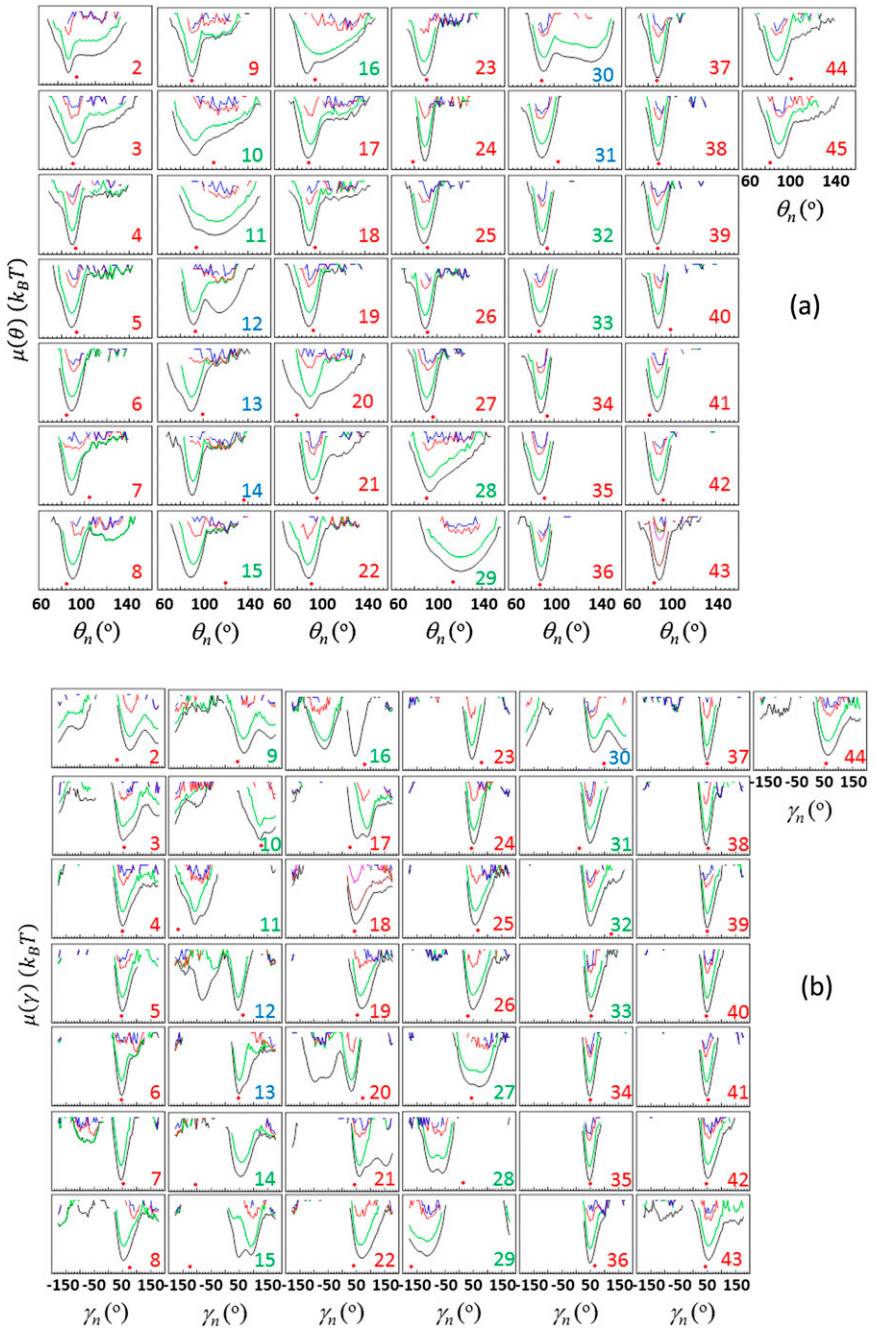

Fig. S5. (Continued)

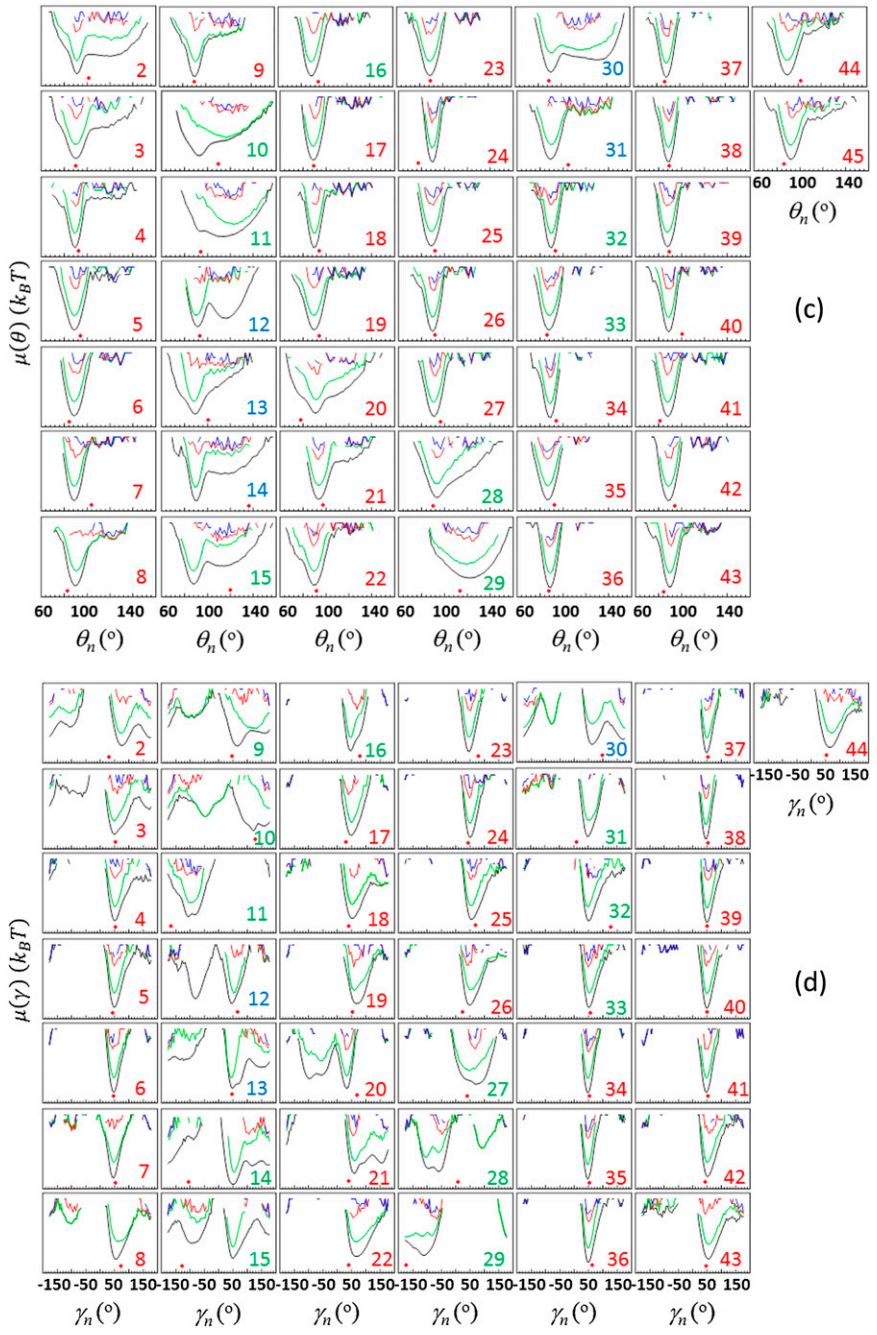

Fig. S5. (*A*–*D*) FEPs, $\mu(\theta)$ and $\mu(\gamma)$, along the θ and γ angles, respectively, for the folding trajectory without a kinetic trap (*A* and *B*, respectively) and the folding trajectory with a kinetic trap (*C* and *D*, respectively) of protein A. Blue, red, green, and black curves correspond to FEPs computed over 300 ps, 650 ps, and 28 ns and over the entire MD trajectories, respectively. The blue numbers pertain to FEPs along the θ and γ angles that include only residues of loops, the red numbers pertain to FEPs along the θ and γ angles that include only residues of α-helices, and the green numbers pertain to FEPs along θ and γ angles that include residues from edges containing both a loop and an α-helix. The NMR-derived structural data (small red solid circles at the bottom of each panel) are computed from the first model of the PDB ID code 1BDD (1).

1. Gouda H, et al. (1992) Three-dimensional solution structure of the B domain of staphylococcal protein A: Comparisons of the solution and crystal structures. *Biochemistry* 31(40):9665–9672.

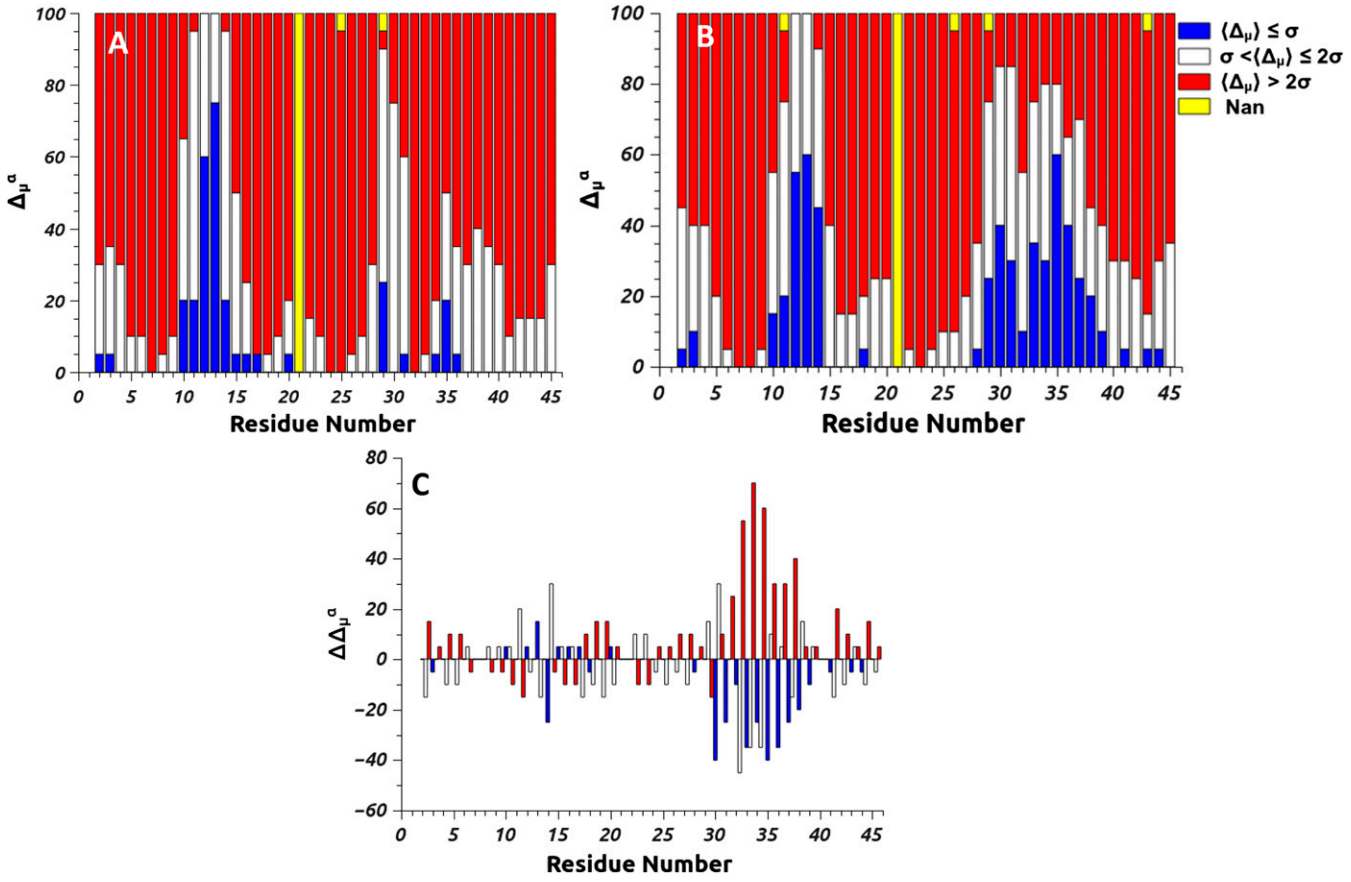

Fig. S6. (Continued)

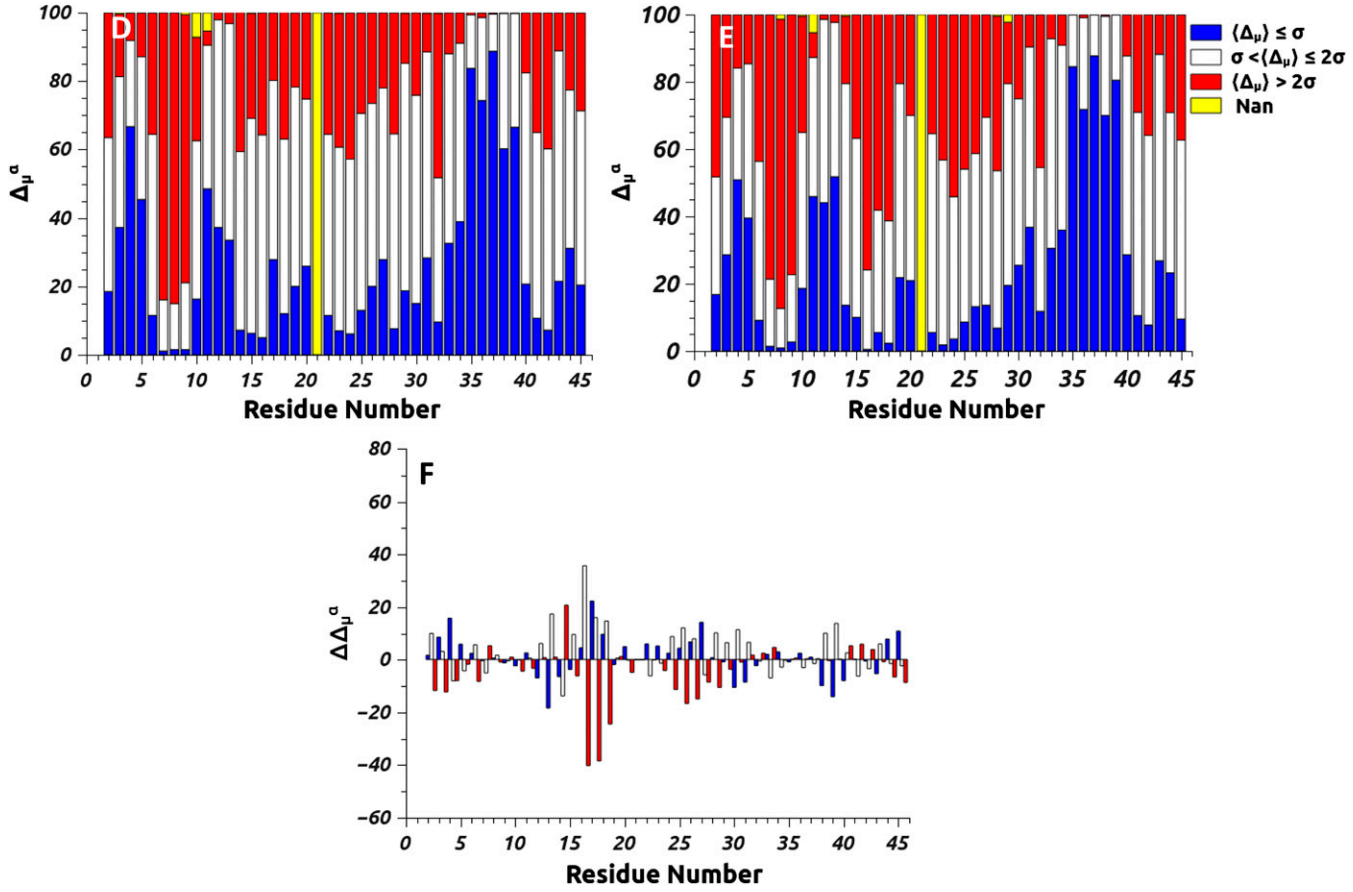

Fig. S6. (Continued)

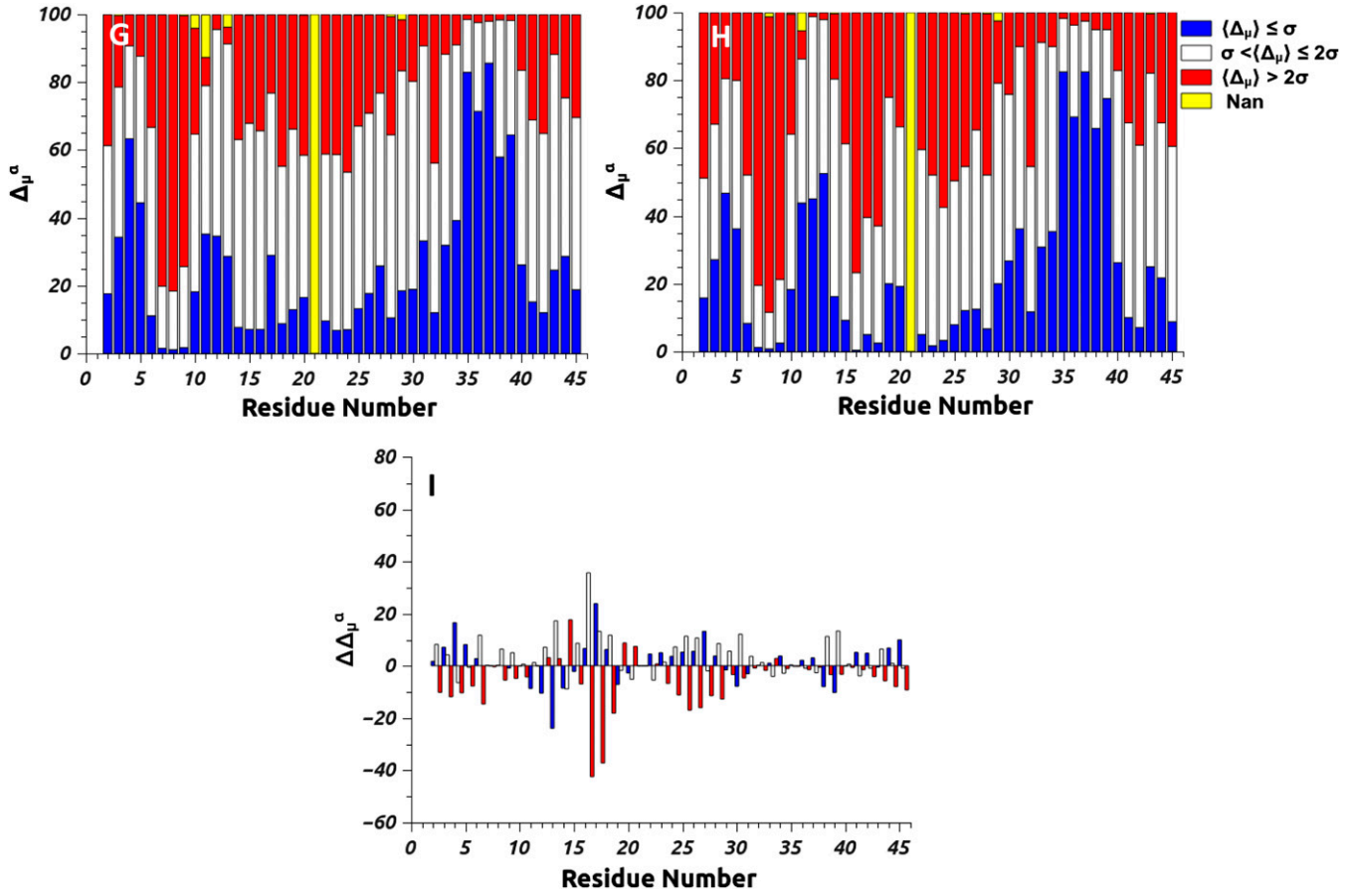

Fig. S6. $\Delta_\mu{}^\alpha$ represents the differences between the observed and theoretical $^{13}C^\alpha$ chemical shifts computed from the MD trajectories, and $\sigma = 1.7$ ppm. The different colors of the bars indicate the magnitude of the differences (Δ_μ) in terms of σ. A full bar highlighted in yellow (Nan, not a number), as for residue Gly21, indicates that the experimental chemical shift was not measured experimentally or that the theoretical value could not be computed for a particular conformation (for all of the remaining partial bars highlighted in yellow). *A*, *D*, and *G* and *B*, *E*, and *H* illustrate the differences in $^{13}C^\alpha$ chemical shift per residue for the trajectories that fold with and without mirror image, respectively. In particular, *A* and *B* correspond to the time interval between the start of the simulation and the collapse of the protein, *D* and *E* correspond to the time interval in which the protein remains at the metastable mirror-image state, and *G* and *H* correspond to the full trajectory of the simulation. *C*, *F*, and *I* illustrate the second-order differences, computed as follows: $\Delta\Delta^\alpha_{\mu(C)} = \Delta^\alpha_{\mu(A)} - \Delta^\alpha_{\mu(B)}$, $\Delta\Delta^\alpha_{\mu(F)} = \Delta^\alpha_{\mu(D)} - \Delta^\alpha_{\mu(E)}$, and $\Delta\Delta^\alpha_{\mu(I)} = \Delta^\alpha_{\mu(G)} - \Delta^\alpha_{\mu(H)}$.

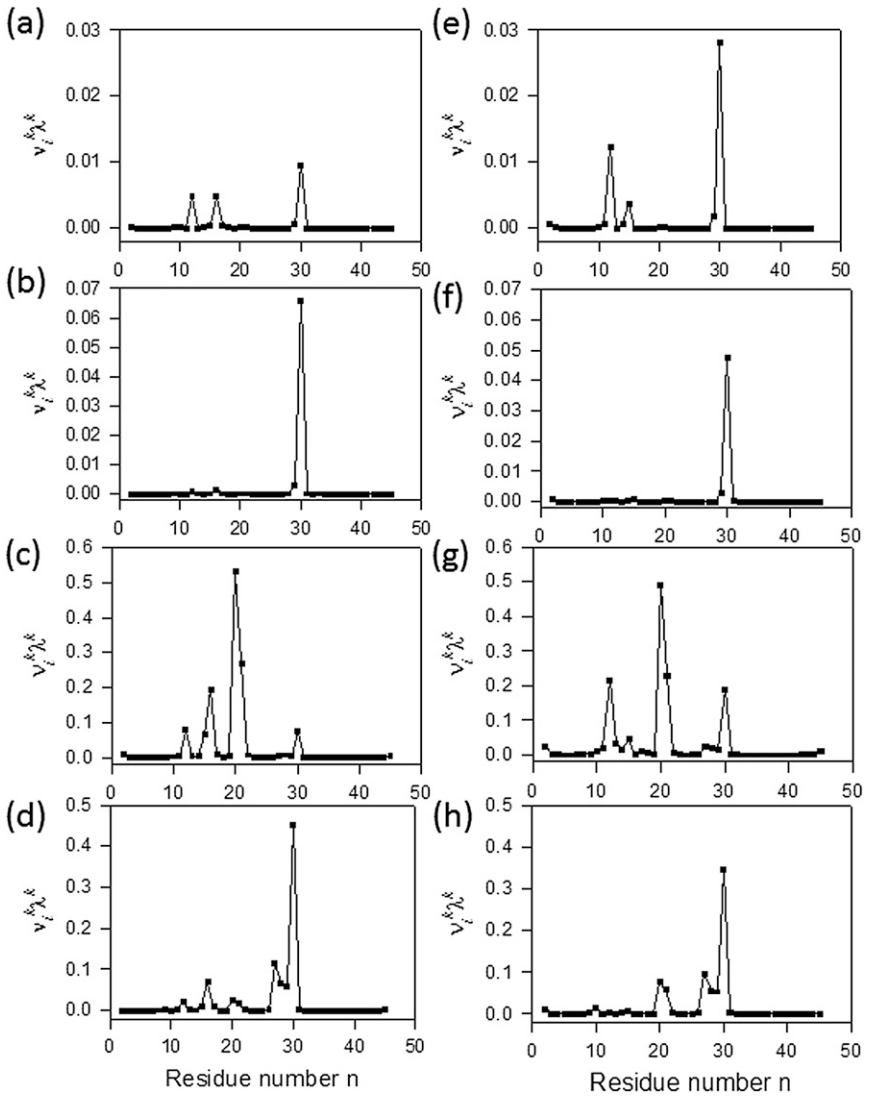

Fig. S7. (*A*–*H*) Contributions $(\nu_i^k \lambda^k)$ of principal mode 1 (*A*, *C*, *E*, and *G*) and mode 2 (*B*, *D*, *F*, and *H*) to the mean-square fluctuations along the θ (*A*, *B*, *E*, and *F*) and γ (*C*, *D*, *G*, and *H*) angles for the folding trajectory without a kinetic trap (*A*–*D*) and with a kinetic trap (*E*–*H*) of protein A.

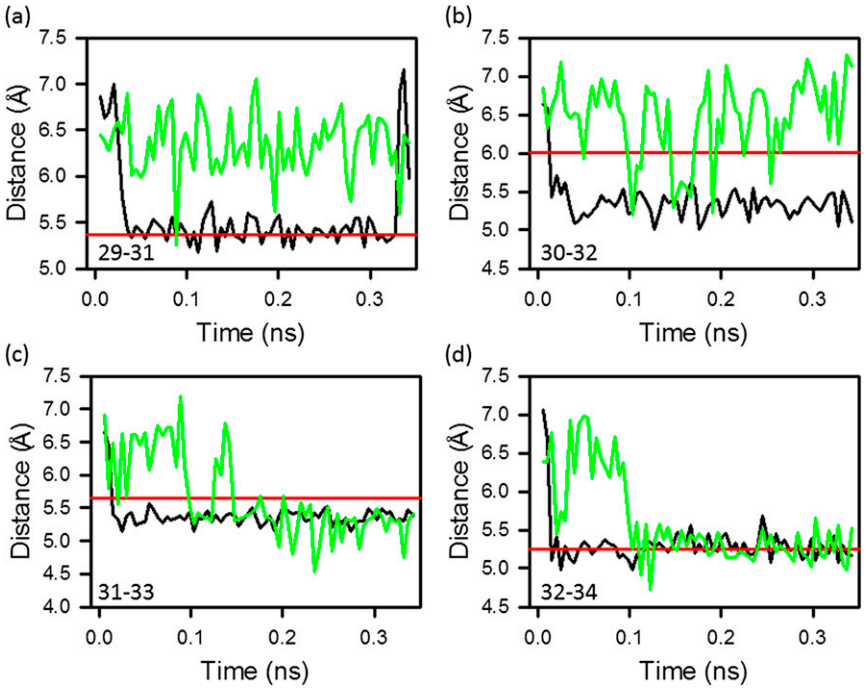

Fig. S8. (*A*–*D*) The distances between C^{α}s of Asp29–Ser31 (*A*), Pro30–Gln32 (*B*), Ser31–Ser33 (*C*), and Gln32–Ala34 (*D*) as function of time for the trajectory without (black line) and with (green line) a kinetic trap. Red line corresponds to experimental distance (1) between the C^{α}s of these selected residues.

1. Gouda H, et al. (1992) Three-dimensional solution structure of the B domain of staphylococcal protein A: Comparisons of the solution and crystal structures. *Biochemistry* 31(40):9665–9672.

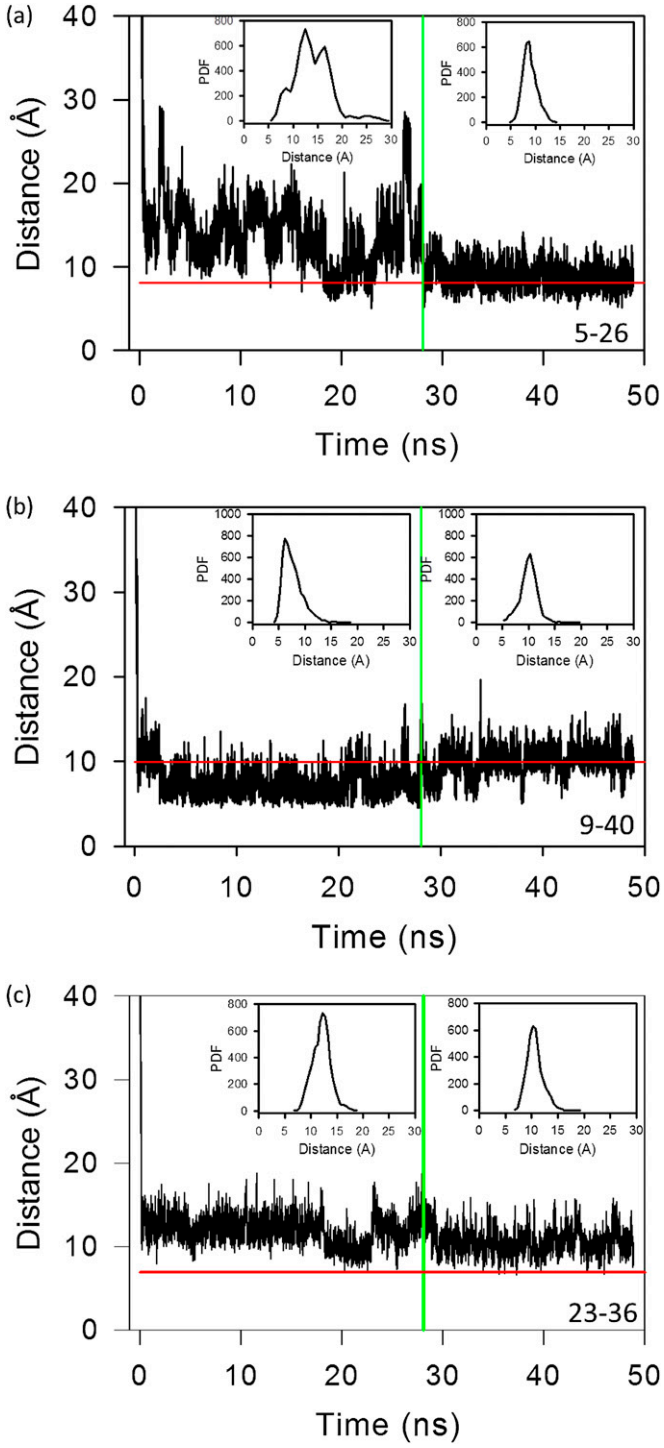

Fig. S9. (*A*–*C*) The distances between C^{α}s of Phe5–Leu26 (*A*), Leu9–Ala40 (*B*), and Ile23–Leu36 (*C*) as functions of time for the 50-ns time interval (the 0- to 28-ns time interval corresponds to the mirror-image state, and the 28- to 50-ns time interval corresponds to the native state). *Insets* show probability distribution functions of distances between these residues computed for both mirror image and native states. Red lines correspond to experimental distances (NMR) (1) between the C^{α}s of these selected residues. Green lines indicate the time when the protein jumps from the metastable mirror-image state into the native state.

1. Gouda H, et al. (1992) Three-dimensional solution structure of the B domain of staphylococcal protein A: Comparisons of the solution and crystal structures. *Biochemistry* 31(40):9665–9672.

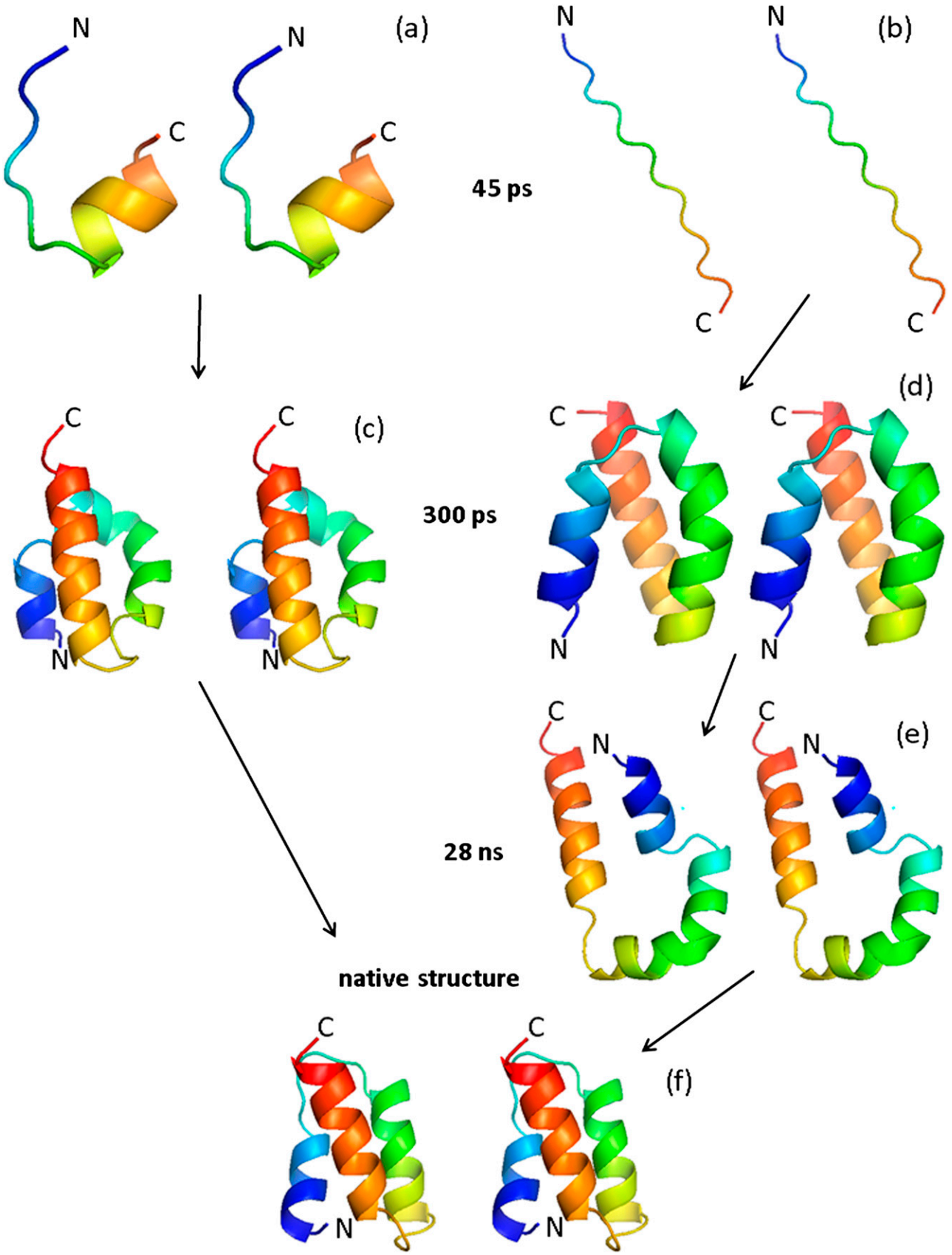

Fig. S10. (*A* and *B*) Stereo diagrams of the 24- to 37-residue portion of protein A at 45 ps of the trajectories without (*A*) and with (*B*) a kinetic trap. (*C* and *D*) At 300 ps the whole protein collapses and forms either a molten globule (*C*) or a mirror-image (*D*) conformation. (*E* and *F*) At 28 ns the protein emerges from a kinetic trap by opening the conformation of the first loop (*E*) and then proceeds to the native state (*F*) by adopting a closed-loop conformation.