



**Role of substrate-product frustration on enzyme functional dynamics**Jiayang Kong, Jiachen Li , Jiajun Lu, Wenfei Li ,\* and Wei Wang<sup>†</sup>*National Laboratory of Solid State Microstructure, Department of Physics, and Collaborative Innovation Center of Advanced Microstructures, Nanjing University, Nanjing 210093, China*

(Received 4 July 2019; published 25 November 2019)

Natural enzymes often have enormous catalytic power developed by evolution. Revealing the underlying physical strategy used by enzymes to achieve high catalysis efficiency is one of the central focuses in the field of biological physics. Our recent work demonstrated that multisubstrate enzymes can utilize steric frustration encountered in the substrate-product cobound complex to overcome the bottleneck of the enzymatic cycle [W. Li *et al.*, *Phys. Rev. Lett.* **122**, 238102 (2019)]. However, the key atomic-level interactions by which the steric frustration contributes to the enzymatic cycle remain elusive. In this work we study the microscopic mechanism for the role of the substrate-product frustration on the key physical steps in the enzymatic cycle of adenylate kinase (AdK), a multisubstrate enzyme catalyzing the reversible phosphoryl transfer reaction  $\text{ATP} + \text{AMP} \rightleftharpoons \text{ADP} + \text{ADP}$ . By using atomistic molecular dynamics simulations with enhanced sampling, we showed that the competitive interactions from the phosphate groups of the substrate ATP and product ADP in the ATP-ADP cobound complex of the AdK lead to local frustration in the binding pockets. Such local frustration disrupts the hydrogen bond network around the binding pockets, which causes lowered barrier height for the opening of the enzyme conformations and expedited release of the bottleneck product ADP. Our results directly demonstrated from the atomistic level that the local frustration in the active sites of the enzyme can be utilized to facilitate the key physical steps of the enzymatic cycle, providing numerical evidence to the predictions of the previous theoretical work.

DOI: [10.1103/PhysRevE.100.052409](https://doi.org/10.1103/PhysRevE.100.052409)**I. INTRODUCTION**

The sequence of a natural protein has evolved to meet different biological requirements from both folding and functions. For robust folding, the frustration of physical interactions in the native structure of a folded protein tends to be minimized as suggested by the principle of minimal frustration [1,2]. On the other hand, because the requirement to the sequence from function can be conflicting with that from folding, the physical interactions in the functionally relevant sites of proteins are often locally frustrated [3–7]. However, the question of how such local frustration contributes to protein functions is not fully understood.

Enzymes are biomolecular machines (mostly proteins) that catalyze various biochemical reactions through the enzymatic cycle [8], including substrate binding, chemical reaction, and product release steps. Natural enzymes have an enormous ability to catalyze chemical reactions. The ability involves not only accelerating individual chemical steps but also coordinating all the chemical and physical steps to complete an enzymatic cycle. For many enzymes, the product release step, which is coupled with conformational motions, is often the rate-limiting step of the enzymatic cycle [9–11]. Understanding the strategies natural proteins use to coordinate these essential dynamic processes of the enzymatic cycle to realize high efficiency catalysis is extremely challenging, and has

been one of the central focuses in the field of biological physics [9,12,13]. In a recent theoretical work, we showed that multisubstrate enzymes can utilize steric frustration encountered in the substrate-product cobound complex to overcome the bottleneck of the enzymatic cycle [14], by using the adenylate kinase (AdK) as a model enzyme.

AdK is a bisubstrate-biproduct enzyme that catalyzes the reversible phosphoryl-transfer reaction  $\text{ATP} + \text{AMP} \rightleftharpoons \text{ADP} + \text{ADP}$  [15]. It plays a key role on the energy balance in the cell by maintaining appropriate ATP concentrations and has been widely used as a model enzyme to study the physical mechanisms of enzymatic catalysis [16–26]. The AdK consists of three domains, the core domain, the LID domain, and the NMP domain [Fig. 1(a)]. The core domain governs the overall stability of the enzyme, whereas the LID domain and NMP domain provide binding sites for the ATP (T) and AMP (M), respectively. After the chemical reaction, the same sites are occupied by two-product ADP (D) molecules. Hereafter, we refer to the two-substrate bound state and two-product bound state as chemical states “TM” and “DD,” respectively. At the apo condition, the AdK prefers an open conformation with the binding pockets exposed to solvent [Fig. 1(a), right]. With the binding of two substrates, it adopts a closed conformation at which the chemical reaction occurs [Fig. 1(a), left]. For this enzyme, product ADP release is the rate-limiting event in the enzymatic cycle, which is tightly coupled with the conformational motions [16,17]. Consequently, large-amplitude opening and closing conformational motions can be crucial for the enzymatic cycle of the AdK [16,17]. In previous works, the

\*wfli@nju.edu.cn

†wangwei@nju.edu.cn

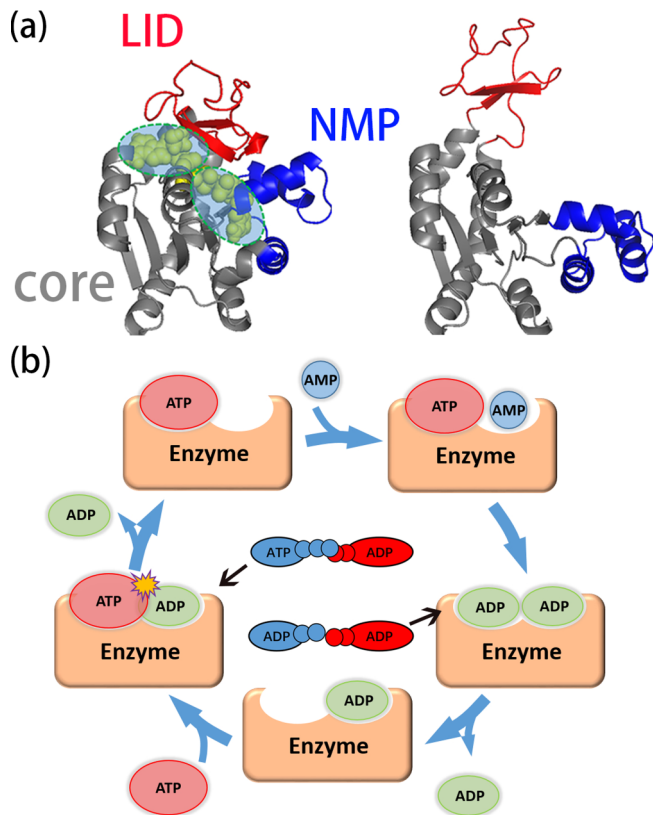


FIG. 1. Enzymatic cycle of the adenylate kinase (AdK). (a) Crystal structure of the AdK at the closed state (left, pdb code: 1ake) and open state (right, pdb code: 4ake). The core domain, LID domain, and the NMP domain are colored gray (left lower), red (top), and blue (right), respectively, in the cartoon structures. The atoms of the substrate molecules in the two neighboring binding pockets are shown by yellow (light gray) spheres. (b) Schematic diagram of the whole enzymatic cycle along the forward direction proposed in Ref. [14] based on a dynamic energy landscape model. At the chemical state with the ATP and ADP cobound, the terminal phosphate groups from the ATP and ADP are sterically incompatible, leading to local frustration in the binding pockets.

properties of the conformational motions of the AdK have been well characterized both computationally and experimentally [18–31]. For example, a coarse-grained molecular dynamics (MD) simulation with structure-based potential showed that the large-amplitude conformational change involves local unfolding of the enzyme, which reduces the free energy barrier for the conformational change [23]. By using atomistic MD simulations, Bagchi and co-workers showed that there exists a half-open half-closed (HOHC) conformation, at which the substrate binding sites can be partially exposed to solvent [26]. Therefore, the substrate-product exchange may occur without fully opening the enzyme conformations. In addition, the free energy landscape and the conformational dynamics at apo state and holo state have been extensively characterized by MD simulations [21,22,24–26,29]. More recently, we showed in a theoretical work with a dynamic energy landscape model that the substrate ATP dominantly binds to the LID domain site before the rate-limiting release of the product ADP at the NMP site, forming a hybrid

complex with substrate-product cobound (i.e., chemical state “TD”) [14] [Fig. 1(b)]. Compared to the canonical chemical states TM and DD, there is an additional phosphate group in the hybrid chemical state TD, which cannot be precisely accommodated by the binding pocket, leading to steric frustration due to conflicting interactions in the pockets [Fig. 1(b)]. Here the steric frustration represents the steric incompatibility of the two charged terminal phosphate groups of the ATP and ADP in the binding pocket at the chemical state TD. Such local frustration can result from the electrosteric repulsion interactions, including the steric effect and electrostatic repulsion.

The resulting substrate-product frustration was predicted to be able to facilitate the bottleneck product release, and therefore speed up the enzymatic cycle [14]. However, the microscopic mechanism of how such local frustration contributes to the conformational dynamics and product release, which are the key physical steps of the enzymatic cycle, is unclear.

In this work, by performing atomistic MD simulations with enhanced sampling, we investigate the underlying microscopic mechanism for the role of the local frustration playing in the conformational motions and rate-limiting product release of the AdK. The results showed that in the canonical DD state, the side chain of the Arg156 forms a stable hydrogen bond with the ADP phosphate groups at the NMP site, which is helpful for the stabilization of the LID-NMP interdomain contacts, particularly the side chain hydrogen bond between the Lys157 and the Asp54, whereas at the hybrid chemical state TD, the electrosteric repulsion arising from the additional phosphate group of the ATP perturbed the location of the phosphate groups of the neighboring ADP, leading to local frustration in the binding pockets. Consequently, the local hydrogen bond network of the binding pocket is disrupted, as featured by the breaking of the hydrogen bonds Arg156-ADP (NMP) and Lys157-Asp54, and the formation of a new hydrogen bond between the Arg156 and the terminal phosphate group of the substrate ATP. The hydrogen bonds Arg156-ADP (NMP) and Lys157-Asp54 have the key contribution to the LID-NMP interdomain interactions, the ruptures of which cause destabilization of the closed conformation of the AdK. Our results suggest that the local frustration can disrupt the hydrogen bond network around the binding pockets and results in lowered barrier heights for the opening of the enzyme conformations and the expedited release of the bottleneck product ADP.

## II. RESULTS AND DISCUSSION

### A. Role of the substrate-product frustration on the free energy landscape

After the chemical reaction step, the two lids (i.e., LID and NMP domains) of the AdK need to open to some extent in order to enable the product release and the new substrate uptake. To understand how the local frustration encountered in the substrate-product cobound complex affects the conformational opening, we constructed the two-dimensional free energy landscapes for the chemical states DD and TD by performing two-dimensional umbrella sampling (US) simulations [32] in combination with the Hamiltonian replica

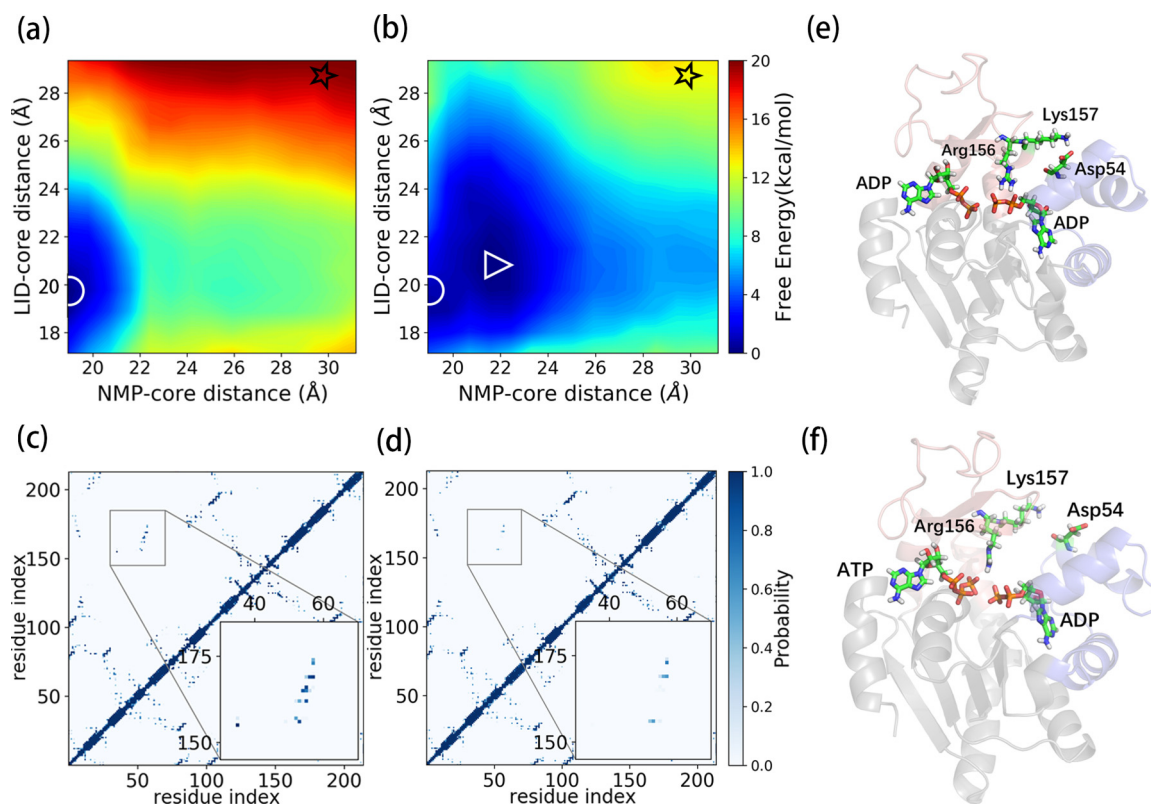


FIG. 2. Local frustration reduces the free energy barrier for the opening of the enzyme conformations. (a,b) Two-dimensional free energy landscapes along the LID-core distance and NMP-core distance at the chemical states DD (a) and TD (b). The white circles, white triangle, and black stars label the location of the closed conformation, partly open conformation, and fully open conformation, respectively. The closed conformation [white circle in panel (a)] corresponds to the global minimum of the free energy profile at the chemical state DD, whereas the partly open conformation [white triangle in panel (b)] corresponds to the global minimum of the free energy profile at the chemical state TD. (c,d) Probabilities of contact formation between two residues of the AdK for the conformations sampled around the global minimum of the free energy profiles at the chemical states DD (c) and TD (d). The residue pairs with large difference of contact probabilities are labeled by black squares, and the enlarged plots are shown in the bottom right of the panel (c,d) for clarity. (e,f) Representative structures of the conformations corresponding to the free energy minima at the chemical states DD (e) and TD (f).

exchange MD [33–35] (see Materials and Method). Figure 2 shows the free energy profiles along the LID-core distance and NMP-core distance for the chemical states DD (a) and TD (b). One can see that the LID domain can fluctuate to a large extent around the closed state (white circles) at both the canonical chemical state DD and the hybrid chemical state TD, suggesting large conformational flexibility of the AdK. Such conformational flexibility makes the substrate-product exchange possible even without fully opening of the AdK lids by exposing the binding sites to solvent, which is consistent with the discussions in Refs. [14,26,30]. In addition, at the chemical state TD, both the LID domain and the NMP domain have larger fluctuations, and there exists an additional free energy minimum with the NMP-CORE distance of 22 Å and LID-CORE distance of 21 Å (white triangle), at which the LID domain and the NMP domain are partly open. Such partly open conformation corresponds to the global minimum of the free energy profile at the chemical state TD and highly resembles the HOHC state observed by Bagchi and co-workers by MD simulations at the apo state [26]. In addition, the free energy landscapes show an overall uphill-like feature for the conformational opening at the chemical states DD and TD. Therefore the free energy

penalty for the conformational change from the closedlike state with the lowest free energy to the fully open state (black stars) gives the free energy barrier height. From Figs. 2(a) and 2(b), we can see that full opening of the two lids involves a high free energy barrier at both the DD and TD states, and therefore is slow. Interestingly, the free energy barrier at the chemical state DD ( $\sim 20.0$  kcal/mol) is much higher than that at the chemical state TD ( $\sim 13.0$  kcal/mol), suggesting that the local frustration in the binding pocket at the chemical state TD reduces the free energy barrier height. The above results indicate that the substrate-product frustration can make the substrate-product exchange easier, therefore contributing to the enzymatic cycle.

To investigate the effects of the substrate-product frustration on the local structure of the binding pockets, we calculated the contacting probabilities between each pair of the residues for the conformations sampled in the free energy basin of the closed state (white circle) at the chemical state DD and in the free energy basin of the partly open state (white triangle) at the chemical state TD. These two states correspond to the global free energy minima of the DD and TD states, respectively. Two residues were considered to form a contact if the distance of the closest heavy atom pair from



the two residues is less than 4.0 Å. Our result showed that despite the overall contact patterns being similar, there are significant differences between the residues indicated by the black squares. Some strong contacts that formed in the DD state cannot be observed at the TD state. In these contacts, the polar interactions between the oppositely charged residue pairs, including the (Asp54, Lys157), (Lys57, Glu170), and (Arg36, Asp158), have large contributions. The representative structures of the sampled conformations taken from the free energy minima at the DD and TD states are shown in Figs. 2(c) and 2(d), respectively. One can see that at the DD state, the side chain of the Arg156 forms a hydrogen bond with the phosphate groups of the ADP at the NMP domain site. Meanwhile, the side chains of the Lys157 and Asp54 form a strong hydrogen bond (part of the salt bridge). These neighboring hydrogen bonds represent the key interactions between the LID and NMP domains, which contribute to the stabilization of the closed conformation. Particularly, the importance of the formation and rupture of the salt bridges between the LID domain and the NMP domain for the conformational transitions has been discussed in several previous works based on computational simulations [36–38]. For example, Beckstein *et al.* showed that the opening-closing conformational transitions follow the zipping-unzipping mechanism, with the formation and rupture of the salt bridges between the LID-NMP domains dominate the kinetics of the transition [36]. In addition, the dissociation of the salt bridges corresponds to the central barrier of the potential of mean force for the conformational motions. The stable hydrogen bond network formed by the above hydrogen bonds at the closed conformation suggests the consistency of the interactions in the binding pockets [Fig. 2(e)]. Strikingly, at the TD state, the positions of the phosphate groups of the ADP were perturbed due to the presence of an additional phosphate group of the ATP at the LID domain site [Fig. 2(f)], leading to local frustration in the binding pockets featured by the incompatible binding of the two terminal phosphate groups from the ADP and ATP. As will be discussed later in this paper, both the steric effect and electrostatic repulsion can contribute to the local frustration. Meanwhile, the side chain of the Arg156 forms a hydrogen bond with the terminal phosphate group of the ATP. As a result, the hydrogen bond Arg156-ADP (NMP) mentioned above was disrupted, which is accompanied by the rupture of the hydrogen bond Lys157-Asp54. Because the above two hydrogen bonds have the key contributions to the interactions between the LID and NMP domains, the closed conformation can be largely destabilized by the breaking of these hydrogen bonds. The above results suggest that the local frustration encountered in the substrate-product cobound state tends to disrupt the local hydrogen bond network, which in turn destabilizes the closed conformation, leading to the lowered free energy barrier for the conformational opening.

### B. Local frustration leads to enhanced exposure of the binding pockets

The disruption of the binding pockets by the local frustration can also be demonstrated by comparing the solvent accessible surface area (SASA). We calculated the SASA of the ADP binding site at the NMP domain, which is

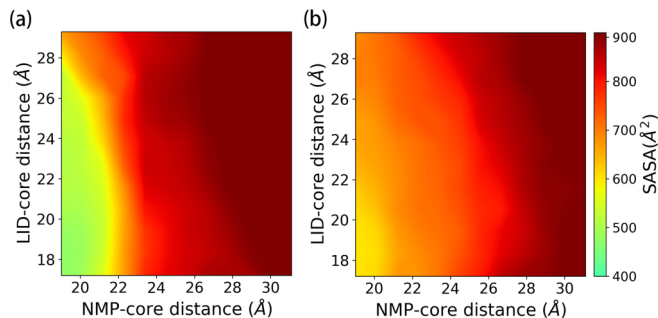


FIG. 3. Averaged solvent accessible surface area of the ADP binding pocket at the NMP domain along the two-dimensional conformational space formed by the LID-core distance and the NMP-core distance at the chemical states DD (a) and TD (b).

represented by the sum of the SASA for all the residues within the binding pocket using the POPS algorithm [39]. Figure 3 shows the SASA values at the chemical states DD (a) and TD (b). As expected, with the increasing of the interdomain distances, the binding site becomes more exposed as suggested by the increased SASA values. Interestingly, even at the same interdomain distances, the SASA can have a large difference between the chemical states DD and TD. At a wide range of conformations, the SASA at the TD state is significantly larger than that at the DD state. Such results indicate that the substrate-product frustration makes the binding pocket more flexible, such that even the otherwise deeply buried ADP site of the NMP domain can be partly exposed at the TD state. The expanded NMP site at the TD state suggests that the product ADP at the NMP domain can dissociate much easier due to the local frustration induced by the binding of the new substrate ATP, enabling expedited product release.

### C. Substrate-product frustration facilitates the conformational opening

In order to more directly characterize the role of the substrate-product frustration on the conformational opening, we investigate the conformational transitions from the closed state to the open state. Since the closed-to-open transition is a rare event, directly sampling the transition events by using conventional MD simulations is extremely difficult. In this work, we employed the parallel cascade selection molecular dynamics simulations (PaCS-MD) developed by Harada and Kitao [40] to sample the transition event (see Materials and Method). A similar method has been successfully used in Ref. [41] to investigate the pathways of some key steps in the enzymatic cycle of the AdK. Figure 4 shows the minimal root mean square deviation (RMSD) values to the open state as a function of the PaCS-MD cycle at the chemical states DD (a) and TD (b). Different colors represent ten different PaCS-MD trajectories. We can see that at the DD state, all of the PaCS-MD trajectories stay at the closedlike conformation at the end of the 150 cycles (RMSD > 3.5 Å). In comparison, in the TD state, the final structures are much more open. Particularly, some trajectories can successfully sample the conformations similar to the fully open conformation (RMSD < 3.0 Å). These results clearly demonstrated that the substrate-product

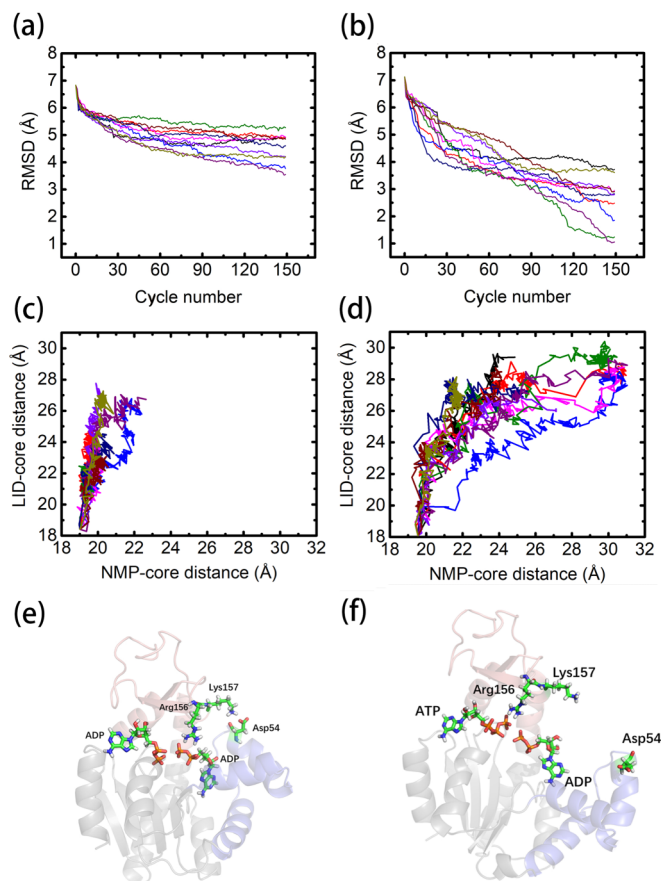


FIG. 4. Local frustration speeds up the opening of the enzyme conformation. (a,b) The minimal RMSD from the open structure as a function of the PaCS-MD cycles at the chemical states DD (a) and TD (b). Ten independent PaCS-MD trajectories are shown. (c,d) The PaCS-MD trajectories projected onto the two-dimensional conformational space formed by the LID-core distance and the NMP-core distance at the chemical states DD (c) and TD (d). (e,f) Representative structures obtained after 150 PaCS-MD cycles at the chemical states DD (e) and TD (f).

frustration encountered in the TD state facilitates the conformational opening.

Based on the PaCS-MD trajectories, we can analyze the pathways of the closed-to-open transitions. At the DD state, the conformational change mainly occurs at the LID domain, and the motions of the NMP domain are very minor [Fig. 4(c)], which is consistent with the results shown in the free energy landscape. At the TD state, the motions of the LID domain around the closed conformation are similar to that of the NMP domain [Fig. 4(d)]. However, with the further progression of the PaCS-MD cycles, the two domains move toward the open conformations cooperatively, during which the motions of the NMP domain are more significant. Among the ten PaCS-MD trajectories, five trajectories arrived at the fully open conformation featured by the large values of the two interdomain distances. We plotted the representative structures at the end of the PaCS-MD simulations for the DD state and TD state. One can see that the hydrogen bonds Arg156-ADP (NMP) and Lys157-Asp54 remain formed at the end of the PaCS-MD simulations for the DD state [Fig. 4(e)],

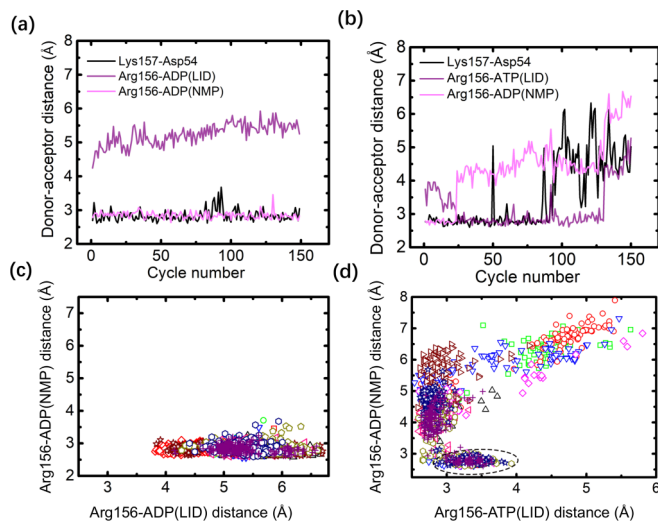


FIG. 5. Pathways of hydrogen bond exchange. (a,b) The donor-acceptor distances for the three key hydrogen bonds as a function of the PaCS-MD cycles for a representative trajectory at the chemical states DD (a) and TD (b). Black: hydrogen bond formed between the side chain of the Lys157 and the side chain of the Asp54; purple (gray): hydrogen bond formed between the side chain of the Arg156 and the terminal phosphate of the ADP (a) or ATP (b) at the LID domain; light magenta (light gray): hydrogen bond formed between the side chain of the Arg156 and the terminal phosphate of the ADP at the NMP domain. (c) Trajectories projected onto the two-dimensional space formed by the Arg156-ADP (NMP) distance and the Arg156-ADP (LID) distance for all ten simulations at chemical state DD. (d) Trajectories projected onto the two-dimensional space formed by the Arg156-ADP (NMP) distance and the Arg156-ATP (LID) distance for all ten simulations at chemical state TD. Different trajectories are represented by different symbols and colors. The dashed ellipse at panel (d) represents the snapshots at the initial stage of the PaCS-MD trajectories.

whereas in the final structure of the TD state with fully open conformation, both the hydrogen bonds Arg156-ADP (NMP) and Lys157-Asp46 were broken, and the overall structure was very close to the open structure at the apo state [Fig. 4(f)].

The above results suggest that the residue Arg156 plays a crucial role for the local frustration induced conformational transitions. It is interesting to investigate how the interactions involving the Arg156 evolve along the conformational transitions. Figure 5 shows the donor-acceptor distances for the above discussed hydrogen bonds. One can see that during the whole trajectory, the donor-acceptor distances have no significant changes at the chemical state DD [Fig. 5(a)], whereas for the TD state, at the early stage of the PaCS-MD simulations, the donor-acceptor distance of the Arg156-ADP (NMP) hydrogen bond has a sharp increase [Fig. 5(b)], suggesting the rupture of the bond. At the same time, the donor-acceptor distance of the Arg156-ATP hydrogen bond has a sharp decrease, indicating the formation of the Arg156-ATP hydrogen bond. Such different behaviors between the DD state and TD state simulations suggest that the presence of the additional phosphate group in the ATP tends to sequester the hydrogen bond donor of the Arg156, therefore disrupting the hydrogen bond Arg156-ADP (NMP). In addition, we showed the

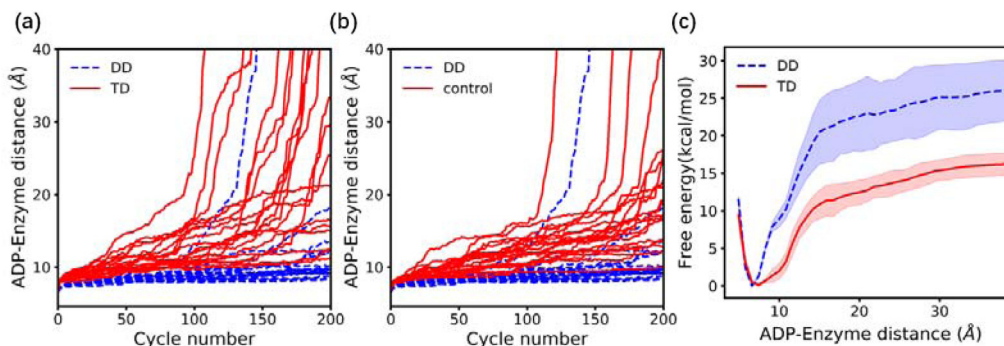


FIG. 6. (a) Distances between the centers of mass of the enzyme and the product ADP at the NMP domain as a function of the PaCS-MD cycles at the chemical states DD (blue) and TD (red) for 20 independent simulations at 315 K; (b) same as (a) but with the electrostatic interactions involving the terminal phosphate group of the ATP being switched off at the chemical state TD; (c) PMF profiles along the distance between the centers of mass of the ADP and the enzyme at the chemical states DD (blue dashed lines) and TD (red solid lines) at 315 K. The errors at the TD state represent the standard deviations from five independent US simulations. The errors at the DD state represent the differences from two independent US simulations.

trajectory of the donor-acceptor distance for the Lys157-Asp54 hydrogen bond, which fluctuates largely after the hydrogen bond Arg156-ADP (NMP) was broken and then encountered a drastic increase, suggesting the rupture of the hydrogen bond. We also plotted the trajectories along the two-dimensional space formed by the above two donor-acceptor distances for all ten simulations [Figs. 5(c) and 5(d)]. The results showed that all ten simulations have similar pathways of the hydrogen bond exchange at the TD state, in which the Arg156-ATP hydrogen bond forms after the rupture of the Arg156-ADP hydrogen bond. In five of the fully opening trajectories, one can also observe the rupture of the Lys157-Asp54 hydrogen bond. These trajectories support the above discussions that disruption of the local hydrogen network around the binding pocket induced by the local frustration contributes to the lowered free energy barrier and the expedited conformational opening.

#### D. Substrate-product frustration speeds up the bottleneck product release

In the catalytic cycle of the AdK, the release of the product ADP is the rate-limiting step. To demonstrate the effect of the frustration on the product release, we conducted PaCS-MD simulations to sample the ADP release event. The distance between the center of mass (COM) of the protein and the COM of the product ADP at the NMP site was used to guide the PaCS-MD simulations. We found that starting from the closed state, it is difficult to successfully sample the ADP dissociation event, mostly because of the strong attractive interactions between the  $Mg^{2+}$  and phosphate group. Therefore, we applied harmonic repulsion potentials when the distances between the  $Mg^{2+}$  and the heavy atoms of the terminal phosphate group of the ADP are closer than  $3.5 \text{ \AA}$  to prevent the  $Mg^{2+}$ -ADP rebinding events in the PaCS-MD simulations. In addition, to accelerate the ADP release events, the simulations were conducted at 315 K. Figure 6(a) shows the ADP-enzyme distances as a function of the PaCS-MD cycles for the chemical states DD (blue) and TD (red). One can see that even without considering the  $Mg^{2+}$  mediated interactions, the product dissociation behaviors of the chemical states DD and TD have a large difference. Compared to the DD state,

the product ADP at the TD state dissociates much easier. For example, at the TD state, the product ADP can fully dissociate from the binding pocket (with the ADP-enzyme distance exceeding  $40 \text{ \AA}$ ) in ten of the 20 PaCS-MD trajectories at the end of the 200 cycles. In comparison, there is only one successful dissociation trajectory in the simulations at the DD state. The above results suggest that the local frustration at the TD state leads to the enhanced product release. Such local frustration at the binding pocket may arise from steric effect and electrostatic repulsion. To provide a deeper understanding to the above effect, we conducted control simulations at the chemical state TD, in which the electrostatic interactions involving the terminal phosphate group were switched off: therefore, only the steric effect was included. As shown in Fig. 6(b), the product ADP can fully dissociate from the binding pocket in four of the 20 PaCS-MD trajectories in the control simulations, which demonstrates relatively weaker effect of the local frustration compared to that of the intact simulations at the chemical state TD, suggesting that both the steric effect and the electrostatic repulsion have important contributions to the local frustration.

By using the same reaction coordinate as that in the PaCS-MD simulations, we performed one-dimensional US simulations to construct the free energy profiles for the product ADP dissociation at 315 K, and the results are shown in Fig. 6(c). The snapshots sampled along the PaCS-MD trajectories were used as the initial structures for the US simulations. For the TD state, we conducted five independent US simulations with the initial structures taken from five different PaCS-MD trajectories, with which we can estimate the errors of the PMF by calculating the standard deviations. For the DD state, since there is only one successful ADP dissociation trajectory within 200 cycles, we elongated one of the PaCS-MD trajectories to get another successful ADP dissociation event. Starting from the snapshots from the above two trajectories, we performed two independent US simulations at the DD state, and the difference of the PMFs was considered as the errors. One can see that the free energy penalties for the separation of the product ADP from the binding site at the chemical state TD are much lower than that at the chemical state DD. These results directly demonstrate that the local frustration at the



binding pocket arising from the additional phosphate group of the ATP at the chemical state TD facilitates the bottleneck product ADP release.

### III. CONCLUSION

The high catalytic efficiency of natural enzymes relies on the ability to facilitate the rate-limiting step of the enzymatic cycle. In this work, for a model multisubstrate enzyme AdK, we showed that the local frustration encountered in the substrate-product cobound complex, which can be dominantly populated during the enzymatic cycle as demonstrated in the previous theoretical work [14], tends to disrupt the local hydrogen bond network in the binding pockets, which in turn leads to a lowered free energy barrier for the domain opening. In addition, by using an enhanced sampling technique, we directly showed that the local frustration speeds up the conformational opening and the product release. All these results suggest that nature enzymes can use frustration to facilitate functions, consistent with the previous theoretical prediction by a dynamic energy landscape model [14]. Interestingly, a recent work based on a statistical survey showed that the active sites of the enzymes are often locally frustrated [6]. The role of the substrate-product frustration on the enzyme conformational motions and product release is in line with the observations of the statistical survey [6]. Differently, here the frustration arises from the incompatible interactions between the substrate and product, instead of the conflicting interactions from the residues of the binding pockets. The results of this work not only provide numerical evidence to the predictions of the previous theoretical work on the role of the local frustrations in the enzymatic cycle, but also provide insights into the microscopic mechanism on the interplay between the local frustration of proteins and their functions.

### IV. MATERIALS AND METHOD

#### A. Details of MD simulations

The initial structure of the MD simulations for the AdK was taken from the crystal structure at the closed conformation (PDB code: 1ake) [42]. The coordinates of the substrate and product molecules were extracted based on the structure of the AP5A molecule in the pdb file. The missing oxygen atoms of the phosphate group in the chemical state TD were added according to the template in the AMBER14 package [43]. As the coordinate of the  $Mg^{2+}$  is lacking in the pdb file, we added it by using the local coordinate of the  $Mg^{2+}$  in the crystal structure (PDB code: 1zio) given in Ref. [44]. The complex structures at the chemical state DD and TD were solvated in a truncated octahedral water box containing  $\sim 8000$  TIP3P water molecules [45]. The AMBER force field ff14SB was used for the protein [46]. The ADP and ATP force field parameters were taken from Ref. [47]. Sodium ions were added to neutralize the simulation system. The particle-mesh Ewald summation [48] was used in calculating the long-range electrostatic interaction and the cutoff was set as 9 Å. The SHAKE algorithms were used to restrain the covalent bond with hydrogen atoms, which enables an integration time step of 2.0 fs.

The simulation system was firstly minimized for 20 000 steps using the steepest descent method. During the

minimization, the harmonic positional restraints were applied to all of the atoms with the force constant of  $100 \text{ kcal/mol/\AA}^2$ . Then the system was heated from 100 to 298 K within 40 ps with the force constant of  $50 \text{ kcal/mol/\AA}^2$ . In the subsequent relaxation simulations, the restraints were gradually reduced within 200 ps. The system was further relaxed for 20 ps in the *NPT* ensemble at the pressure of 1.0 atm without positional restraints. The resulting system was then used for the subsequent enhanced sampling simulations. We used PYMOL [49] and VMD [50] software for the structure visualization.

#### B. 2D H-REMD-US simulations

In order to investigate the role of the substrate-product frustration on the free energy landscape, we performed two-dimensional umbrella sampling (US) simulations [32] in combination with the Hamiltonian replica exchange MD (H-REMD) method [33–35]. This simulation scheme has been successfully used in calculating the free energy landscapes of the AdK under various substrate binding conditions in Ref. [19]. The US simulations were performed in the *NVT* ensemble at 298 K. The LID-core interdomain distance and NMP-core interdomain distance were used as the collective variables for the US simulations. The LID-core distance was represented by the distance between the center of mass (COM) of the  $C_\alpha$  atoms of the residues 1–28 in the core domain and the center of mass of the  $C_\alpha$  atoms of the residues 125–152 in the LID domain. Similarly, the NMP-core distance was represented by the distance between the COM of the  $C_\alpha$  atoms of the residues 112–121, 160–175 in the core domain and the COM of the  $C_\alpha$  atoms of the residues 33–58 in the NMP domain. The harmonic potentials were applied to the windows along the NMP-core distance ranging from 19.0 Å to 31.1 Å and along the LID-core distance ranging from 17.2 to 29.3 Å with a spacing of 1.1 Å, leading to a total of 288 windows. The force constant of the harmonic umbrella potential was set as  $2 \text{ kcal/mol/\AA}^2$ . In preparing the initial structures of the US simulations at each window, the equilibrated structure obtained in the previous subsection was stretched by a harmonic potential from the closed to the open conformation. Then we generated a second set of initial structures by a similar harmonic potential but from the open to closed conformation following the scheme used in Ref. [19]. Two sets of US sampling were run in parallel with the above two sets of initial structures. To reduce the dependence of the US sampling on the initial structures, the simulation system at the neighboring window within and between the two sets of US simulations were allowed to exchange configurations with the attempt interval of 1.0 ps. The length of the simulations at each window was 6 ns, leading to an accumulation time of 1.7  $\mu\text{s}$ . The resulting successful exchange rates were between 0.05 and 0.20. The sampled conformations of the 2D H-REMD-US simulations were used to construct the two-dimensional free energy landscapes by using the weighted histogram analysis method [51].

#### C. Parallel cascade selection molecular dynamics simulations (PaCS-MD)

PaCS-MD is an enhanced sampling method without applying biasing potential proposed by Harada and Kitao [40].

A similar approach called the weighted ensemble simulation toolkit with parallelization and analysis was also proposed by Zuckerman and Chong [52]. It can be used to sample the conformational transition pathways between two preset structures by using a genetic-type algorithm. In this work, in order to sample the rare events of the conformational transitions from the closed conformation to the open conformation, we used the PaCS-MD for the AdK at the chemical states DD and TD. The simulations consisted of the following steps: (1) Starting from the closed structure, we ran a short MD simulation with the length of 100 ps. The snapshots of the short trajectory were saved with the interval of 1 ps. (2) From the sampled structures, we chose the top ten structures with the smallest RMSD to the open structure, and ran ten independent short MD simulations starting from the chosen structures. We repeated steps (2) and (3) for 150 cycles, and ten independent PaCS-MD simulations were carried out for each of the two chemical states at 298 K. We also conducted the PaCS-MD simulations to sample the dissociation event of the product

ADP of the NMP site from the ADP-ADP bound state and the ATP-ADP bound state at 315 K. The COM distance between the protein and the ADP at the NMP site was used to guide the PaCS-MD simulations. We performed 200 cycles for each of the PaCS-MD simulations and 20 independent PaCS-MD simulations were carried out for each of the chemical states DD and TD. For simplicity, the positions of the ATP/ADP and  $Mg^{2+}$  at the LID site were restrained to the positions in the initial structure.

#### ACKNOWLEDGMENTS

We thank Shoji Takada for the stimulating discussions and Zhiyong Zhang for expert assistance in implementing the PaCS-MD sampling method. This work was supported by the National Natural Science Foundation of China (Grants No. 11574132, No. 11974173, and No. 11934008), the HPC center of Nanjing University, and the HPC center of Collaborative Center of Advanced Microstructures.

- 
- [1] J. D. Bryngelson, J. N. Onuchic, N. D. Socci, and P. G. Wolynes, *Proteins* **21**, 167 (1995).
- [2] J. N. Onuchic, Z. Luthey-Schulten, and P. G. Wolynes, *Annu. Rev. Phys. Chem.* **48**, 545 (1997).
- [3] D. U. Ferreira, E. A. Komives, and P. G. Wolynes, *Q. Rev. Biophys.* **47**, 285 (2014).
- [4] W. Li, P. G. Wolynes, and S. Takada, *Proc. Natl. Acad. Sci. USA* **108**, 3504 (2011).
- [5] D. U. Ferreira, J. A. Hegler, E. A. Komives, and P. G. Wolynes, *Proc. Natl. Acad. Sci. USA* **108**, 3499 (2011).
- [6] M. I. Freiberger, A. Brenda Guzovsky, P. G. Wolynes, R. Gonzalo Parra, and D. U. Ferreira, *Proc. Natl. Acad. Sci. USA* **116**, 4037 (2019).
- [7] A. Das and S. S. Plotkin, *Proc. Natl. Acad. Sci. USA* **110**, 3871 (2013).
- [8] I. Segel, *Enzyme Kinetics* (Wiley, New York, 1993).
- [9] E. D. Watt, H. Shimada, E. L. Kovrigin, and J. P. Loria, *Proc. Natl. Acad. Sci. USA* **104**, 11981 (2007).
- [10] S. J. Benkovic and S. Hammes-Schiffer, *Science* **301**, 1196 (2003).
- [11] W. W. Cleland, *Acc. Chem. Res.* **8**, 145 (1975).
- [12] A. V. Pislakov, J. Cao, S. C. Kamerlin, and A. Warshel, *Proc. Natl. Acad. Sci. USA* **106**, 17359 (2009).
- [13] A. Ariyaratne, C. Wu, C. Y. Tseng, and G. Zocchi, *Phys. Rev. Lett.* **113**, 198101 (2014).
- [14] W. Li, J. Wang, J. Zhang, S. Takada, and W. Wang, *Phys. Rev. Lett.* **122**, 238102 (2019).
- [15] C. Müller, G. Schlauderer, J. Reinstein, and G. E. Schulz, *Structure* **4**, 147 (1996).
- [16] K. A. Henzler-Wildman, M. Lei, V. Thai, S. J. Kerns, M. Karplus, and D. Kern, *Nature* **450**, 913 (2007).
- [17] J. A. Hanson, K. Duderstadt, L. P. Watkins, S. Bhattacharyya, J. Brokaw, J.-W. Chu, and H. Yang, *Proc. Natl. Acad. Sci. USA* **104**, 18055 (2007).
- [18] S. J. Kerns, R. V. Agafonov, Y. J. Cho, F. Pontiggia, R. Otten, D. V. Pachov, S. Kutter, L. A. Phung, P. N. Murphy, V. Thai *et al.*, *Nat. Struct. Mol. Biol.* **22**, 124 (2015).
- [19] B. Pelz, G. Zoldak, F. Zeller, M. Zacharias, and M. Rief, *Nat. Commun.* **7**, 10848 (2016).
- [20] H. Y. Aviram, M. Pirchi, H. Mazal, Y. Barak, I. Riven, and G. Haran, *Proc. Natl. Acad. Sci. USA* **115**, 3243 (2018).
- [21] M. D. Daily, G. N. Phillips, Jr., and Q. Cui, *J. Mol. Biol.* **400**, 618 (2010).
- [22] Q. Lu and J. Wang, *J. Am. Chem. Soc.* **130**, 4772 (2008).
- [23] P. C. Whitford, O. Miyashita, Y. Levy, and J. N. Onuchic, *J. Mol. Biol.* **366**, 1661 (2007).
- [24] M. B. Kubitzki and B. L. de Groot, *Structure* **16**, 1175 (2008).
- [25] Y. Matsunaga, H. Fujisaki, T. Terada, T. Furuta, K. Moritsugu, and A. Kidera, *PLoS Comput. Biol.* **8**, e1002555 (2012).
- [26] B. V. Adkar, B. Jana, and B. Bagchi, *J. Phys. Chem. A* **115**, 3691 (2011).
- [27] J. Aden, A. Verma, A. Schug, and M. Wolf-Watz, *J. Am. Chem. Soc.* **134**, 16562 (2012).
- [28] B. Jana, B. V. Adkar, R. Biswas, and B. Bagchi, *J. Chem. Phys.* **134**, 035101 (2011).
- [29] Y. Wang, L. Gan, E. Wang, and J. Wang, *J. Chem. Theory Comput.* **9**, 84 (2013).
- [30] M. Kovermann, C. Grundstrom, A. E. Sauer-Eriksson, U. H. Sauer, and M. Wolf-Watz, *Proc. Natl. Acad. Sci. USA* **114**, 6298 (2017).
- [31] P. Maragakis and M. Karplus, *J. Mol. Biol.* **352**, 807 (2005).
- [32] G. M. Torrie and J. P. Valleau, *J. Comput. Phys.* **23**, 187 (1977).
- [33] H. Fukunishi, O. Watanabe, and S. Takada, *J. Chem. Phys.* **116**, 9058 (2002).
- [34] Y. Sugita and Y. Okamoto, *Chem. Phys. Lett.* **314**, 141 (1999).
- [35] W. Nadler and U. H. E. Hansmann, *Phys. Rev. E* **75**, 026109 (2007).
- [36] O. Beckstein, E. J. Denning, J. R. Perilla, and T. B. Woolf, *J. Mol. Biol.* **394**, 160 (2009).
- [37] S. L. Seyler, A. Kumar, M. F. Thorpe, and O. Beckstein, *PLoS Comput. Biol.* **11**, e1004568 (2015).
- [38] H. D. Song and F. Zhu, *PLoS One* **8**, e68023 (2013).
- [39] L. Cavallo, J. Kleinjung, and F. Fraternali, *Nucleic Acids Res.* **31**, 3364 (2003).



- [40] R. Harada and A. Kitao, *J. Chem. Phys.* **139**, 035103 (2013).
- [41] C. Ye, C. Ding, R. Ma, J. Wang, and Z. Zhang, *Proteins* **87**, 337 (2019).
- [42] M. Bellinzoni, A. Haouz, M. Grana, H. Munier-Lehmann, W. Shepard, and P. M. Alzari, *Protein Sci.* **15**, 1489 (2006).
- [43] D. A. Case, V. Babin, J. Berryman, R. Betz, Q. Cai, D. Cerutti, T. Cheatham III, T. Darden, R. Duke, H. Gohlke *et al.*, AMBER 14 (University of California, San Francisco, 2014).
- [44] M. B. Berry and G. N. Phillips, *Proteins* **32**, 276 (1998).
- [45] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, *J. Chem. Phys.* **79**, 926 (1983).
- [46] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling, *Proteins* **65**, 712 (2006).
- [47] K. L. Meagher, L. T. Redman, and H. A. Carlson, *J. Comput. Chem.* **24**, 1016 (2003).
- [48] T. Darden, D. York, and L. Pedersen, *J. Chem. Phys.* **98**, 10089 (1993).
- [49] W. L. DeLano, *The PYMOL Molecular Graphics System* (DeLano Scientific, San Carlos, CA, 2009).
- [50] W. Humphrey, A. Dalke, and K. Schulten, *J. Mol. Graphics Modell.* **14**, 33 (1996).
- [51] S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman, and J. M. Rosenberg, *J. Comput. Chem.* **13**, 1011 (1992).
- [52] D. M. Zuckerman and L. T. Chong, *Annu. Rev. Biophys.* **46**, 43 (2017).