JCTC Journal of Chemical Theory and Computation

Mechanism of Coupled Folding and Binding in the siRNA-PAZ Complex

Hai-Feng Chen*

College of Life Sciences and Biotechnology, Shanghai Jiaotong University, 800 Dongchuan Road, Shanghai 200240, China

Received January 29, 2008

Abstract: The PAZ domain plays a key role in gene silencing pathway. The PAZ domain binds with siRNAs to form the multimeric RNA-induced silencing complex (RISC). RISC identifies mRNAs homologous to the siRNAs and promotes their degradation. It was found that binding with siRNA significantly enhances apo-PAZ folding. However, the mechanism by which folding is coupled to binding is poorly understood. Thus, the coupling relationship between binding and folding is very important for understanding the function of gene silencing. We have performed molecular dynamics (MD) of both bound and apo-PAZ to study the coupling mechanism between binding and folding in the siRNA-PAZ complex. Room-temperature MD simulations suggest that both PAZ and siRNA become more rigid and stable upon siRNA binding. Kinetic analysis of high-temperature MD simulations shows that both bound and apo-PAZ unfold via a two-state process. The unfolding pathways are different between bound and apo-PAZ: the order of helix III and helices I & II unfolding is switched. Furthermore, transition probability was used to determine the transition state ensemble for both bound and apo-PAZ. It was found that the transition state of bound PAZ is more compact than that of apo-PAZ. The predicted Φ-values suggest that the Φ-values of helix III and sheets of β3-β7 for bound PAZ are more native-like than those of apo-PAZ upon the binding of siRNA. The results can help us to understand the mechanism of gene silencing.

Introduction

Short RNAs mediate gene silencing, a process associated with virus resistance, developmental control, and heterochromatin formation in eukaryotes.[1–6] At first, the RNase III-like enzyme Dicer cleaves dsRNA into 21−23-nucleotide small interfering RNAs (siRNAs).[7–10] Then, siRNAs bind with the PAZ domain to form the multimeric RNA-induced silencing complex (RISC).[11–13] Finally, RISC identifies mRNAs homologous to the siRNAs and promotes their degradation.[7,14] Therefore, the PAZ domain from the Argonaute 2 protein is a critical component of the gene silencing pathway.[8,9,15,16] Because of its importance in gene silencing, the PAZ-siRNA complex is a possible target to study gene silencing.

The crystal structure siRNA-PAZ complex was reported in 2004 (pdb code: 1SI2).[17] The complex has three α-helices and eight $\hat{\beta}$ strands. The PAZ domain consists of helix I from 228 to 237, helix II from 249 to 261, and helix III from 306 to 313. The PAZ domain in the complex adopts a heart-shaped globular topology, with a twisted β-barrel consisting of six β-strands (β1−β3, β6−β8), capped by two amino-terminal α-helices (H I and H II) on one side and connected to an αβ module (β4−β5-HIII) on the other side. The RNA is bound in the cleft between the β-sheet surface (β2, β5) and the β3, β4, H3 module of the PAZ domain (shown in Figure 1).

The NMR experiment indicates that the PAZ domain residues undergo chemical shift changes upon RNA binding.[18] The pocket of the PAZ domain elongates and consists of helix HIII and strand β4 along one face and strand β7 and loop β2−β3 along the opposite face. This suggests that siRNA binding induces significant conformational change

* Corresponding author phone: 0086-21-34204348; fax: 0086-21-34204348; e-mail: haifengchen@sjtu.edu.cn.

10.1021/ct800030s CCC: $40.75
Published on Web 07/02/2008

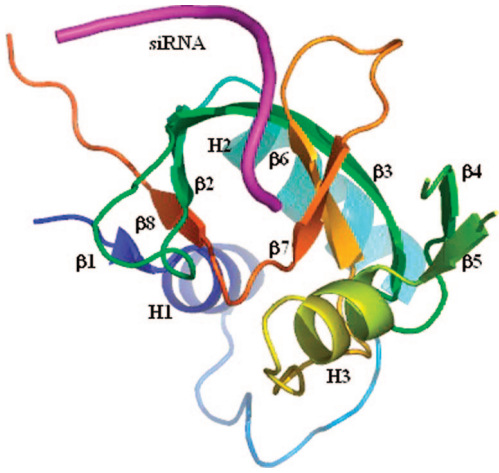

Figure 1. Ribbon representation of crystal structure of PAZ-siRNA (pdb code: 1SI2).[17] The locations of main secondary structures are indicated.

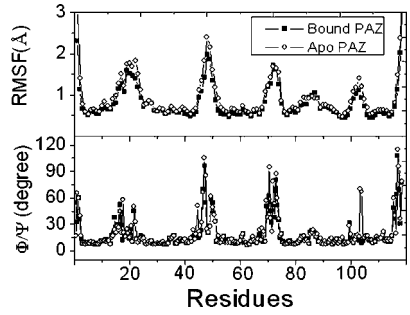

Figure 2. Cα and Ψ/Φ variations at the folded state for bound and apo-PAZ, respectively.

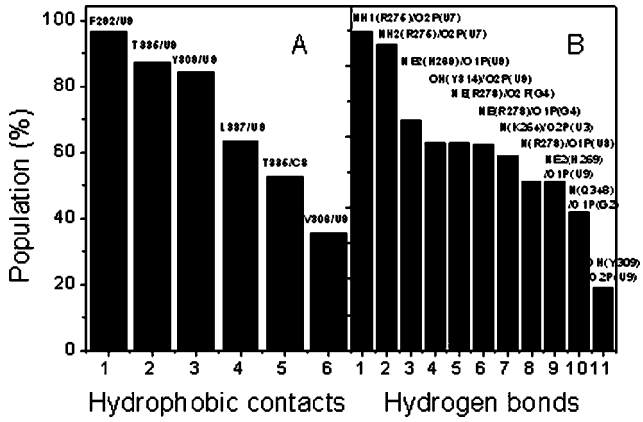

Figure 3. Hydrophobic contacts and hydrogen bonds between siRNA and PAZ. A: hydrophobic contacts and B: hydrogen bonds.

in the PAZ domain. These experimental observations raise a series of interesting questions. Some of those questions are as follows: (i) How does the folded state of PAZ change upon siRNA binding? (ii) What is the difference in the folding pathway between bound and apo-PAZ? (iii) If the binding of siRNA influences the unfolded state of PAZ? To shed light on these questions, we utilize molecular dynamics (MD) simulations in explicit solvent to analyze the coupling between binding and folding[19,20] in the siRNA-PAZ complex.

However, MD simulations are currently restricted to time scales of less than 1 μs, which is much shorter than folding

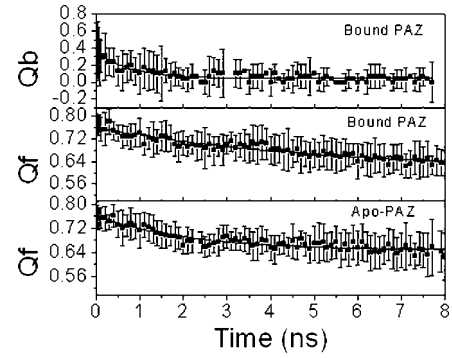

Figure 4. Kinetics fitting for bound and apo-PAZ.

Table 1. Unfolding Kinetics Constants[a]

		τ (ns)	A	B	R^2
bound PAZ	Qb	0.77	0.31	0.046	0.83
	Qf	5.26	0.15	0.61	0.91
	H (I & II)	9.70	0.28	0.51	0.72
	H (III)	12.93	0.094	0.84	0.70
Apo PAZ	Qf	2.09	0.11	0.65	0.89
	H (I & II)	8.74	0.34	0.44	0.72
	H (III)	6.97	0.35	0.63	0.67

[a] All curves are fitted by A exp($-$t/τ) + B.

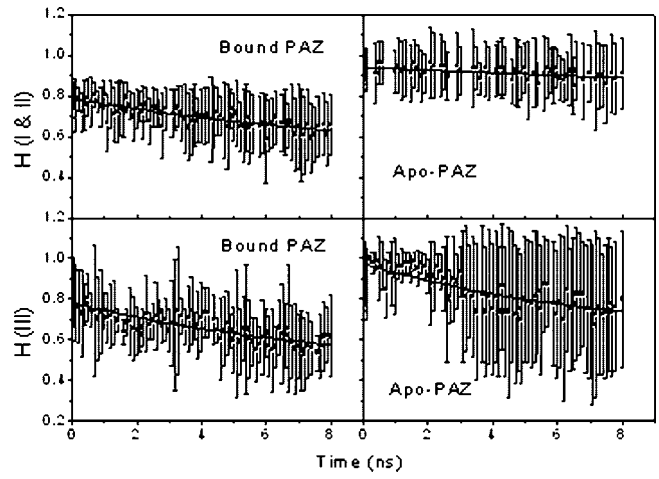

Figure 5. The unfolding kinetics of two group helices.

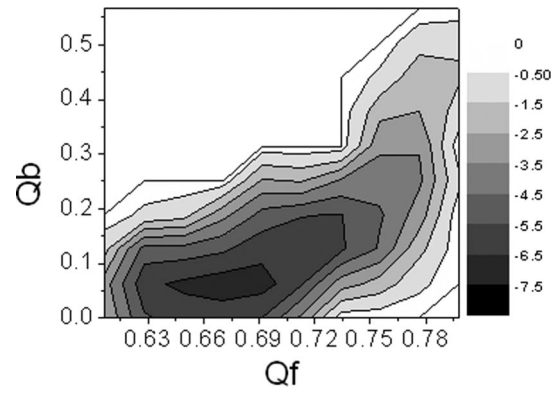

Figure 6. Unfolding landscapes with respect to Qb and Qf for bound PAZ.

and unfolding half-times of most proteins (at least 1 ms).[21,22] As we known, the rate of unfolding increases at high temperature, so most proteins unfold in the ns time scale at 498 K.[21] Therefore, simulations of protein unfolding at high

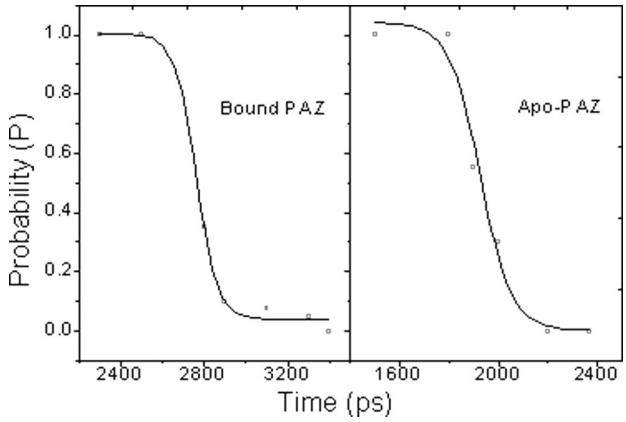

Figure 7. A representative transition probability *P* calculated at 498 K for the F ⇔ U transition for snapshot in the transition region for one of trajectories for bound and apo-PAZ, respectively. The red line is the fit to $P = 1/\{1 + \exp[(\tau - \tau_{TS})/\tau_{trans}]\}$.

temperature have been widely used[23–34] to study protein folding within a reasonable time. Furthermore, experiment and simulation confirm that the transition state for folding and unfolding is expected to be the same from the principle of microscopic reversibility.[21,35] Consequently, unfolding simulations at high temperature have been used in the current study.

In the following, we will discuss the folded states, unfolding kinetics, unfolding landscapes, transition states, prediction of Φ-values, unfolded states, and possible folding pathway for both bound and apo-PAZ to understand the influence of siRNA binding in the folding of PAZ.

Materials and Method

Room-Temperature and High-Temperature Molecular Dynamics Simulations. The atomic coordinates of the PAZ-siRNA complex were obtained from the crystal structure (pdb code: 1SI2).[17] Hydrogen atoms were added using the LEAP module of AMBER8.[36] Counterions were used to maintain system neutrality. All systems were solvated in a truncated octahedron box of TIP3P[37] waters with a buffer of 8 Å. The dimensions of the water box were 61.4 × 61.4 × 61.4 Å for bound PAZ. Particle Mesh Ewald (PME)[38] was employed to treat long-range electrostatic interactions with the default setting in AMBER8.[36] The parm99 force field was used for intramolecular interactions.[39,40] The SHAKE algorithm[41] was used to constrain bonds involving hydrogen atoms. 1000-step steepest descent minimization was performed to relieve any structural clash in the solvated systems. This was followed by heating up and brief equilibration for 20 ps in the NVT ensemble at 293 K with SANDER of AMBER8. Langevin dynamics with a time step of 2 fs was used in the heating and equilibration runs with a friction constant of 1 ps^{-1}.

To study the folded state of each solvated system, ten independent trajectories of 10.0 ns each in the NPT ensemble[42] at 293 K were then simulated with PMEMD of AMBER8. Here molecular dynamics with a time step of 2 fs was used for all production dynamics runs. To study unfolding pathways of each solvated system, ten independent unfolding trajectories of 20 ns each were performed in the NVT ensemble at 498 K but with the water density at 293 K (i.e., all high-temperature simulations were started from the end of the 10 ns 293 K trajectories). A total of 800 ns trajectories were collected for three solvated systems (bound PAZ, apo-PAZ, and siRNA) at both 293 K and 498 K, respectively, taking about 38,290 CPU hours on the in-house Xeon (1.86 GHz) cluster.

Native contacts for the bound and apo-PAZ were monitored to detect the beginning of unfolded state simulations. It was found that 8 ns at 498 K were needed to reach the equilibrium stage for both bound and apo-PAZ, so that the first 8 ns (a total of 80 ns for each system) were used to study unfolding kinetics and the remaining 12 ns (a total of 120 ns for each system) were used for equilibrium simulations at an unfolded state for each system. The siRNA is far less stable; it was found that only 3 ns at 498 K was enough to reach the unfolded state, so that the first 3 ns (a total of 30 ns) were used to study unfolding kinetics and the remaining 7 ns (a total of 70 ns) were used to study the unfolded equilibrium state.

Transition State Simulations. According to the definition of transition state (TS), 40 test MD runs for each candidate snapshot were performed to calculate the transition probability (*P*).[43–45] All test simulations have the same initial conformation for protein and solvent atoms but differing initial velocities. Because transition state structures agree with experiment under varying simulation temperatures,[46] TS simulations were done at 498 K in order to accelerate a simulated folding/unfolding rate. Each test trajectory is terminated when a conformation reaches the folded or unfolded state, as determined by Cα-rmsd. The Cα-rmsd vs simulation time at room and high temperature are shown in Figure 1s (Supporting Information). The results suggest that the folded state defines as the Cα-rmsd within 3.0 Å from the average structure at a folded state at 293 K. Up to 1 ns simulation at 498 K was found to be sufficient for each test trajectory, i.e. *P* values for tested snapshots were no longer changing when longer trajectories were run.

Free Energy Landscape Analysis. The unfolding landscapes were determined by calculating normalized probability from a histogram analysis.[43] Here we used a fraction of native tertiary contacts Qf and a fraction of native binding contacts Qb to map the unfolding landscape.

Data Analysis. Tertiary contact assignment was handled with in-house software. Two nonadjacent residues are in contact when their Cα atoms are closer than 6.5 Å. Secondary structure assignment was performed by using the DSSP program.[47] All rmsd and distance calculations were performed with PTRAJ in AMBER8.[36] The unfolding kinetics was fitted in Origin 7.0. Representative structures at unfolding half-times were used to construct unfolding pathways. Each representative structure is the closest snapshot to the average of all chosen snapshots at a given half-time (within ± its standard deviation).

Φ-values were computed with a strategy similar to those used in other studies[23,44,48]

$$\Phi_i^{calc} = \frac{N_i^{TS} - N_i^U}{N_i^F - N_i^U} \quad (1)$$

where N_i^{TS} is the number of native contacts of residue i at transition state, and N_i^F and N_i^U are the number of native contacts of residue i at folded and unfolded states, respectively. As defined by Calfisch and co-workers, native contacts were counted when the side chain heavy atoms of two nonadjacent residues are closer than 6.5 Å.[44]

Conformational clustering of bound and apo-PAZ at 293 K and 498 K was performed by using the MMTSB program.[49]

Results

Folded State. As a reference for the unfolding simulations, 10 trajectories of 10.0 ns each were simulated at 293 K to analyze the folded state of apo-PAZ, apo-siRNA, and their complex, respectively. To study the influence of siRNA binding on the stability of the folded PAZ, Cα and Φ/ψ variations for bound and apo-PAZ are illustrated in Figure 2. The Cα variation of bound PAZ is smaller than that of apo-PAZ, especially in the domain of the siRNA binding site. This suggests that bound PAZ become less flexible and more stable upon siRNA binding, consistent with experiment.[17] The Φ/ψ variation of bound PAZ is also smaller than that of apo-PAZ at the siRNA binding region, suggesting that the secondary structure stabilities have visible change upon siRNA binding.

To study the drive force for binding induced stable change in the folded state, the hydrophobic and hydrogen-bonding interactions between siRNA and PAZ were analyzed. All possible hydrophobic contacts and hydrogen bonds between siRNA and PAZ in the crystal structure were identified with Ligplot,[50] shown in Figure 2s (Supporting Information). The populations of six hydrophobic contacts in the simulation of ten trajectories are shown in Figure 3A. Five stable hydrophobic interactions can be found: F292/U9, T335/U9, Y309/U9, L337/U9, and T335/C8, with populations higher than 50%. Another one is rather unstable. Besides hydrophobic interactions, eleven possible hydrogen bonds were also identified with Ligplot.[50] Their populations in simulation are shown in Figure 3B. The results suggest that there are seven stable hydrogen bonds with a population higher than 50%. The other four hydrogen bonds are very weak. These strong hydrogen bonds between residues K264, H269, R275, R278, and Y314 and the Watson—Crick edges of the guanine and uracil bases (U9, U7, G4, and U3) and hydrophobic interactions are responsible for the stability in the PAZ domain.[17] This is consistent with the structural analysis by Ma et al.[17] These hydrophobic and hydrogen-bonding interactions play a key role in stabilizing the siRNA-protein interface. In summary, siRNA binding introduced more hydrophobic contacts and hydrogen bonds at the interface which are responsible for the higher stability in siRNA and PAZ.

Unfolding Kinetics. To investigate unfolding kinetics of bound and apo-PAZ, native tertiary contacts (Qf) and native binding contacts (Qb) are used to monitor unfolding and unbinding kinetics. The time scale of Qb and Qf for bound PAZ unfolding is illustrated in Figure 4. Apparently, the tertiary unfolding and unbinding kinetics can be represented well by single exponential functions, indicating that they are first order kinetics in the NVT ensemble at 498 K but with the room-temperature water density. The fitted kinetics data are listed in Table 1. Analysis shows that the unbinding half-time is 0.77 ns, and the unfolding half-time is 5.26 ns, suggesting that unbinding is faster than tertiary unfolding. The time evolution of Qf for apo-PAZ is also shown in Figure 4. It is found that tertiary unfolding of apo-PAZ also obeys first order kinetics, with a half-time of 2.09 ns, which is obviously faster than tertiary unfolding of bound PAZ. This suggests that the binding of siRNA significant postpones the tertiary unfolding of PAZ. This is consistent with the experimental observation.[17]

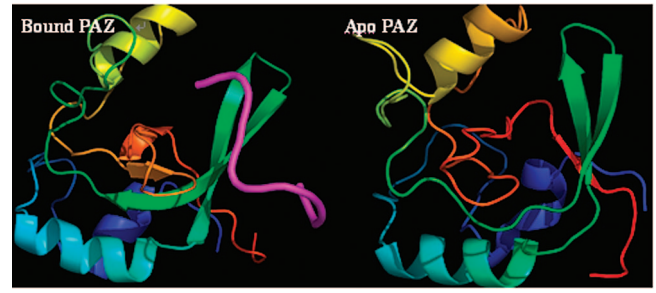

Figure 8. Average TSE structures for bound and apo-PAZ, respectively.

The unfolding kinetics of all three helices is also analyzed. The three helices can be grouped into two categories: helices (I & II) and helix III according to the position relative to siRNA in complex. Helix III is located at the binding pocket of siRNA and helices I and II are apart from siRNA. Time evolution curves are shown in Figure 5 for two groups of helixes in both bound and apo-PAZ, respectively. Our analysis shows that helical unfolding also obeys first order kinetics under the high-temperature simulation condition. THe unfolding half-time is 6.97 ns for helix III and 8.74 ns for helices (I and II), respectively, in apo-PAZ. The unfolding half-time is 12.93 ns for helix III and 9.70 ns for helices (I and II), respectively, in bound PAZ. Not surprisingly, helical unfolding is slower than tertiary unfolding and unbinding for both. This is consistent with other unfolding simulations of helical proteins, for example chymotrypsin inhibitor 2 and MDM2.[34,46] Note also that the helical unfolding half-times of the bound PAZ are longer than those of apo-PAZ, suggesting that the binding to siRNA stabilizes the helices in PAZ.[51] Furthermore, the unfolding half-time is reversed for helices (I and II) and helix III upon the binding of siRNA. This suggests that siRNA binding significantly changes the pathway of PAZ folding. Finally, since the seven β strands are very flexible, they are not monitored during unfolding.

Unfolding Landscapes. Furthermore, to understand the coupling mechanism between unfolding and unbinding, the unfolding landscape of bound PAZ was analyzed with the variables Qf and Qb (shown in Figure 6). The unfolding landscape shows that unbinding happens first while tertiary contacts are held stable and then followed by tertiary unfolding and unbinding. This is in agreement with the

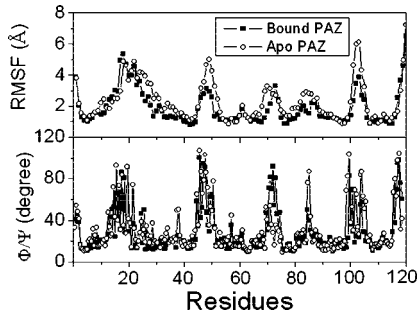

Figure 9. Ca and Ψ/Φ variations at TSE for bound and apo-PAZ, respectively.

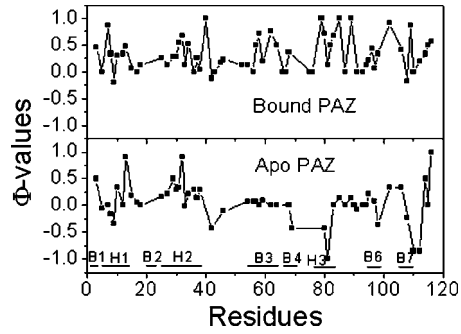

Figure 10. Predicted Φ-values of bound and apo-PAZ.

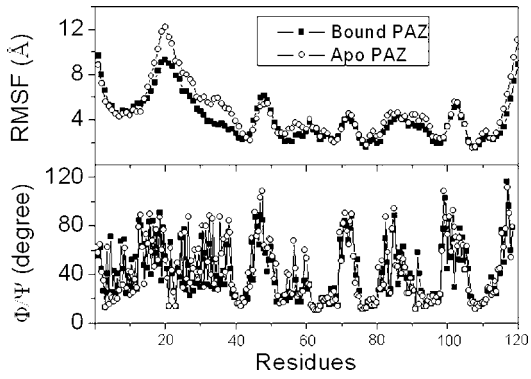

Figure 11. Structural variation of an unfolded state for bound and apo PAZ.

unfolding kinetics analysis for the bound PAZ. This suggests the formation of binding interface partly depends on formation of tertiary contacts.

The coupling between secondary and tertiary unfolding is also investigated for both bound and apo-PAZ, shown in Figure 3s (Supporting Information). Since seven β strands are too short or too flexible, only helical structures were monitored during unfolding. For bound PAZ, it is found that Qf decreases first while the content of helices (I and II) does not change. This is followed by a simultaneous decrease of tertiary contacts and the content of helices (I and II). This suggests that the tertiary unfolding is followed by secondary and tertiary unfolding. This is also consistent with the above unfolding kinetics analysis of bound PAZ. For apo-PAZ, Qf and the content of helices (I and II) simultaneously decrease, and then Qf accelerates to decrease. The coupling between the content of helix III and tertiary contacts is also monitored. We found that helix III keeps stable during tertiary unfolding for bound PAZ. For apo-PAZ, it is found that Qf decreases first, while the content of helix III does not change. Then Qf and the content of helix III simultaneously decrease. The different unfolding landscapes between Qf and helix III for bound and apo-PAZ suggest that helix III changes more stability upon siRNA binding and consistent with unfolding kinetics.

Transition State. Kinetics analysis shows that tertiary unfolding of both bound and apo-PAZ obeys first order kinetics. This suggests that bound and apo-PAZ unfold via a two-state process. Therefore, there is a transition state which corresponds to the free energy maximum between two states along each of their unfolding pathways. The transition state ensemble (TSE) structures can either fold or unfold, and the transition probability (P) will be 50%. According to the definition of TSE, we have scanned TSE structures from MD snapshots in all 10 unfolding trajectories for each of the bound and apo-PAZ, respectively.[43] The transition probability curves are further fitted by the Boltzmann equation and shown in Figure 7. The equation is $P = 1/(1 + \exp((t - \tau_{TS})/\tau_{trans}))$, where τ_{TS} is the time when $P = 50\%$, and τ_{trans} sets the period when P is between 0.4 and 0.6.[43–45] Our analysis yields 681 snapshots for bound PAZ TSE and 414 snapshots for apo-PAZ TSE, respectively.

Figure 8 illustrates the average structures for all TSE snapshots for bound and apo-PAZ, respectively. There are 60.7% native hydrophobic contacts and 72.3% native helical content for apo-PAZ TSE, while there are 73.5% native hydrophobic contacts and 88.2% native helical content for bound PAZ. Apparently, it can be concluded that the TSE of bound PAZ is more native-like than that of apo-PAZ. This indicates that the folding activation free energy for bound PAZ is probably smaller than that for apo-PAZ, leading to a relatively faster folding rate for bound PAZ. Combining with the above analysis that the unfolding rates for bound PAZ are slower than that of apo-PAZ, this leads to the conclusion that bound PAZ is more stable than apo-PAZ, based on the following relationship for two-state folding systems,[22,52] $K_{eq} = k_f/k_u$, where K_{eq} is folding equilibrium constant, and k_f and k_u are folding and unfolding rates, respectively. This conclusion on the relative stability between bound and apo-PAZ is consistent with the previous experimental observation.[17]

Figure 9 illustrates structural variations of TSE structures for bound and apo-PAZ, respectively. The Cα fluctuation of bound PAZ is smaller than that of apo-PAZ, suggesting that bound PAZ is more stable than that of apo-PAZ at the transition state. Thus TSE for bound PAZ is more native-like than that for apo-PAZ. However, the Φ/ψ fluctuation of bound PAZ is similar to that of apo-PAZ: there are some regions of helix and sheet with low fluctuation in both states.

All TSE snapshots were used to predict Φ-values of bound and apo-PAZ as shown in Figure 10. The sequence distribution of Φ-values suggests that the helices of helix I and II are more native-like than other loop regions for both bound and apo-PAZ. Furthermore, the Φ-values of helix III, β3, β4, β7, and the loop (314−334) of bound PAZ are significantly larger than those of apo-PAZ. This suggests that these regions of bound PAZ are more native-like than those

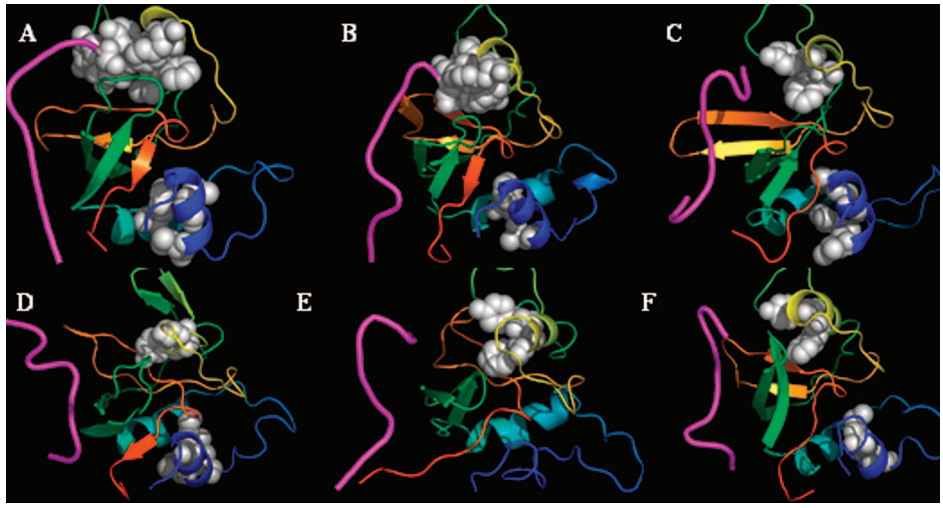

Figure 12. Unfolding pathway of bound PAZ. A: <0 ns (F), B: 0.77 ns (τ_{Qb}), C: 5.26 ns (τ_{Qf}), D: 9.70 ns (τ_{H12}), E: 12.93 ns (τ_{H3}), and F: >15 ns (U).

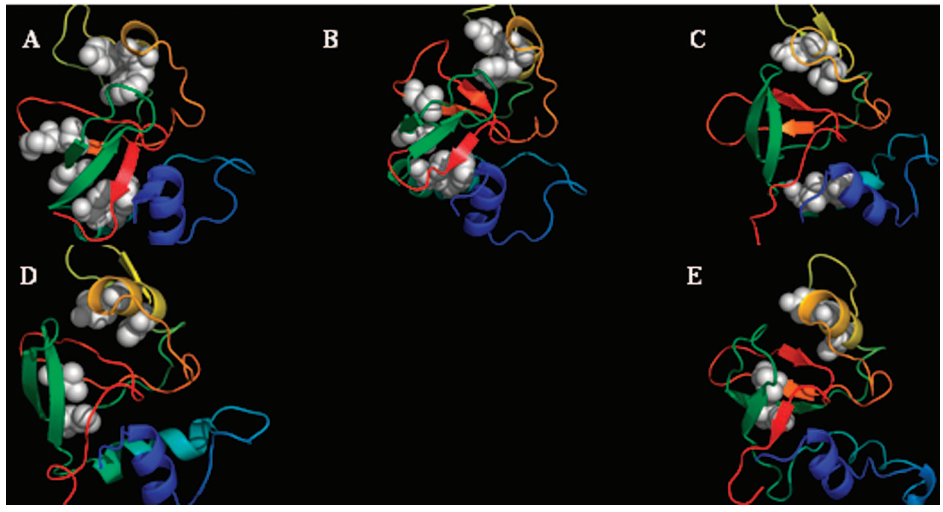

Figure 13. Unfolding pathway of Apo-PAZ. A: <0 ns (F), B: 2.09 ns (τ_{Qf}), C: 6.97 ns(τ_{H3}), D: 8.74 ns (τ_{H12}), and E: >15 ns (U).

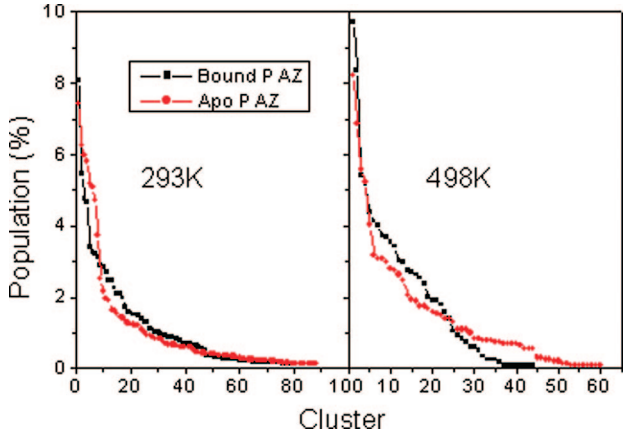

Figure 14. Cluster of bound and apo-PAZ.

of apo-PAZ upon the binding of siRNA and consistent with the unfolding kinetics analysis.

Unfolded State. All unfolding simulations eventually reach the unfolded equilibrium state under the high-temperature unfolding condition. For the unfolded state of apo-PAZ, there are 50.0% native hydrophobic contacts and 54.2% helical content remaining. Worth noting is the fact that the native hydrophobic contacts of H (I & II) all disappear. This is different from that of the bound-PAZ, for which the part hydrophobic contacts of H(I & II) remain. There are 57.8% native hydrophobic contacts and 60.4% helical content remaining.

Conformational fluctuation was also analyzed using structural variation from the average structures for the unfolded state (Figure 11). The Cα fluctuation of bound PAZ is smaller than that of apo-PAZ, indicating that bound molecules at an unfolded state are also more stable than apo molecules. However, there is little difference in the Φ/ψ fluctuation between bound and apo-PAZ.

Discussion

Comparison with Experiment. Structural analysis suggests that F292 and L337 are critical residues for efficient binding of siRNA to PAZ.[17] Our folded simulation illustrates that there are two stable hydrophobic contacts between F292 and U9 and L337 and U9. This suggests that simulation

results are consistent with structural analysis.[17] Beside these native hydrophobic contacts, there is also stable hydrophobic contact between Y309 and U9. This is in agreement with the mutation result that Y309F/Y314F causes a 46-fold reduction of binding affinity.[17]

Furthermore, it has been observed in experiment that PAZ must undergo an expansion upon siRNA binding, achieved through a rearrangement and outward expansion of the helices and β-sheets surrounding the binding pocket. In order to make quantitative comparison, the distance between the centers of β4-β5 and loop of 267−276 at the binding pocket opening in the experimental structures is measured and is found to increase by 0.48 Å upon siRNA binding. In simulation, the average distance between the centers of the same regions increases by about 0.25 Å upon siRNA binding, in qualitative agreement with experiment.[13,17]

Unfolding Pathways and Likely Folding Pathways. Based on the unfolding kinetics, landscape analysis, and transition state analysis, unfolding pathways for bound PAZ can now be constructed as shown in Figure 12. A) At the unbinding half-time, there are 36 out of 45 (folded state) native hydrophobic contacts within PAZ. The lost hydrophobic contact is within N-terminal and C-terminal. The native hydrophobic contacts between siRNA and PAZ also start to disappear: only 1 out of 4 exists. There is 81.1% native helical content remaining. B) At the half-time of tertiary unfolding, there are 32 native hydrophobic contacts within PAZ. The hydrophobic core is located mostly among H (I & II) and H III. All native hydrophobic contacts between siRNA and PAZ have disappeared. siRNA is partly structured. There is 75.9% helical content remaining. C) At the half-time of H (I & II) unfolding, there are 30 native hydrophobic contacts within PAZ. Two of the native hydrophobic contacts within H (I & II) have disappeared. siRNA moves away from the cavity of PAZ. There is 65.0% helical content remaining. D) At the half-time of H III unfolding, there are 29 native hydrophobic contacts within PAZ. The native hydrophobic contacts within H III began to unfold. There is 60.3% helical content remaining.

Similarly, the unfolding pathway of apo-PAZ is constructed and shown in Figure 13. A) At the half-time of tertiary unfolding, there are 35 out of 44 (folded state) native hydrophobic contacts. The hydrophobic core among β4-β5 almost disappears. The remaining hydrophobic interactions are within the regions of H III and H (I & II). There is 76.8% helical content remaining. B) At the half-time of H III unfolding, there are 33 native hydrophobic contacts within PAZ. Two of the native hydrophobic contacts within H III have disappeared. There is 65.3% helical content remaining. C) At the half-time of H (I & II) unfolding, there are 28 native hydrophobic contacts within PAZ. The hydrophobic core for siRNA binding disrupts. There is 64.7% helical content remaining.

If we assume that folding is the reverse of unfolding, the proposed folding/binding pathway of bound PAZ is the H III folding, H (I & II) folding, tertiary folding, and then siRNA binding. The folding order for apo-PAZ is H (I & II) folding, H III folding, and tertiary folding. The different order between the folding pathways of bound and apo-PAZ is the reverse of H III folding and H (I & II) folding upon the binding of siRNA.

Entropy Effect in the Binding/Folding Coupling. Figure 14 illustrates the cluster of bound and apo-PAZ at 293 K and 498 K, respectively. In the folded state, bound PAZ is less heterogeneous than apo-PAZ, indicating that the entropy decreases in the presence of siRNA, which induces from specific binding interactions between PAZ and siRNA. The strong and specific binding interactions make the specific binding favorable thermodynamically. In the unfolded state, however, there is no specific binding interaction between the two molecules. Nevertheless, the cluster of bound PAZ is also smaller than that of apo-PAZ, so that the entropy still decreases in the presence of siRNA. Thus the nonspecific binding of the two molecules is also favorable thermodynamically. Therefore, the entropic effect is significantly unfavorable in the coupled folding and binding in the PAZ-siRNA complex, and this unfavorable entropy effect weakens the strong contribution of hydrophobic and hydrogen bonding interactions between siRNA and PAZ. This is consistent in structural analysis that siRNA binding partly stabilizes the complex of siRNA-PAZ.[17]

Conclusion

Room-temperature MD simulations were used to study the stability of the PAZ domain uponing the binding of siRNA. The results suggest that both PAZ and siRNA become more rigid and stable upon siRNA binding. High-temperature MD simulations were used to investigate the folding kinetic of the PAZ domain. The results show that both bound and apo-PAZ unfold via a two-state process. The unfolding pathways are different between bound and apo-PAZ: the order of helix III and helices I & II unfolding is switched. Furthermore, transition probability was used to determine the transition state ensemble for both bound and apo-PAZ. It was found that the transition state of bound PAZ is more compact than that of apo-PAZ. The predicted Φ-values suggest that the Φ-values of helix III and sheets of β3-β7 for bound PAZ are more native-like than those of apo-PAZ upon the binding of siRNA. The results can help us to understand the mechanism of gene silencing.

Acknowledgment. This work is supported by the National Natural Science Foundation of China (Grants No. 30770502 and No. 20773085).

Supporting Information Available: Cα-RMSD vs simulation time for one trajectory (Figure 1s), two-dimensional representation for the interactions mode between siRNA and PAZ (Figure 2s), and unfolding landscapes for bound and apo-PAZ (Figure 3s). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Denli, A. M.; Hannon, G. J. RNAi: an ever-growing puzzle. *Trends Biochem. Sci.* **2003**, *28*, 196.

(2) Bartel, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **2004**, *116*, 281.

(3) Voinnet, O. RNA silencing as a plant immune system against viruses. *Trends Genet.* **2001**, *17*, 449.

(4) Volpe, T. A.; Kidner, C.; Hall, I. M.; Teng, G.; Grewal, S. I.; Martienssen, R. A. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **2002**, *297*, 1833.

(5) Peters, L.; Meister, G. Argonaute proteins: mediators of RNA silencing. *Mol. Cell* **2007**, *26*, 611.

(6) Hall, I. M.; Shankaranarayana, G. D.; Noma, K.; Ayoub, N.; Cohen, A.; Grewal, S. I. Establishment and maintenance of a heterochromatin domain. *Science* **2002**, *297*, 2232.

(7) Hammond, S. M.; Bernstein, E.; Beach, D.; Hannon, G. J. An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature* **2000**, *404*, 293.

(8) Zamore, P. D.; Tuschl, T.; Sharp, P. A.; Bartel, D. P. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **2000**, *101*, 25.

(9) Bernstein, E.; Caudy, A. A.; Hammond, S. M.; Hannon, G. J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **2001**, *409*, 363.

(10) Elbashir, S. M.; Lendeckel, W.; Tuschl, T. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.* **2001**, *15*, 188.

(11) Agrawal, N.; Dasaradhi, P. V.; Mohmmed, A.; Malhotra, P.; Bhatnagar, R. K.; Mukherjee, S. K. RNA interference: biology, mechanism, and applications. *Microbiol. Mol. Biol. Rev.* **2003**, *67*, 657.

(12) Song, J. J.; Liu, J.; Tolia, N. H.; Schneiderman, J.; Smith, S. K.; Martienssen, R. A.; Hannon, G. J.; Joshua-Tor, L. The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nat. Struct. Biol.* **2003**, *10*, 1026.

(13) Lingel, A.; Simon, B.; Izaurralde, E.; Sattler, M. Structure and nucleic-acid binding of the Drosophila Argonaute 2 PAZ domain. *Nature* **2003**, *426*, 465.

(14) Nykanen, A.; Haley, B.; Zamore, P. D. ATP requirements and small interfering RNA structure in the RNA interference pathway. *Cell* **2001**, *107*, 309.

(15) Hammond, S. M.; Boettcher, S.; Caudy, A. A.; Kobayashi, R.; Hannon, G. J. Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* **2001**, *293*, 1146.

(16) Knight, S. W.; Bass, B. L. A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in Caenorhabditis elegans. *Science* **2001**, *293*, 2269.

(17) Ma, J. B.; Ye, K.; Patel, D. J. Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature* **2004**, *429*, 318.

(18) Yan, K. S.; Yan, S.; Farooq, A.; Han, A.; Zeng, L.; Zhou, M. M. Structure and conserved RNA binding of the PAZ domain. *Nature* **2003**, *426*, 468.

(19) Henkels, C. H.; Kurz, J. C.; Fierke, C. A.; Oas, T. G. Linked folding and anion binding of the Bacillus subtilis ribonuclease P protein. *Biochemistry* **2001**, *40*, 2777.

(20) Henkels, C. H.; Oas, T. G. Ligation-state hydrogen exchange: Coupled binding and folding equilibria in ribonuclease P protein. *J. Am. Chem. Soc.* **2006**, *128*, 7772.

(21) Fersht, A. R.; Daggett, V. Protein folding and unfolding at atomic resolution. *Cell* **2002**, *108*, 573.

(22) Baker, D. Metastable states and folding free energy barriers. *Nat. Struct. Biol.* **1998**, *5*, 1021.

(23) Caflisch, A.; Karplus, M. Molecular-dynamics simulation of protein denaturation-solvation of the hydrophobic cores and secondary structure of barnase. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 1746.

(24) Caflisch, A.; Karplus, M. Acid and thermal-denaturation of barnase investigated by molecular-dynamics simulations. *J. Mol. Biol.* **1995**, *252*, 672.

(25) Daggett, V.; Li, A. J.; Itzhaki, L. S.; Otzen, D. E.; Fersht, A. R. Structure of the transition state for folding of a protein derived from experiment and simulation. *J. Mol. Biol.* **1996**, *257*, 430.

(26) Ladurner, A. G.; Itzhaki, L. S.; Daggett, V.; Fersht, A. R. Synergy between simulation and experiment in describing the energy landscape of protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 8473.

(27) Gsponer, J.; Caflisch, A. Role of native topology investigated by multiple unfolding simulations of four SH3 domains. *J. Mol. Biol.* **2001**, *309*, 285.

(28) Mayor, U.; Guydosh, N. R.; Johnson, C. M.; Grossmann, J. G.; Sato, S.; Jas, G. S.; Freund, S. M. V.; Alonso, D. O. V.; Daggett, V.; Fersht, A. R. The complete folding pathway of a protein from nanoseconds to microseconds. *Nature* **2003**, *421*, 863.

(29) Mayor, U.; Johnson, C. M.; Daggett, V.; Fersht, A. R. Protein folding and unfolding in microseconds to nanoseconds by experiment and simulation. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 13518.

(30) Akanuma, S.; Miyagawa, H.; Kitamura, K.; Yamagishi, A. A detailed unfolding pathway of a (beta/alpha)8-barrel protein as studied by molecular dynamics simulations. *Proteins* **2005**, *58*, 538.

(31) Scott, K. A.; Randles, L. G.; Moran, S. J.; Daggett, V.; Clarke, J. The folding pathway of spectrin R17 from experiment and simulation: using experimentally validated MD simulations to characterize states hinted at by experiment. *J. Mol. Biol.* **2006**, *359*, 159.

(32) Oard, S.; Karki, B. Mechanism of beta-purothionin antimicrobial peptide inhibition by metal ions: molecular dynamics simulation study. *Biophys. Chem.* **2006**, *121*, 30.

(33) Tsai, J.; Levitt, M.; Baker, D. Hierarchy of structure loss in MD simulations of src SH3 domain unfolding. *J. Mol. Biol.* **1999**, *291*, 215.

(34) Chen, H. F.; Luo, R. Binding induced folding in p53-MDM2 complex. *J. Am. Chem. Soc.* **2007**, *129*, 2930.

(35) Day, R.; Daggett, V. Direct observation of microscopic reversibility in single-molecule protein folding. *J. Mol. Biol.* **2007**, *366*, 677.

(36) Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L., III; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Wang, B. Pearlman, D. A.; Crowley, M.; Brozell, S. Tsui, V.; Gohlke, H. Mongan, J.; Hornak, V. Cui, G. Beroza, P.; Schafmeister, C.; Caldwell, J. W. Ross, W. S.; Kollman, P. A. *AMBER 8*; University of California: San Francisco, 2004.

(37) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926.

(38) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: an N log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089.

(39) Wang, J. M.; Cieplak, P.; Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules. *J. Comput. Chem.* **2000**, *21*, 1049.

(40) Lwin, T. Z.; Lu, Q.; Luo, R. Force field influences in protein folding simulations. *Protein Sci.* **2006**, *15*, 2642.

(41) Rychaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Comput. Phys.* **1977**, *23*, 327.

(42) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684.

(43) Pande, V. S.; Rokhsar, D. S. Molecular dynamics simulations of unfolding and refolding of a beta-hairpin fragment of protein G. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9062.

(44) Gsponer, J.; Caflisch, A. Molecular dynamics simulations of protein folding from the transition state. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 6719.

(45) Chong, L. T.; Snow, C. D.; Rhee, Y. M.; Pande, V. S. Dimerization of the p53 oligomerization domain: identification of a folding nucleus by molecular dynamics simulations. *J. Mol. Biol.* **2005**, *345*, 869.

(46) Day, R.; Daggett, V. Sensitivity of the folding/unfolding transition state ensemble of chymotrypsin inhibitor 2 to changes in temperature and solvent. *Protein Sci.* **2005**, *14*, 1242.

(47) Kabsch, W.; Sander, C. Dictionary of protein secondary structure-pattern-recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983**, *22*, 2577.

(48) Vendruscolo, M.; Paci, E.; Dobson, C. M.; Karplus, M. Three key residues form a critical contact network in a protein folding transition state. *Nature* **2001**, *409*, 641.

(49) Feig, M.; Karanicolas, J. ; Charles, L. ; Brooks, I. I. I. *MMTSB Tool Set. MMTSB NIH Research Resource*; The Scripps Research Institute: 2001.

(50) Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* **1995**, *8*, 127.

(51) Petsko, G. A.; Ringe, D. *Protein Structure and Function. Chapter 2. From Structure to Function*; New Science Press: London, 2003.

(52) Nolting, B. *Protein Folding Kinetics, Biophysical Methods*; Springer: Berlin, 1999.

CT800030S