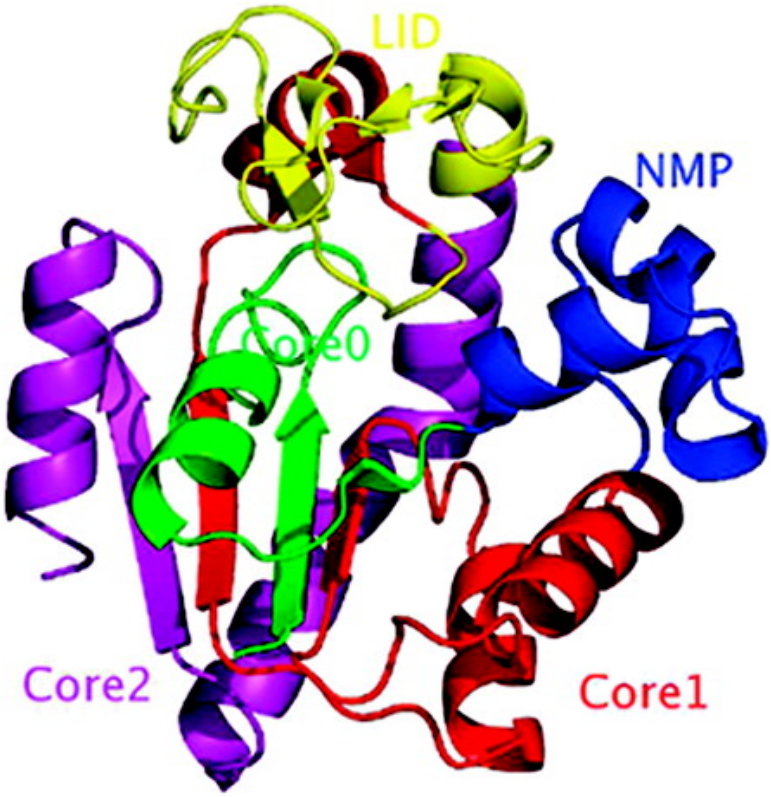
Article

Single Molecule Conformational Dynamics of Adenylate Kinase: Energy Landscape, Structural Correlations, and Transition State Ensembles

Qiang Lu, and Jin Wang

J. Am. Chem. Soc., **2008**, 130 (14), 4772-4783 • DOI: 10.1021/ja0780481



J|A|C|S
ARTICLES
Published on Web 03/14/2008

Single Molecule Conformational Dynamics of Adenylate Kinase: Energy Landscape, Structural Correlations, and Transition State Ensembles

Qiang Lu[†] and Jin Wang*[,†,‡,§]

Department of Chemistry and Department of Physics and Astronomy, State University of New York at Stony Brook, Stony Brook, New York 11794, and State Key Laboratory of Electroanalytical Chemistry, Changchun Institute of Applied Chemistry, Chinese Academy of Sciences, Changchun, Jilin 130022, People's Republic of China

Received October 20, 2007; E-mail: jin.wang.1@stonybrook.edu

Abstract: We developed a coarse grained two-well model to study the single molecule protein conformational dynamics in microscopic detail at the residue level, overcoming the often encountered computational bottleneck. In particular, we explored the underlying conformational energy landscape of adenylate kinase, a crucial protein for signal transduction in the cell, and identified two major kinetic pathways for the conformational switch between open and closed states through either the intermediate state or the transient state. Based on the parameters fitted to the room-temperature experimental data, we predicted open and closed kinetic rates at the whole temperature ranges from 10 to 50 °C, which agree well with the experimental turnover numbers. After uncovering the underlying mechanism for conformational dynamics and exploring the structural correlations, we found the crucial dynamical interplay between the nucleoside monophosphate binding domain (NMP) and the ATP-binding domain (LID) in controlling the conformational switch. The key residues and contacts responsible for the conformational transitions are identified by following the time evolution of the two-dimensional spatial contact maps and characterizing the transition state as well as intermediate structure ensembles through ϕ value analysis. Our model provides a general framework to study the conformational dynamics of biomolecules and can be applied to many other systems.

1. Introduction

Conformational dynamics has been a central issue in understanding many important processes in molecular biology. It often links biological function with the associated collective conformational changes and reactions.[1] Single molecule measurements can provide direct probes to the conformational dynamics. The intrinsic fluctuations and conformational dynamics are not weighted down by the number of molecules, as in the bulk, and can be probed and extracted directly through either long time or multiple short time measurements.[2,3] Interesting properties about many molecular processes such as protein folding, molecular recognition, and conformational dynamics have been revealed with the analysis of single molecule time series.[4−7,21]

In bulk studies, it is hard to distinguish whether complex kinetic behavior is due to the inhomogeneity of each molecule or the intrinsic properties themselves. Single molecule experiments can determine the underlying kinetic mechanisms and identify the pathways directly without ambiguities. This provides a framework for studying single molecule conformational dynamics at the microscopic level.

Theoretical studies on single molecule conformational dynamics so far have been mostly focused on qualitative behaviors; therefore, they are on the phenomenological level, without structural details.[2,8,9] Processes in conformational dynamics

[†] Department of Chemistry, State University of New York at Stony Brook.
[‡] Department of Physics and Astronomy, State University of New York at Stony Brook.
[§] Chinese Academy of Sciences.

(1) McCammon, J. A.; Wolynes, P. G.; Karplus, M. *Biochemistry* **1979**, *18*, 927−942.
(2) Wang, J.; Wolynes, P. *Phys. Rev. Lett.* **1995**, *74*, 4317−4320.
(3) Lu, H. P.; Xun, L. Y.; Xie, X. S. *Science* **1998**, *282*, 1877−1882.
(4) Leite, V. B.; Onuchic, J. N.; Stell, G.; Wang, J. *Biophys. J.* **2004**, *87*, 3633−3641.
(5) Lu, Q.; Lu, H. P.; Wang, J. *Phys. Rev. Lett.* **2007**, *98*, 128105−128108.
(6) Lipman, E. A; Schuler, B.; Bakajin, O.; Eaton, W. A. *Science* **2003**, *301*, 1233−1235.
(7) Zhuang, X. W. *Annu. Rev. Biophy. Biomol. Struct.* **2005**, *34*, 399−414.
(8) Geva, E.; Skinner, J. L. *Chem. Phys. Lett*. **1998**, *288*, 225−229.
(9) Yang, S. L.; Cao, J. *J. Chem. Phys.* **2004**, *121*, 572−581.
(10) Miyashita, O.; Onuchic, J. N.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12570−12575.
(11) Zuckerman, D. M. *J. Phys. Chem. B* **2004**, *108*, 5127−5137.
(12) Okazaki, K.; Koga, N.; Takada, S.; Onuchic, J. N.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 11844−11849.
(13) Whitford, P. C.; Miyashita, O.; Levy, Y.; Onuchic, J. N. *J. Mol. Biol.* **2007**, *366*, 1661−1671.
(14) Best, R. B.; Chen, Y. G.; Hummer, G. *Structure* **2005**, *13*, 1755−1763.
(15) Maragakis, P.; Karplus, M. *J. Mol. Biol.* **2005**, *352*, 807−822.
(16) Temiz, N. A.; Meirovitch, E.; Bahar, I. *Proteins* **2004**, *57*, 468−480.
(17) Arora, K.; Brooks, C. L. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 18496−18501.
(18) Henzler-Wildman, K. A.; Thai, V.; Lei, M.; Ott, M.; Wolf-Watz, M.; Fenn, T.; Pozharski, E.; Wilson, M. A.; Petsko, G. A.; Karplus, M.; Hübner, C. G.; Kern, D. *Nature* **2007**, *450*, 838−844.
(19) Henzler-Wildman, K. A.; Lei, M.; Thai, V.; Kerns, S. J.; Karplus, M.; Kern, D. *Nature* **2007**, *450*, 913−916.
(20) Wolf-Watz, M.; Thai, V.; Henzler-Wildman, K.; Hadjipavlou, G.; Eisenmesser, E. Z.; Kern, D. *Nat. Struct. Mol. Biol.* **2004**, *11*, 945−949.
(21) Hanson, J. A.; Duderstadt, K.; Bhattacharyya, S.; Brokaw, J.; Chu, J. W.; Yang, H. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 18055−18060.

10.1021/ja0780481 CCC: $40.75

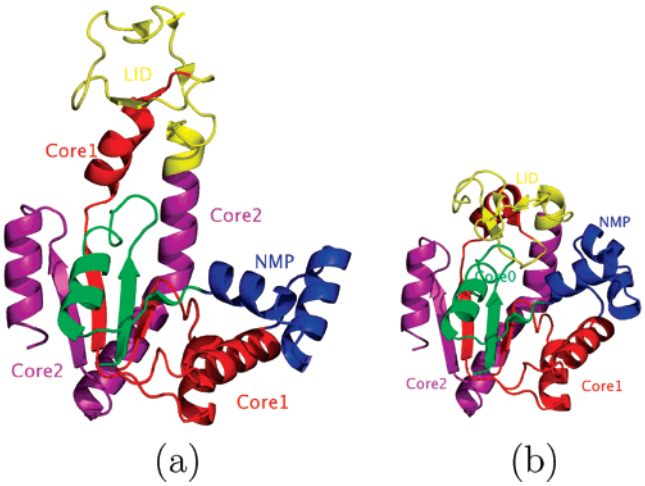


Figure 1. Ribbon diagrams of the crystal structures of ADK: (a) open conformation (4AKE);[24] (b) the closed conformation (1ANK).[22] The AMP binding domain, NMP (residue 31−60), is colored in blue, and LID (residue 127−164) is colored in yellow. The CORE domains are separated into three regions: Core0 (residue 1−30), Core1 (residue 61−126), and Core2 (residue 165−217). They are colored in green, red, and pink.

typically happen in milliseconds (sometimes seconds to minutes) while currently all atom molecular dynamics simulations can only reach up to microsecond time scales at best. Therefore, it is difficult to explore conformational dynamics in atomic detail. Thus, there have been increasing efforts recently toward understanding conformational dynamics at the coarse grained level[10−17] by biasing the resulting macroscopic potential of the whole system to two reference states, such as open and closed structures. In this study, we developed a model with microscopic two-well potential between two residues based on the known open and closed structures, rather than with resulting macroscopic potentials of the whole system, to correlate the structures and conformational dynamics. In this way, one can bridge the gap between theories and experiments by exploring the relationship among structure, dynamics, and function.

Using this model, we studied adenylate kinase (ADK),[22,24] which is a phosphotransferase enzyme that catalyzes the production of ATP, whose conformational dynamics is crucial for its function, and for which NMR[18−20] and single molecule experiments[21] are being carried out. ADK is a protein vital to the function of signal transduction cellular networks. The conformational change is between open (pdb: 4AKE) and closed (pdb: 1ANK) structures, shown in parts a and b, respectively, of Figure 1.

By exploring the underlying conformational energy landscape of ADK, we found the landscape has two major basins with one basin bias toward the open and the other toward the closed conformation of the protein. The formation of the basin probably reflects of the evolutionary constraint for biological function. We also observed two major kinetic pathways, with one through a stable intermediate state and another through an unstable transient state. Both intermediate states have shallower basins than those of open and closed states. The saddle points between the basins provide information on the location and height of the free energy barrier for the conformational transitions from open to closed or vice versa. The entire conformational dynamics of the system is thus determined by the interbasin dynamics, which depends on the barriers and basins. Furthermore, we can uncover the structural origin for the conformational change in the molecule and explore physical mechanisms of multiple time scales, which is often observed. The landscape approach has the advantages of going beyond the harmonic oscillator approximation or normal-mode analysis[16,25] for the protein conformations to reflect the anharmonic nature of the conformational changes and going beyond the smooth basin approximation to reflect the local roughness critical for the detailed dynamics and function.

2. Results and Discussion

Our microscopic two-well model simulated the detailed process of conformational changes. Note that the conformational energy landscape here is defined as the conformational energy as a function of the $3N$ dimensional space. The chemical reactions usually refer to bond breaking and bond forming or particular conformational switching events. The reaction coordinate refers to the order parameter optimally describing the reaction event. The reaction landscape only refers to the coordinates relevant to the reactions.

Our model gives a statistical correlation of $R^2 = 0.787$ with the open structure crystal *B*-factor and gives the statistical correlation 0.35 with the closed structure, which is consistent with the *B*-factor analysis in ref 13.

2.1. Conformational Energy Landscapes and Pathways. Figure 2a shows a typical trajectory in time for RMSD1 (root-mean-square deviation from closed structure) and RMSD2 (rmsd from nucleoside monophosphate-binding domain (NMP)-closed and ATP-binding domain (LID)-opening structures). We clearly see conformation switching between open and closed states through intermediates. Figure 2b shows the 2-D free energy landscape as a function of RMSD1 and RMSD2. Free energies are correlated with the steady-state probabilities of the specific states which can be probed in single molecule experiments through the fluorescence intensity distribution. The distribution is observed to peak at open and closed structures, and sometimes in the intermediate states.[19,21] The results here are consistent with single molecule experiments and provide physical origin in terms of the underlying energy landscapes.

From Figure 2b, one can find two major pathways connecting the open and closed structures. Each pathway has intermediate or transient structures. Here we label the deeper intermediate as "LID-closing'', where the NMP domain is open but LID is closed, and the corresponding path as the "LID-closing pathway''. We label the shallow intermediate as "NMP-closing'', where LID is open but NMP is closed, and the corresponding path as the "NMP-closing pathway''. Note that NMP closes faster than LID does in the NMP-closing pathway and opens first in the LID-closing pathway, while LID closes faster than NMP does in the LID-closing pathway and opens first in the NMP-closing pathway. LID-closing is an intermediate structural ensemble which has barrier heights over $3kT$ to both open and closed basins. The NMP-closing state has barrier heights of about $1kT$ and $3kT$ to open and closed basins, respectively. One of the barrier heights is in the range of thermal motion (kT) and can be easily crossed. This means the NMP-closing

(22) Muller, C. W.; Schulz, G. E. *J. Mol. Biol.* **1992**, *224*, 159−177.
(23) Åden, J.; Wolf-Watz, M. *J. Am. Chem. Soc.* **2007**, *129*, 14003−14012.
(24) Muller, C. W.; Schlauderer, G. J.; Reinstein, J.; Schulz, G. E. *Structure* **1996**, *4*, 147−156.
(25) Ma, J. *Structure* **2005**, *13*, 373−380.

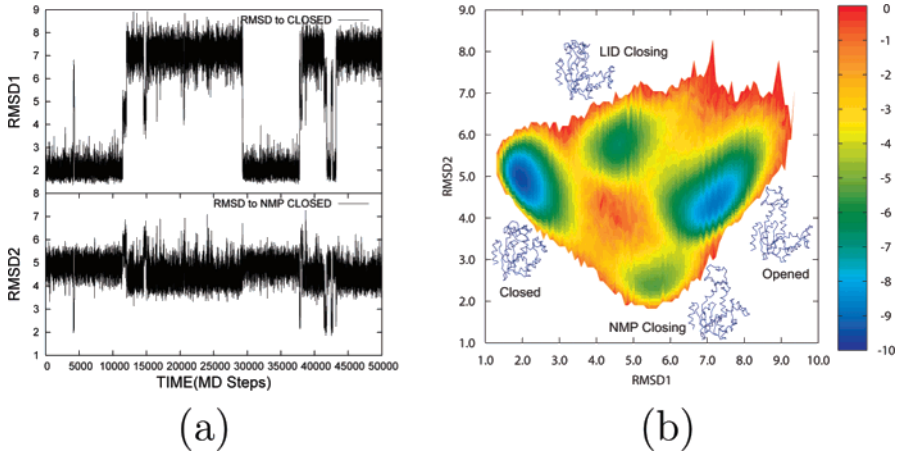


Figure 2. At room temperature, part a shows the trajectory in time for RMSD1 (the root-mean-square deviation from the closed structure) and RMSD2 (from the NMP-closing conformation) and part b shows the 2-D free energy as function of RMSD1 and RMSD2. There are two pathways connecting the open and closed structures. These two pathways go through LID-closing intermediate and NMP-closing transient structures.

structures are not thermodynamically stable and can be considered as transient conformations. Previous static investigations led to open and closed structures with only one intermediate state,[13,15] with no further studies on alternative pathways. Our microscopic dynamic study results in different pathways with intermediate and transient states. This adds to the complexity of ADK conformational dynamics. Different pathways reflect the roughness of the underlying conformational energy landscape. This can be observed by our two-well microscopic model, but hard to be captured in the models of the resulting macroscopic potential of the whole system.[12−15] This is because, in our model, the resulting macroscopic potential is a sum of all the microscopic double wells between different residue pairs. This leads to a rougher landscape with more possibilities of intermediate states and discrete kinetic paths than the two-well macroscopic potential models. The NMR experiments[23] also provide the evidence of both LID-closing and NMP-closing pathways and intermediate states.[19]

We ran simulations at 12 different temperatures and obtained the 2-D free energy profile for each temperature. After extracting the 1-D free energy profile from the 2-D free energy profile along both pathways and plotting them together, we obtained Figure 3a for the LID-closing path and Figure 3b) for the NMP-closing pathway. Here, ΔG_{closed} (ΔG_{open}) is the free energy value of the closed (open) basin, ΔG_{bar1} (ΔG_{bar2}) is the free energy value of the transition state barrier near the closed (open) state, and ΔG_{mid} is the free energy value of the intermediate or transient state. There are four free energy barrier heights along both pathways labeled as $F_{\text{close}} = \Delta G_{\text{bar1}} - \Delta G_{\text{closed}}$; $F_{\text{open}} = \Delta G_{\text{bar2}} - \Delta G_{\text{open}}$; $F_{\text{mid1}} = \Delta G_{\text{bar1}} - \Delta G_{\text{mid}}$; $F_{\text{mid2}} = \Delta G_{\text{bar2}} - \Delta G_{\text{mid}}$. The free energy barrier heights along the LID-closing pathway and along the NMP-closing pathway as a function of temperature are shown in parts c and d, respectively, of Figure 3. We observe that the closed state free energy is more sensitive to the temperature changes. In addition, the four relative barrier heights, in units of kT, F_{closed}, F_{open}, F_{mid1}, and F_{mid2}, all decrease, when the temperature increases. This implies that a higher temperature leads to faster kinetics of conformational switching. Along the LID-closing pathway in the opening direction, the opening barrier F_{closed} is slightly lower compared to that of the NMP-closing pathway. It would seem easier for ADK to go across this barrier. However, due to the significant barrier height of the second transition state, $F_{\text{mid2}} = 3kT$, many first-barrier crossing events are stopped at the intermediate state. Along the NMP-closing pathway, although the first barrier height F_{closed} is slightly higher than that of the other path, the second barrier height F_{mid2} is only $1kT$ and much easier to cross. This makes the NMP-closing pathway the dominant one for ADK opening. The NMP-closing pathway appears 87% of the time, while the LID-closing pathway appears 13% of the time. Along the NMP-closing pathway, in the closing direction, both the first barrier height F_{open} and the second barrier height F_{mid1} are lower than the corresponding ones in the LID-closing pathway. This leads to the NMP-closing pathway occurring 86% of the time and the LID-closing pathway occurring 14% of the time. In other words, in the closing direction, the NMP domain closes first more often. On the other hand, in the opening direction, LID opens first more often. In bulk experimental studies, the two pathways are hard to distinguish distinctly due to the possible inhomogeneous distribution of each individual molecule. As we see in Figure 2a, the single molecule trajectories can provide direct probes of multiple conformational states and multiple pathways for conformational transitions without ambiguities encountered in the bulk studies.

Figure 4 describes the free energies ΔG_{closed} and ΔG_{open} at the closed state and the open state as a function of temperature. ΔG_{mid1} and ΔG_{mid2} are the free energies of the intermediate states of the LID-closing and NMP-closing pathways, respectively. The free energies are $\Delta G_{\text{closed}} = -9.63kT$ at the closed state and $\Delta G_{\text{open}} = -8.88kT$ at the open state. Therefore, the system prefers the closed state at room temperature. This agrees with the NMR experimental results.[20,21,23] However, at higher temperatures beyond 32 °C, we predict that the population of the open state will become higher than that of the closed state (in Figure 4), while the kinetic rate for closing is about the same as the kinetic rate for opening only at 37 °C, which is body temperature (Figure 3d). The free energy values are $-6.64kT$ at the LID-closing state and $-5.43kT$ at the NMP-closing conformations at room temperature. Therefore, the LID-closing conformations have populations about 3.4 times larger than those of the NMP-closing conformations, which is comparable to the recent single molecular experimental measurements[21] that LID-closing conformations have about 3.7 times larger populations than NMP-closing conformations have.

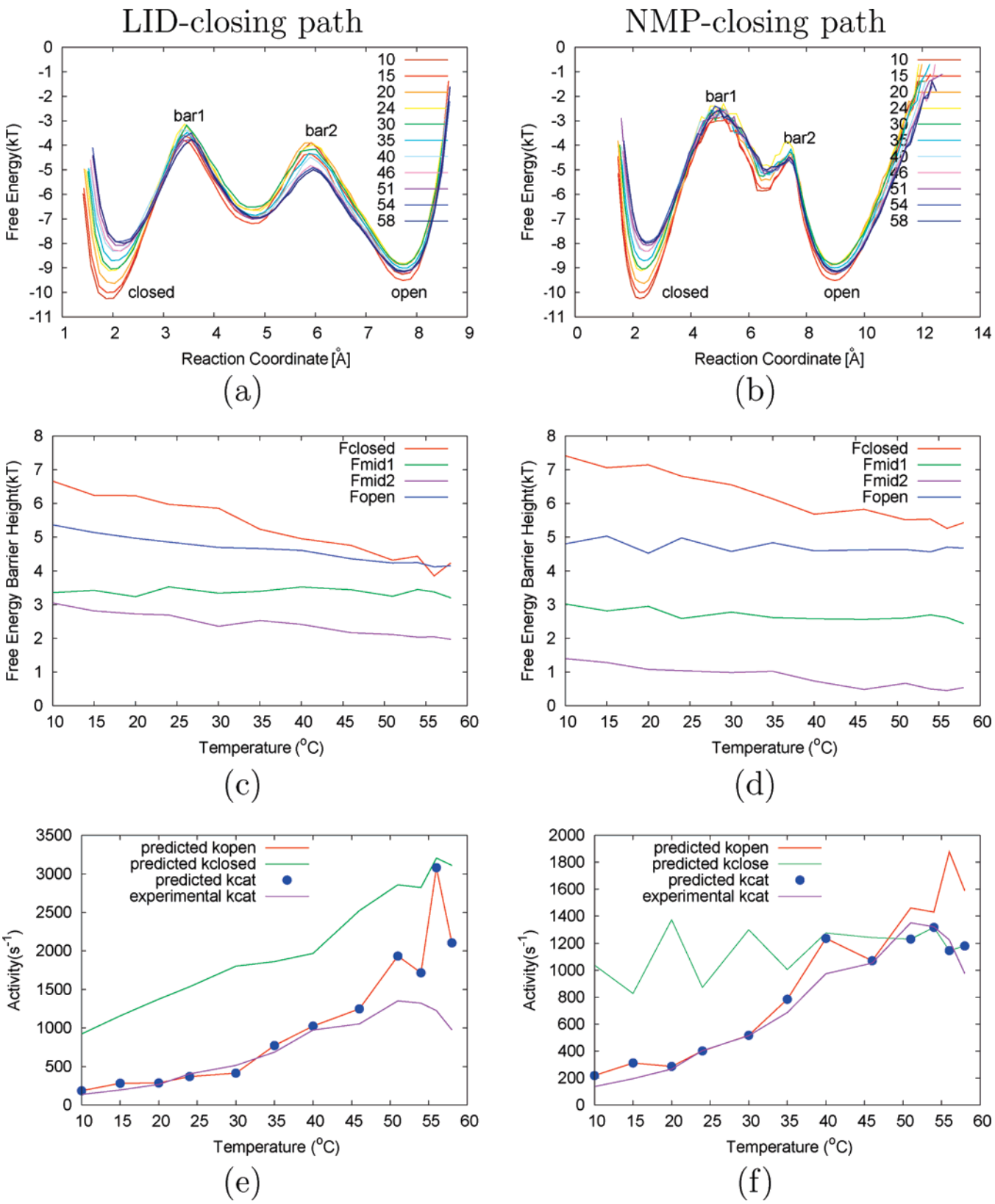


Figure 3. At different temperatures, part a shows the 1-D free energy profile along the LID-closing pathway, part b shows the 1-D free energy profile along the NMP-closing pathway, and part c shows the barrier heights along the LID-closing pathway. There are four barrier heights, $F_{closed} = \Delta G_{bar1} - \Delta G_{closed}$; $F_{open} = \Delta G_{bar2} - \Delta G_{open}$; $F_{mid1} = \Delta G_{bar1} - \Delta G_{mid}$; and $F_{mid2} = \Delta G_{bar2} - \Delta G_{mid}$. Part d shows the barrier heights along the NMP-closing path, part e shows the comparison between the predicted k_{open}, k_{closed}, and the experimental turnover number, k_{cat}, along the LID-closing path, and part f shows the comparison between the predicted k_{open}, k_{closed}, and the experimental turnover number, k_{cat}, along the NMP-closing path.

Based on our model, we also followed a recent single molecule experiment[21] to study the time series of a specific residue pair distance and did not find evidence for the intermediate state of ADK accordingly, which is consistent with the experimental result. However, this does not mean that there is no intermediate state, because, according to our model, the

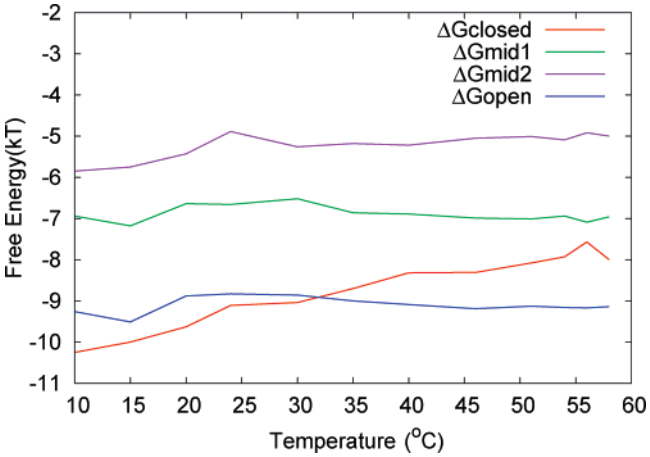


Figure 4. At different temperatures, ΔG_{closed} and ΔG_{open} are the free energies of the closed state and the open state, and ΔG_{mid1} and ΔG_{mid2} represent the free energies of the intermediate states of the LID-closing and NMP-closing pathways.

intermediate state is an ensemble of collective events of many residue pairs forming stable contacts instead of specific individual ones. The collective counting of the contacts gives a possible reaction coordinate. This suggests for the future single molecule experiments that, instead of monitoring a single residue pair, one should either monitor multiple residue pairs simultaneously or make multiple time series measurements with each focusing on a different single residue pair; this would pin down the underlying mechanisms of conformational switching. Our model implies that both pathways can be explored and identified by monitoring multiple residue contact pairs.

2.2. Reaction Rate Prediction. In comparison with experimental results by Kern's group,[20] we predicted the catalytic rate constant, also called the turnover number or k_{cat}, i.e. a first-order rate constant corresponding to the slowest step in the overall catalytic pathway, at different temperatures. Here we used the highest barrier height to determine its value. In other words, smallest value of k_{closed} and k_{open} is used as the approximation of k_{cat}, in terms of the relations $k_{open} = k_1 \exp^{-F_{closed}/kT}$ and $k_{closed} = k_2 \exp^{-F_{open}/kT}$. However, the prefactors k_1 and k_2 are not available in experiments. We have to use our predicted F_{closed} and F_{open} (in Figure 3b and d) and experimental $k_{open} = 286\ s^{-1}$ and $k_{closed} = 1374\ s^{-1}$ (ref 20) to deduce k_1 and k_2 at room temperature and apply them to the predictions of k_{open}, k_{closed}, and k_{cat} at other temperatures. The comparisons of experimental k_{cat} values and our predicted k_{cat}, k_{open}, and k_{closed} values are plotted in Figure 3e for the LID-closing pathway and in Figure 3f for the NMP-closing pathway. Through comparisons, we found that, below 50 °C, k_{open} is the rate-limiting step in both pathways. k_{cat} increases with temperature due to decreasing of barrier heights (shown in Figure 3c and d). Beyond 50 °C, the slowing down of ADK closure makes it the rate-limiting step along the NMP-closing pathway (green line in Figure 3f). We believe it is the reason for the lowering of k_{cat} in the experiment. Beyond 54 °C, both k_{open} and k_{closed} decrease, and k_{cat} continues to drop as the temperature increases. This could be because the system becomes less stable near the melting temperature.

In our model, the NMP-closing pathway is dominant in both the closing and opening directions, and LID opening in this pathway is the rate-limiting step, which agrees with the NMP

Table 1. Path Weight of ADK Opening and Closing for the Two Pathways[a]

mut among	NMP open	LID open	NMP close	LID close
no mut	0.87	0.13	0.86	0.14
NMP−LID	0.48	0.52	0.52	0.48
LID−Core0	0.93	0.07	0.92	0.08
LID−Core2	0.76	0.24	0.85	0.15
NMP−Core0	0.79	0.21	0.87	0.13
NMP−Core2	0.15	0.85	0.11	0.89
NMP−NMP	0.65	0.35	0.70	0.30

[a] "NMP'' and "LID'' represent the NMP-closing and LID-closing pathways, respectively. The "open'' ("close'') means along the direction of opening (closing).

experimental conclusion that LID opening was found to be the rate-limiting step.[19,20] In addition, normal-mode analysis[16] found that LID has the slowest mode and NMP has the second slowest mode for an open structure; this supports our results in the way that, when the closing barrier height F_{closed} of NMP is comparable to, although slightly higher than, that of the LID closing barrier, the normal-mode frequency will become important for kinetics in the form of the prefactor. A higher normal-mode frequency of NMP will significantly help NMP close first more often than LID does. This provides another reason for the prevailing of the NMP-closing pathway in addition to the barrier height explanation. Furthermore, in the LID-closing pathway, the barrier steepness of the NMP opening is higher than that of LID closing (Figure 3a); this also agrees with the elastic network model analysis.[10]

2.3. Characterizations of Transition and Intermediate State Ensembles. We now characterize the transition and intermediate state ensembles through ϕ value and residue contact pair probabilistic analysis. ϕ value analysis can provide the importance of a particular residue upon mutation. Contact probability provides the importance of particular residue pair interactions. For the LID-closing pathway, in Figure 5, the left column shows the high ϕ value for the first transition state, the intermediate, and the second transition state. Figure 6 shows the corresponding contact pairs' probability (left) and typical structures (right). These analyses together indicate the link between the structures and conformational dynamics along the LID-closing pathway.

At the first transition state from the open structure, in Figure 5a and Figure 6a and b, LID starts to move toward the Core domain, beginning from the Hinge region of residue R119, which connects G12 in the Core0 region and F137 in LID. These three residues all interact with the ligand in the closed X-ray structure.

At the intermediate state, in Figures 5b and 6c and d, LID binds to the Core domain, with help from the long side-chain residues I120, R123, and R119 in the Hinge region, forming contacts with Core0 region residues, P9, G10, and G12, as well as LID residues, R123, V132, and F137. Dragged by the Hinge region, LID residues, S129, G130, R131, V132, and N138, form contacts with Core0 residues, G14, T15, Q18, and Q28. The only contacts between LID and Core2 are F137(LID)−V202(Core2) and N138(LID)−V202(Core2). These residues' ϕ values are already higher than 0.7 at this stage. Meanwhile, the hinge residue D33 of NMP is within the reach of LID residues A127, S129, and G130 and is preparing for the NMP closing action.

At the second transition state, in Figures 5c and Figure 6e and f, LID forms stable contacts with Core. NMP hinge residue

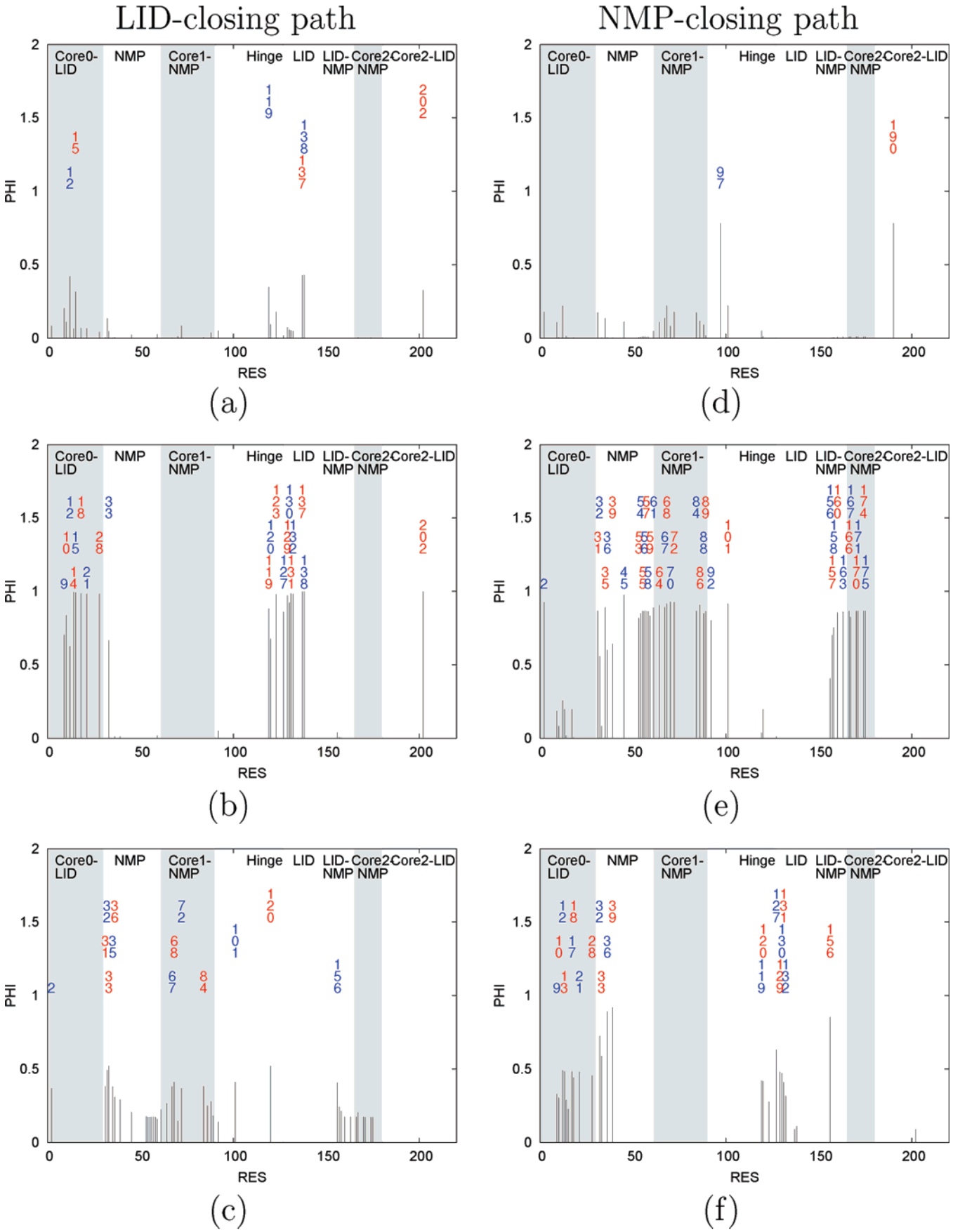


Figure 5. (a−c) ϕ values along the LID-closing pathway and (d−f) ϕ values along the NMP-closing pathway: (a and d) at the first transition state; (b and e) at the intermediate state; (c and f) at the second transition state. The high ϕ values (>0.8) in parts b and e are not shown in parts c and f to avoid redundancy.

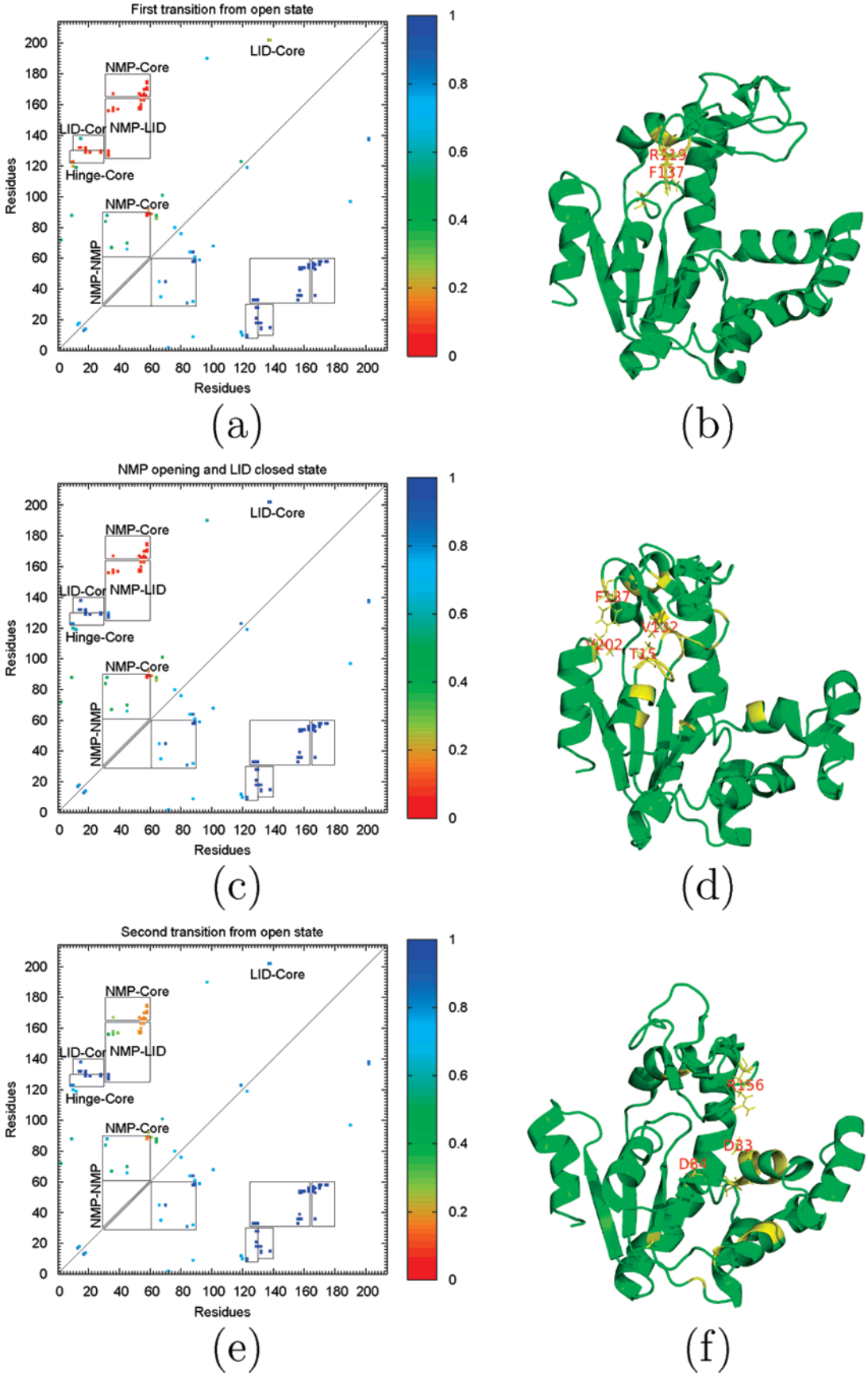


Figure 6. Along the LID-closing pathway, there are three conformational state ensembles: first transition state, the LID-closing intermediate state (NMP opened, LID closing), and the second transition state. The left column shows the residues contact map at the three configurational ensembles. Blue means full contact, and red means no contact. The lower right corner of the contact map shows the closed structural contacts, and the upper left corner shows the contacts at a certain state. The right column shows the typical structures at the three configurational ensembles. The high-ϕ residues are colored in yellow.

D33 involves more association with LID by the salt bridge D33−R156 and with Core1 by the contacts T31−D84, T31−L67, and T31−V68. The interactions between LID and NMP introduce NMP to Core1. The ϕ values of the NMP residue are relatively small, yet they imply that the NMP domain is starting to close.

For the NMP-closing pathway, at the first transition state, in Figures 5d and Figure 7a and b, the Core1 region is adjusting its structure before the NMP-closing step. The polar residue pair K97(Core1)−N190(Core2) is of particular interest, which locates at the bottom and on the surface of ADK, connecting two separated Core regions. It is easily affected by the solvent environment. The interesting point is that, before NMP closing, the pair has a small C_α distance of 9.83 Å. When NMP is closed, the C_α distance becomes 10.84 Å. Similarly to the LID-closing pathway, K97 and N190 are adjacent to each other only at the second transition state right before the NMP closing and enlarged after the closing. Thus, we believe that the pair K97−N190 is highly related to the NMP closing behavior, mimicking a switch deciding which pathway the ADK will take. Given that, the environment could significantly affect the ADK's pathway decision.

At the intermediate state, in Figures 5e and Figure 7c and d, NMP residues T31, L35, L45, M53, L58, and V59 are forming mainly hydrophobic contacts with Core1 residues V64, A66, L67, V68, and F86. NMP is analoguous to a cap covering the hydrophobic surface of Core1. We believe the function of these hydrophobic residues is to help the closing of NMP and stabilize the closed structure. Furthermore, NMP also made contacts with LID and its Hinge region at this stage. There are mainly charged contacts among LID and NMP, such as those of NMP residues R36, D54, and K57 with LID residues R156, K157, D158, Q160, R167, and E170. These charged residues are almost all involved in the ligand interactions in the closed structure. These interactions are observed in the contact list of our simulations as well. The ligand seems to be the primary switch to control the binding of NMP and LID. Although many LID/Core and NMP/Core contacts are related to the ligand, there are not as many charged residue contacts as the NMP−LID has. NMP−LID association should be more sensitive to the ligand charge modification before and after the catalysis, which can trigger the association or disassociation of LID and NMP simultaneously, so that LID and NMP could close and open cooperatively. Once the NMP is closed and bound to LID, it will further help and stabilize LID closing together with the help from ligand attraction.

At the second transition state, in Figures 5f and Figure 7e and f, more NMP and LID contacts are formed, including D33(NMP)−R156(LID), which are ligand related. LID's hinge residue R119 first contacts with Core0 residue G12. Both of them contact with ligand also. Triggered by the LID hinge and NMP, LID (S129, G130, R131) begins to approach Core0 (Q18, Q28, M21). From the 1-D free energy profile (Figure 3b), the barrier height is about $3.0kT$, indicating that it is much easier for LID to close with the help of NMP than with itself facing the $6.0kT$ barrier height alone in the LID-closing pathway.

All the residues mentioned above have high ϕ values and contact probabilities among each other in Figures 5−7. Through single molecule measurements, the distribution of fluorescence intensity can be measured and used to infer the equilibrium free energy profiles along the reaction coordinate. The single molecule time trajectories and ϕ value analysis can unambiguously identify the multiple conformational states and distinct transition pathways as well as the key structural elements compared to the bulk studies.

2.4. Mutation Studies. Figure 8 shows the 2-D free energy as a function of RMSD1 (root-mean-square deviation from closed structure (1ANK)) and RMSD2 (root-mean-square deviation from the LID-opening and NMP-closed structure) in the cases of (a) nonmutation and in the cases of having all the contacts deleted between (b) NMP and LID, (c) LID and Core0, (d) LID and Core2, (e) NMP and Core0, and (f) NMP and Core2, and (g) in the case of having all NMP intradomain contacts deleted. There are two pathways connecting the open and closed structures that go through intermediate or transient states, which are the LID-closing intermediate state and the NMP-closing transient state. Figure 8h illustrates where LID, NMP, Core0, Core1, and Core2 are on an open structure of ADK.

In Figure 8b, when NMP-LID contacts are deleted, both the open state and the LID-closing state have lower free energies; this suggests that the whole system prefers the open state, and the weight of the LID-closing pathway increases, as has been confirmed by Table 1. Another noticeable difference from the nonmutation free energy profile is that all the barriers between open and closed states decrease, which should result in all the actions including LID closing/opening and NMP closing/opening becoming easier. The easier closure could be due to the deletion of the second well, which mimicks the long-range water-mediated interactions among NMP−LID contacts. This effectively eliminates the traps toward the closed structure and leaves only the Core domain interactions intact. The increased possibility of ADK opening could be due to the removal of the short-range interactions among LID−NMP contacts. Therefore, both LID and NMP lose many mutual restraints and have more freedom after mutation. Our results show that both pathways have almost equal chances of appearing now (Table 1).

Comparing Figure 8c (LID−Core0) with Figure 8a (nonmutation), we find that the open state free energy is lower than the closed state free energy, reflecting the system preference for the open state. There is an uphill barrier from the open state to the LID-closing intermediate state, so the LID is easy to open and hard to close. From Table 1, almost all closing and opening events are through the NMP-closing pathway, indicating that NMP often closes first and then drags LID, together with help from Core, to close. Without the other's help, it would be hard for LID to close. Correspondingly, in the opening direction, LID often opens first and reaches the NMP-closing transient state faster (due to smaller barriers along the pathway) than with no mutations. This is consistent with the conclusion drawn above from Figure 8c.

Comparing Figure 8d (LID−Core2) with Figure 8a (nonmutation), we find that there are no significant differences between the two, suggesting that LID−Core2 interactions are not as important as LID−Core0 interactions. Similar conclusions can be drawn for the NMP−Core0 mutation from the free energy profile (Figure 8e) as well. From Table 1, when the mutations occur between LID−Core2, the LID closing is not affected but the LID opening becomes slightly harder. This could also be due to the second well effect. But there are only a few contacts

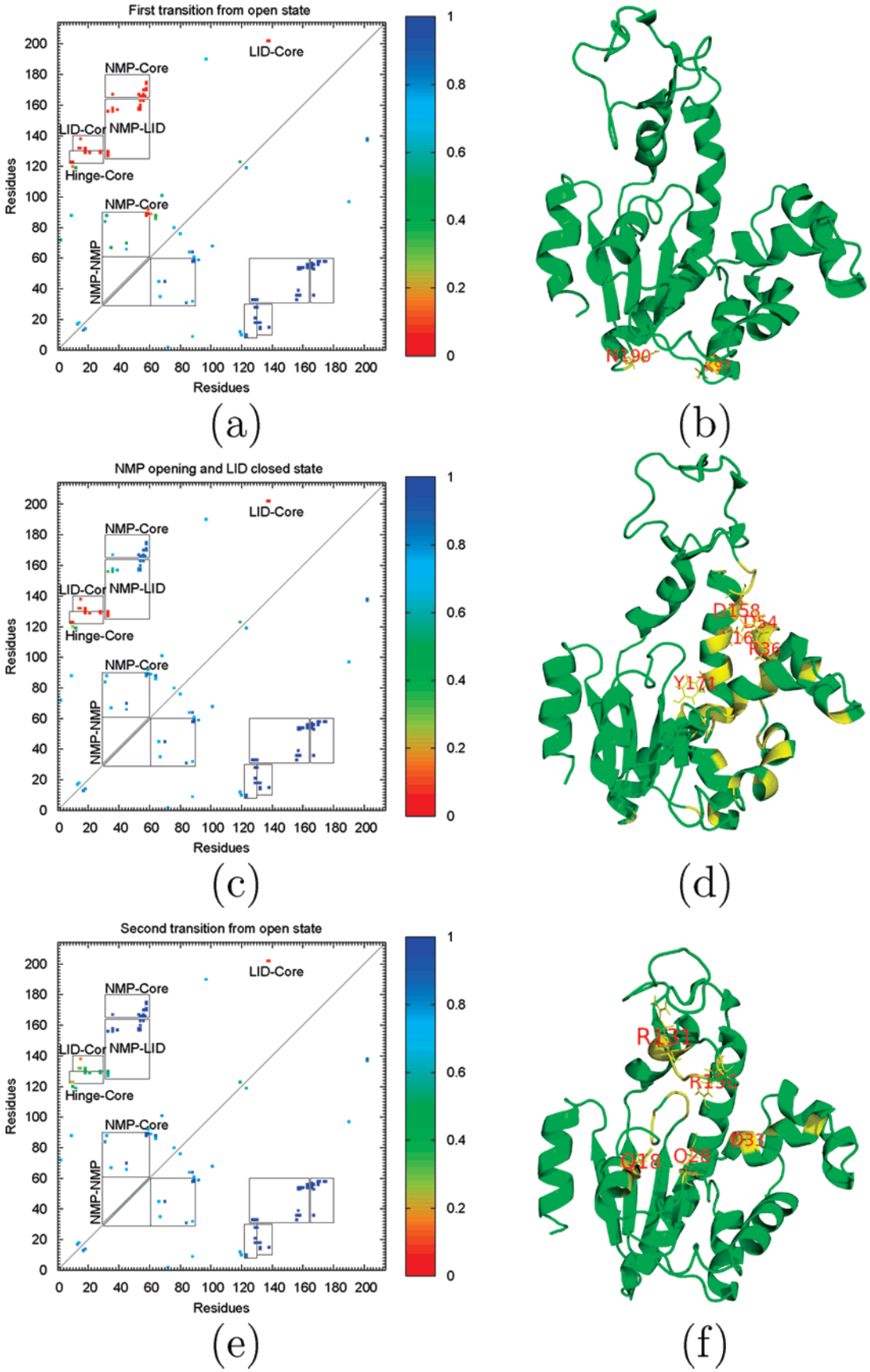


Figure 7. Along the NMP-closing pathway, there are three conformational state ensembles: the first transition state, the NMP-closing transient state (NMP closing, LID opening), and the second transition state. The left column shows the residue contact map at the three configurational ensembles. Blue means full contact, and red means no contact. The lower right corner of the contact map shows the closed structural contacts, and the upper left corner shows the contacts at a certain state. The right column shows the typical structures at the three configurational ensembles. The high-ϕ residues are colored in yellow.

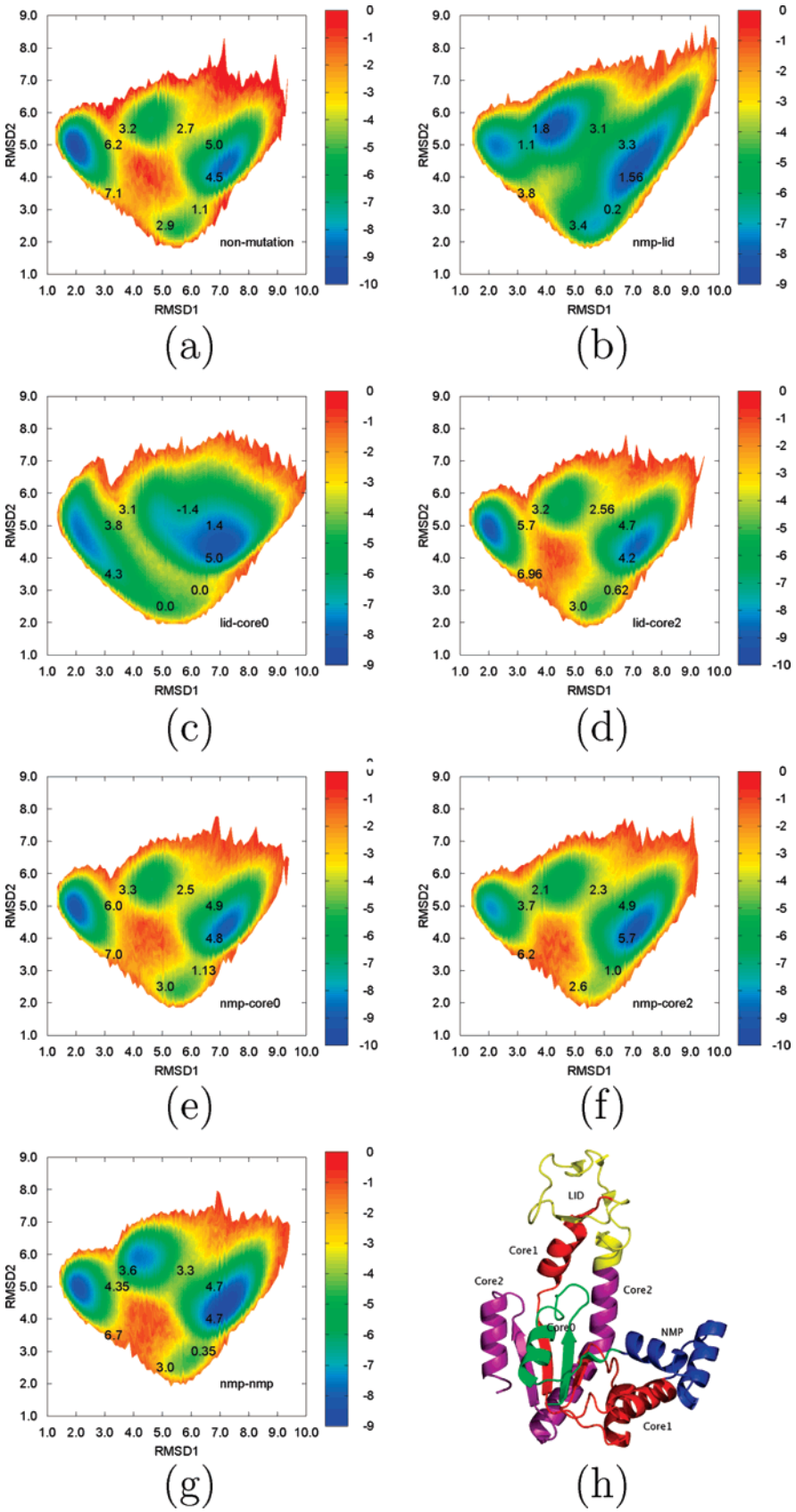


Figure 8. 2-D free energy as a function of RMSD1 (root-mean-square deviation from closed) and RMSD2 (NMP closing conformation). In the cases of: (a) nonmutation, all the contact interactions are deleted between (b) NMP and LID domains, (c) LID and Core0 domains, (d) LID and Core2 domains, (e) NMP and Core0 domains, (f) NMP and Core2 domains, and (g) the NMP intradomain. There are two pathways connecting the opening and closing structures. These two pathways go through intermediate or transient structures; they are the LID-closing intermediate and NMP-closing transient structures. Part h illustrates where the LID, NMP, Core0, Core1, and Core2 domains are. The values labeled in the free energy profiles are the free energy differences between the adjacent transition state and the bottom of the free energy well.

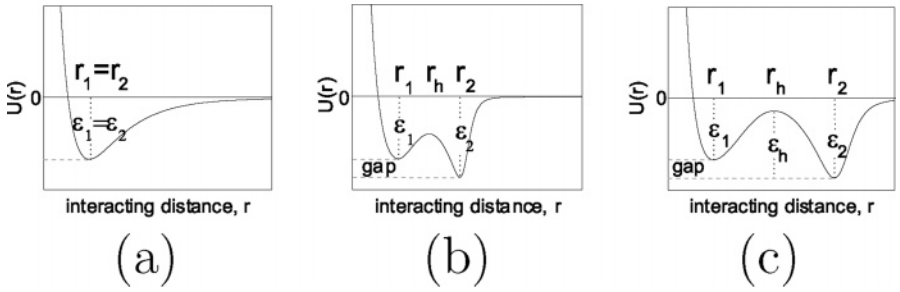


Figure 9. Coarse grained microscopic interaction energy between the two residues, when the native closed contacts and the corresponding native open contacts are (a) at the same distance, (b) less than 2 Å, or (c) larger than 2 Å.

with large distance interactions, and most of them only have short range interactions. That is why the LID closing is not affected much but the opening becomes slightly harder.

In Figure 8f (NMP−Core2), we find that the open state has a lower free energy than the closed state, which is in contrast to the nonmutation free energy profile; this shows that the system becomes reluctant to close after the mutation. From Table 1, once the NMP−Core2 interactions are deleted, the closing and opening are mostly through the LID-closing pathway. This suggests that NMP closes often after the LID closing and opens often before the LID opening, which is opposite to the nonmutation case where NMP closes before the LID closing and opens after the LID opening. In this situation, it should be more easy to find the LID-closing intermediate structure for experimentalists.

In Figure 8g, when the intradomain interactions are deleted within the NMP domain, the open state and LID-closing state have lower free energies, indicating that NMP is harder to close and easier to open. However, from Table 1, the weight of the NMP-closing path is still higher than the weight of the LID-closing path, indicating the NMP still closes before the LID closing and opens after the LID opening. By eliminating the intranal interactions within NMP, we believe that the NMP structure becomes extended and unstable. Therefore, NMP is reluctant to close and more inclined to open.

3. Conclusions

In this paper, we developed a general coarse grained model to study the conformational dynamics of ADK. In particular, we explored the underlying conformational energy landscape; we predicted open and closed rates at the whole temperature ranges that compared well with the experimental turnover number. We found both NMP-closing and LID-closing pathways, whose dynamics are controlled by the free energy barriers of open, intermediate, and closed conformations. Additionally, we uncovered the structural correlations to the conformational changes and identified the hot residues as well as contact interactions. This provides rich information about the microscopic origin for the conformational change at the single molecule level and, furthermore, may open a way for engineering and designing biomolecules and enzymes for specific functions and effective conformational change.

4. Method and Materials: Computational Details

In order to capture the crucial details of the topography of the global basins of the energy landscape and monitor the detailed intra- and interbasin conformational dynamics in the whole time domain, we developed a coarse grained two-well model, which takes into account native interactions that exist in both open and closed structures. Due to evolution constraints, the energetic frustration is assumed to be small. By construction, we include the topological frustration or structural (entropy) heterogeneity from native structure constraints.

4.1. Two-Well Model. We developed here an off-lattice model, where each residue is represented by a single bead centered on its α-carbon (C_α) position. The interaction energy U at a given protein conformation Γ is

$$U(\Gamma,\Gamma_1,\Gamma_2) = \sum_{\text{bonds}}^{N-1} K_b(b_i - b_{0i})^2 + \sum_{\text{angles}}^{N-2} K_\theta(\theta_i - \theta_{1i})^2(\theta_i - \theta_{2i})^2 + \sum_{\text{dihedrals}}^{N-3} K_\phi\left[\cos\left(\phi_i - \frac{\phi_{1i}+\phi_{2i}}{2}\right) - \cos\left(\frac{\phi_{1i}-\phi_{2i}}{2}\right)\right]^2 + \sum_{|i-j|>3}^{\text{nat}} U_{\text{nat}}(r_{ij}) + \sum_{|i-j|>3}^{\text{non-native}} \epsilon\left(\frac{C}{r_{ij}}\right)^{12} \quad (1)$$

The first three terms control the bond, angle vibration, and dihedral rotation within four adjacent residues. The parameter b_{0i} is the average of the bonds of both conformations Γ_1 and Γ_2 (native open and closed). θ_{1i} (θ_{2i}) and ϕ_{1i} (ϕ_{2i}) stand for the corresponding variables at the native open (closed) structure Γ_1 (Γ_2). For the energy parameters, $K_b = 100.0$, $K_\theta = 20.0$, and $K_\phi = 1.0$.[5,26] The fourth and fifth terms represent the native and non-native interaction energy (repulsion) between two nonadjacent residues, respectively. C (=4.0 Å) parametrizes the excluded volume repulsion between residue pairs. $U_{\text{nat}}(r_{ij})$, inspired by ref 27, is given explicitly in the Supporting Information and discussed below. All second, third, and forth terms bias toward both native open and closed structures. Figure 9 shows our coarse grained microscopic two-well potential between two residues. We can clearly see the potential biased toward both native open and closed structures. When the same specific residue pair forming a contact in the closed structure is the same as that in the open structure, a single well potential emerges biasing toward the same open and closed structure native contact in Figure 9a. When the same specific residue pair forming a contact in the closed structure is not the same but within a cutoff distance (2 Å) with the contact formed in the open structure, a two-well potential with a shallow barrier emerges with each biasing toward the closed structural and the open structural native contact, as shown in Figure 9b. For the same specific residue pair forming a contact in the closed structure, the spatial distance between the pair in the open structure can be large ($|r_{ij}^{\text{open}} - r_{ij}^{\text{closed}}| > 2$ Å). In this case, a two-well potential still emerges with each biasing toward the closed structure native contact and the open structure long distance contact, as shown in Figure 9c. The native contact maps for both open and closed structures are derived with the CSU software.[28]

Figure 9b represents the biasing toward native structure with small fluctuations between open and closed structures. The reason for the

(26) Levy, Y.; Wolynes, P. G.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 511−516.
(27) Cheung, M. D.; Garcia, A. E.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 685−690.
(28) Sobolev, V.; Wade, R. C.; Vriend, G.; Edelman, M. *Proteins: Struct., Funct., Genet.* **1996**, *25*, 120−129.

presence of the second well at a large distance that could be over 30 Å, as shown in Figure 9c, is because there exists a balance between repulsive interactions and structural constraints. The repulsive interactions originate from the presence of charged hydrophilic residues among LID, NMP, and Core. Water molecules can bridge two hydrophilic or charged residues separated by relatively large distances.[29] The structural constraints from the open structure are mainly caused by the backbone secondary structures and the geometries. This effectively leads to a potential wall around the open structure and, when combined with a long distance water mediated repulsion, results in the second well. The repulsion mimicked by a downhill potential pushes the contact distance further to long distances; the wall from the open structure constraint mimicked by the steep uphill potential stops it at a certain place.

Although the ligand was not explicitly included to calculate the contact list, some ligand contacts are implicitly considered by including some long distance interactions without the presence of the ligand to resemble the contact interactions mediated by the ligand.

Molecular dynamics simulations were performed, by using AMBER6 as an integrator, starting from either the open or closed conformation. Fifty trajectories with 50 million steps each were performed to determine the parameters and for free energy profiles.

4.2. Parametrization. There are three parameters ϵ_1, ϵ_2, and ϵ_h. ϵ_h determines at which temperature the closing and opening activity will occur. Here we choose 275 K, which is slightly higher than 0 °C, to determine $\epsilon_h = 0.55$ kcal/mol. It agreed with the fitting result of data in Figure 1d of ref 20. Note that the barrier height is the height difference between the barrier and the deepest well. The deepest well, or the maximum of ϵ_1 or ϵ_2, determines the melting temperature, which is about 58 °C.[20] Furthermore, the open and closed rate at room temperature[20] can help us to determine the free energy "GAP" between the open and closed states. After running the simulations at both room temperature and 58 °C, we determined the double wells' depths are $\epsilon_1 = 0.64$ kcal/mol and $\epsilon_2 = 0.53$ kcal/mol separately, so that our model can have melting behavior at 58 °C and the resulting free energy landscape and reaction rate of closing and opening agree with experimental results at room temperature. Fifty trajectories with 50 million steps each were performed in order to obtain ϵ_1, ϵ_2, and the free energy profiles by $F = -\log(P)$. P is the statistical population accumulated from all conformational dynamic runs.

4.3. Advantages, Limitations, and Applications. The advantage of our model is that it is a microscopic one, where each pair of the residues is interacting with a two-well potential. This is in contrast with the macroscopic models where the resulting potential of all the pairs of interactions together is two well in nature. As we see, our microscopic two-well model can capture the intermediate states between the open and close conformational states caused by the inherent roughness of the resulting energy landscape. A limitation of the current model is that it is coarse grained and residue based. The detailed side chain properties are not considered. An atomic based microscopic two-well model can be developed to study further more details.

Another improvement of our model is that we considered the long-range water mediated interactions, which have not been studied so far. The simulation can be performed at room temperature and compared to experimental data. Our model is a structural based one still dependent on the native open and closed structures rather than an ab initio based one. Another limitation of our model is that the interactions between substrate (ligand) and the protein are implicitly treated rather than explicitly studied.

Our model can be applied to other enzyme systems even if the experimental data on k_{open} and k_{closed} are not known. There is an approximate way to determine the two-well difference. Due to the fact that most of the closed structures of the enzymes correspond to the conformations for biological functions of catalysis, we may assume that at approximately body temperature the closed state population is higher than that of the open state. In this study for the ADK system, only below 32 °C does the closed state have a higher population. Once the melting temperature is known, one can test different ϵ_1 and ϵ_2 values around the body temperature and find a set of parameters resulting in the population of the closed state being higher than or almost equal to that of the open state.

After this work was finished,[30] we heard that another model with a microscopic double well was proposed.[31] The difference seems to lie in the fact that we considered the water mediated interactions and structural constraints as long range (without a predetermined cutoff) interactions while the other group did not.[11,30,31]

Acknowledgment. J.W. and Q.L. ackowledge support from a National Science Foundation Career Award and the American Chemical Society Petroleum Fund. Most of the simulations were performed on a Pacific Northwest National Laboratory supercomputer, and some were performed on the SUNYSB Seawolf Supercomputer Cluster.

Supporting Information Available: Mathematical form of the two-well potential, parametrization, the method for ϕ value calculations, path weight at different temperatures, and mutation studies. This material is available free of charge via the Internet at http://pubs.acs.org.

JA0780481

(29) Papoian, G. A.; Ulander, J.; Eastwook, M. P.; Luthey-Schulten, Z.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 3352−3357.

(30) Wang, J.; Lu, Q. *Abstr. Pap., 234th ACS Natl. Meet.* **2007**.

(31) Chu, J. W.; Voth, G. A. *Biophys. J.* **2007**, *97*, 3860−3871.