Quantifying the Intrinsic Conformation Energy Landscape Topography of Proteins with Large-Scale Open–Closed Transition

Wen-Ting Chu† and Jin Wang‡,†,‡

1 State Key Laboratory of Electroanalytical Chemistry, Changchun Institute of Applied Chemistry, Chinese Academy of Sciences, Changchun, Jilin 130022, China
2 Department of Chemistry & Physics, State University of New York at Stony Brook, Stony Brook, New York 11794, United States

ABSTRACT: Large-scale conformational changes of proteins, including the open–closed transitions, are crucial for a variety of protein functions. These open–closed transitions are often associated with ligand binding. However, the understandings of the underlying mechanisms of the conformational changes within proteins during the open–closed transitions are still challenging at present. In this study, we quantified the intrinsic underlying conformational energy landscapes of five different proteins with large-scale open–closed transitions. This is realized by exploring the underlying density of states and the intrinsic conformational energy landscape topography measure \( \Lambda \). \( \Lambda \) is a dimensionless ratio of conformational energy gap \( \delta E \) versus conformational energy roughness \( \delta E \) and configurational entropy \( S \) or size of the intrinsic conformational energy landscape. By quantifying the \( \Lambda \) of intrinsic open–closed conformational (\( \Lambda^{oc} \)) and intrinsic global folding (\( \Lambda^{glob} \)) energy landscapes, we show that both intrinsic open–closed conformation energy and entropy landscapes are funneled toward the closed state. Furthermore, our results indicate the strong correlations between \( \Lambda \) and thermodynamics (conformational state transition temperature against trapping temperature) as well as between \( \Lambda \) and kinetics (open–closed kinetic time) of these proteins. This shows that the intrinsic conformational landscape topography determines both the conformational thermodynamic stability and kinetic speed of the conformational dynamics. Our investigations provide important insights for understanding the fundamental mechanisms of the protein conformational dynamics in a physical and global way.

INTRODUCTION

Large-scale structural rearrangements in biomolecules have been observed in experiments, during the processes of ligand binding, catalysis, signal transduction, and regulation. These structural changes are often found to be critical to the functions of biomolecules. The intrinsic flexibility of the protein during binding is critical for uncovering the underlying mechanisms in kinetics and thermodynamics of the conformational changes.

By X-ray crystallography, NMR, EM, and other biophysical techniques, ligand-induced conformational changes from apo state (usually open) to ligand-bound state (usually closed) have been observed at the atomic level. In the “Database of Macromolecular Movements”, the conformational changes can be classified into five main types, referenced as predominantly shear, predominantly hinge, not hinge or shear, involving partial refolding of the structure, or unclassified. For ligand binding, the most often observed types of motions of the proteins belong to shear and hinge ones. However, understanding the underlying mechanisms of open–closed transitions, in particular the relationship between local flexibility and global flexibility of the proteins, is still challenging.

For biomolecules, large-scale conformational change within the molecule is a rate-limiting step. The time scales of protein motion (binding/unbinding, folding, etc.) span from femtosecond to beyond seconds. Though molecular dynamics (MD) simulation provides a good way to investigate the protein systems and give atomic structural information, it is rather time-consuming for conventional MD simulations to deal with the protein systems with large-scale conformational changes and effectively sample enough of the conformational states. The structure-based models based on the energy landscape theory have become a powerful tool for the studies of the mechanisms of protein folding and binding processes, which are often associated with large conformational changes. Recently, two-basin structure-based models with two reference states, and even multibasin structure-based models with multiple reference states, have been developed and successfully applied to many typical allosteric systems, for example, adenylate kinase (ADK, two-basin), DNA Y-family polymerase IV (DPO4, two-basin), glutamine-binding protein (GBP, two-basin), calmodulin (CaM, two-basin), maltose-binding protein (MBP, multibasin), and protein kinase A (PKA, multibasin), etc. By using these methods, simulation results can provide valuable information about energy, entropy, free energy, binding
constants, and other physical quantities from effective sampling of the conformational states.

Energy landscape theory has guided our understanding of biomolecules as well as their kinetic and thermodynamic processes. In the previous studies, by quantifying the underlying density of states (DOS), we have shown that the dimensionless ratio $\Lambda$ between the energy gap $\delta E$, energy roughness $\delta E$, and configurational entropy $S$ of the system \[ \Lambda = \delta E / (\Delta E_{\text{f}}/2S) \] quantifies the topography of the underlying energy landscape and measures the degree of its funneledness. It has been demonstrated that $\Lambda$, as the energy landscape topology measure, determines the thermodynamic stability and kinetic rate of the folding process of the proteins.\(^{29}\) For the protein-binding process, we have suggested that the topography of binding energy landscape $\Lambda$ (intrinsic specificity) also dictates with the thermodynamic and kinetic specificity.\(^{30-33}\) In addition, for the coupled binding--folding process, the entire energy landscapes have been proposed to be the combined one of the binding and folding energy landscapes if the binding and folding are weakly coupled in the recognition process.\(^{34}\) Upon quantification of the topology of the individual effective binding and folding landscapes, as well as the global binding--folding energy landscapes, landscape topology $\Lambda$ is shown again to govern both thermodynamic feasibility and kinetic binding--folding rate.\(^{35,36}\) Therefore, the $\Lambda$ will also be a valuable quantity to study the large-scale conformational change within proteins between open and closed states.

Here in this study, for the uncovering of the underlying relationship between the thermodynamics and kinetics and the topography of intrinsic energy landscape $\Lambda$ of protein conformational changes (between open and closed states), topography of the intrinsic local open--closed conformation and intrinsic global folding energy landscapes will be quantified for several different proteins. Five different proteins with reference structures of open and closed conformations were selected for this study: lysine/arginine/ornithine-binding protein (LAOBP),\(^{37,38}\) adenylate kinase (ADK),\(^{39}\) DNA Y-family polymerase IV (DPO4),\(^{40}\) lipase 1 (LIP1),\(^{40,41}\) and phosphonate-binding protein (PhnD).\(^{32,42}\) As shown in Figure S1, these proteins are with different sizes and different motions between open and closed states. We will quantify the intrinsic global energy landscape as well as the intrinsic local open--closed conformation energy landscape from the underlying DOS extracted from the conformational dynamics simulation, by applying the two-basin structure-based model. The topography of the conformational landscape for each individual protein can be represented by the dimensionless ratio $\Lambda$ between the energy gap between the native (closed) state and the average non-native states ($\delta E = E_n - (E_{\text{non-native}})$), the roughness of the conformation energy landscape or the width of the energy distribution of the non-native states ($\delta E$), and the size of the funnel measured by the configurational entropy of non-native states of the conformation energy landscape ($S$). We focus on the topography of both intrinsic local open--closed conformation and intrinsic global folding landscapes ($\Lambda^\text{local}$ and $\Lambda^\text{global}$), as well as the relationships between them and the thermodynamic and kinetic properties of the proteins. This study will be essential for uncovering the fundamental mechanisms of large-scale conformational changes between open and closed states within proteins, revealing the physical effects and significance of landscape topography measure $\Lambda$.

### RESULTS AND DISCUSSIONS

#### Intrinsic Energy and Entropy Landscapes Are Funneled toward the Basin at Closed State.

The intrinsic conformation energy landscape can be quantified by the density of states (DOS), a statistical energy distribution in microcanonical ensemble. With the help of the WHAM algorithm,\(^{43-45}\) the intrinsic conformation energy landscape is quantified by transforming the canonical ensemble representation to microcanonical representation. In general, the intrinsic energy landscape probes the underlying interactions and usually has a very weak dependence on the temperature. First, the energy landscape is illustrated by the energy spectrum directly through the zero dimensional projection of energy to itself, as shown in Figure S2A,B. In this study, both the intrinsic local conformation (open--closed) and intrinsic global folding (folded and unfolded) energy landscapes are shown in terms of the energy spectrum for the 5 different proteins. They both have only one native state basin (as illustrated in Figure S2C,D). The minimum of intrinsic energy landscape spectrum of the non-native ensemble is higher than that of the native ensemble. Our results show that LIP1 has the largest energy gap $\delta E$ in intrinsic local open--closed conformational and intrinsic global folding energy landscapes; DPO4 has the lowest energy gap $\delta E$ in intrinsic local open--closed conformation energy landscapes while ADK has the lowest energy gap $\delta E$ in intrinsic global folding energy landscapes (listed in Table S1). Figure S2C,D also demonstrates that the size of the intrinsic energy landscape measured by entropy through DOS decreases as energy goes down. This clearly shows a funnel toward native states. Meanwhile, many other valuable quantities, such as average energy, heat capacity, and free energy, can be obtained from DOS.

We further show the two-dimensional DOS by projecting it onto the fraction of native contacts $Q$ and energy $E$, illustrated in Figure 1. It is observed that the number of states decreases as either $Q$ increases or $E$ decreases. This indicates that the size of the intrinsic energy landscape measured by the number of states shrinks as the energy decreases and/or the state gets closer to the native with the lowest energy, as a funnel toward native state. It is obvious that closed states and folded states have the lowest DOS and stay at the bottom of the energy funnel. In addition, the transitions between open and closed states, as well as between folded and unfolded states, can be visualized in the heat capacity curves. As shown in Figure S3, all the 5 proteins have a higher folding--unfolding transition temperature ($T_{\text{trans}}^\text{global}$) and a lower open--closed transition temperature ($T_{\text{trans}}^\text{local}$). With the aim of further showing the details of these transitions, the fraction of native closed contacts $Q_{\text{closed}} > 0.8$ locate at the basin with the lowest free energy at lower temperature (Figure S5, temperature lower than $T_{\text{trans}}$), open states ($Q_{\text{closed}} < 0.2$ and RMSD$_{\text{open}} < 1.0$) at intermediate temperature (Figure S6, temperature between $T_{\text{trans}}$ and $T_{\text{global}}^\text{local}$), and unfolded states (RMSD$_{\text{open}} > 5.0$) at high temperature (temperature higher than $T_{\text{global}}^\text{global}$). The details of free energy analyses are included in the Supporting Information. However, both average intrinsic energy and DOS landscapes have only one basin that points to the closed state (Figure 2 and Figure S4). This shows a clear funneled
intrinsic energy landscape toward the native closed state. In general, free energy depends on the energy, entropy or DOS, and temperature \((G = H - TS)\). As a result, at low temperature, the effect of energy contributes to the closed basin on free energy landscape; at intermediate and high temperature, the effect of entropy leads to the open and unfolded basins on free energy landscape.

**Intrinsic Conformation Energy Landscape Topography \(\Lambda\) Determines the Thermodynamics of Protein Conformational Changes.** According to the energy landscape theory, we can quantify the topography of the intrinsic energy landscape by a dimensionless quantity \(\Lambda\), which can be calculated by \(\Lambda = \Delta E / (\Delta E \sqrt{2S})\) \(^{30,34,46-50}\). We analyzed the data of \(\Lambda\), as well as the important thermodynamic characteristics of protein conformational changes such as the glassy trapping temperature \(T_g\) and the conformation state transition temperature \(T_{trans}\) from the DOS of the intrinsic open–closed conformation and intrinsic global folding energy landscapes (details are referred to in the Supporting Information).

All the related data of the intrinsic energy landscapes and the thermodynamic characteristics are listed in Table S1. Different superscripts are used for different intrinsic energy landscapes, os for intrinsic open–closed conformation energy landscape, global for intrinsic global folding energy landscape. Of all the 5 proteins in our studies, ADK has the highest \(\Lambda^{oc}\), indicating the most funneled intrinsic open–closed conformation energy landscape and the strongest open–closed conformation stability against trapping. DPO4 has the lowest \(\Lambda^{os}\), which may be the one most easily trapped into non-native states. As shown in Figure 3, there are significant differences between the topography of the intrinsic open–closed conformation energy landscapes of ADK and DPO4. Though they have similar entropy \(S^{os}\), the energy gap \(\Delta E^{oc}\) of ADK landscape is much greater than that of DPO4 landscape, and the roughness \(\Delta E^{oc}\) of ADK landscape is much less than that of DPO4 landscape.

---

**Figure 1.** Logarithm of open–closed (A) and global (B) DOS of the 5 different proteins as a function of \(Q\) and energy. Here, \(Q\) is calculated as the fraction of native closed contacts for panel A and native global contacts for panel B.

**Figure 2.** Average energy of intrinsic local conformation energy landscape as a function of \(Q^{closed}\) and RMSD\(_{open}\) of LIP1 (A), DPO4 (B), LAOBP (C), ADK (D), and PhnD (E). Closed states are located at the right bottom part; open states are located at the left bottom part, and unfolded states are located at the left upper part.
In general, the intrinsic open−closed conformation energy landscape can be considered as a part of intrinsic global folding energy landscape, locating at the bottom of the intrinsic global folding funnel (see Figure 3). However, the significant difference between ADK and DPO4 in the intrinsic open−closed conformation energy landscapes does not exist in their intrinsic global folding energy landscapes. The $\Lambda_{\text{global}}$ of DPO4 is slightly higher than that of DPO4, with much greater energy gap $\delta E_{\text{global}}$ and entropy $S_{\text{global}}$.

To further analyze the 5 proteins individually, it should be noted that ADK has the lowest residue number as compared to the others but relatively high number of open−closed contacts (high $\text{NC}_{\text{oc}}/N$, see the details in the Supporting Information). The open−closed contact number of DPO4 is the lowest one of all. In the “Database of Macromolecular Movements”, 4 of the 5 proteins in our studies (LIP1, LAOBP, ADK, and PhnD) can be classified as “hinge motions”, except for the DPO4. DPO4 transfers from the “stable” open state (with native open contacts) via the “unstable” intermediate state (without native open or closed contacts) to the “stable” closed state (with closed states), which may be related to the relatively low $\Lambda_{\text{oc}}$ of DPO4. According to the open and closed structures, ADK

Figure 3. Intrinsic local open−closed conformation energy landscape as well as the intrinsic global folding energy landscape of ADK (blue and purple funnels, top part panel) and DPO4 (yellow and red funnels, bottom part panel). A simplified funnel is represented for each energy landscape. The intrinsic global and open−closed energy landscapes are located in the same coordinate system with different graph scales in the lower left corner, respectively. The depth of the funnel in the $z$ axis corresponds to the energy gap $\delta E$. The opening width of the funnel is described with the entropy $S$. For each intrinsic energy landscape, we use the number of the local (small) basins to roughly show the roughness $\delta E$. The script for generating this funneled energy landscape was used from https://oaslab.com/drawing_funnels.html.

Figure 4. Correlations between the intrinsic energy landscape topographic measure ($\Lambda_{\text{oc}}$ and $\Lambda_{\text{global}}$) and thermodynamic characteristics. (A) Correlation between the intrinsic open−closed landscape topography measure ($\Lambda_{\text{oc}}$) and open−closed transition temperature against open−closed glassy trapping temperature ($T_{\text{trans}}/T_{g}$). (B) Correlation between the intrinsic global landscape topography measure ($\Lambda_{\text{global}}$) and global transition temperature against global glassy trapping temperature ($T_{\text{trans}}/T_{g}$). The red solid line is the linear fitting results; the blue solid line is the analytical mean field theory prediction of the relationship between $T_{\text{trans}}/T_{g}$ and $\Lambda$: $T_{\text{trans}}/T_{g} = \Lambda + (\Lambda^2 - 1)^{1/2}$.
In the previous studies, $\Lambda_{\text{oc}}$ correlates with $\Lambda_{\text{fg}}$ or $\Lambda_{\text{global}}$ for the thermodynamic folding/stability against trapping (Figure S7). As mentioned above, $\Lambda_{\text{oc}}$ and $\Lambda_{\text{global}}$ are correlated with $T_{\text{oc}}/T_{\text{fg}}$ and $T_{\text{oc}}/T_{\text{global}}$, respectively. Therefore, $\Lambda_{\text{oc}}/\Lambda_{\text{global}}$ is shown to have positive correlation with $T_{\text{oc}}/T_{\text{global}}$ and $(T_{\text{oc}}/T_{\text{global}})^{-1}$ (Figure S7A,B). Likewise, according to the equation $\Lambda = \delta E/\sqrt{2S}$, $\Lambda_{\text{oc}}/\Lambda_{\text{global}}$ is suggested to have a positive correlation with $\delta E_{\text{oc}}/\delta E_{\text{global}}$ and $((S_{\text{oc}})^{1/2}/2\delta E_{\text{oc}}/((S_{\text{global}})^{1/2}/2\delta E_{\text{global}}))^{-1}$. Figure S7C,F). Though $\delta E_{\text{oc}}/\delta E_{\text{global}}$ and $((S_{\text{oc}})^{1/2}/2\delta E_{\text{oc}}/((S_{\text{global}})^{1/2}/2\delta E_{\text{global}}))$ have the opposite trends with $\Lambda_{\text{oc}}/\Lambda_{\text{global}}$, their product $((S_{\text{oc}})^{1/2}/2\delta E_{\text{oc}}/((S_{\text{global}})^{1/2}/2\delta E_{\text{global}}))$ has the reciprocal correlation with $\Lambda_{\text{oc}}/\Lambda_{\text{global}}$. The $\Lambda_{\text{oc}}/\Lambda_{\text{global}}$ shows the coupling effect between intrinsic open—closed and intrinsic folded—unfolded landscapes, but it has the same order with $\Lambda_{\text{oc}}$ (among the 5 proteins, $\Lambda_{\text{oc}}$ changes much greater than $\Lambda_{\text{global}}$).

The results indicate that, of the 5 proteins in this study, ADK has relatively more of a funneled intrinsic open—closed conformation landscape and less of a funneled intrinsic folded—unfolded landscape. In panels A and B, the green dashed line and blue solid line represent the non-exponential and exponential fitting results, respectively. In panels C–D, the red solid line shows the line fitting results.

Figure 5. Correlations between the properties of intrinsic open—closed conformation energy landscape topography and the kinetics of open—closed transitions. (A) Logarithm of mean first passage time from open to closed states ($\ln\langle\tau^q\rangle$) correlates with the intrinsic open—closed landscape topography measure ($\Lambda_{\text{oc}}$). (B) Logarithm of the standard deviation of $\tau^q$ ($\ln\sqrt{\langle\tau^q\rangle^2 - \langle\tau^q\rangle^2}$) correlates with $\Lambda_{\text{oc}}$. (C) Second-order moments of $\tau^q$ ($\langle\tau^q\rangle^2/\langle\tau^q\rangle^2$) correlate with roughness $\delta E_{\text{oc}}$ of intrinsic open—closed conformation energy landscape. (D) $\langle\tau^q\rangle/\langle\tau^q\rangle^2$ correlates with configuration entropy $S$ of intrinsic open—closed conformation energy landscape. In panels A and B, the green dashed line and blue solid line represent the non-exponential and reciprocal fitting results, respectively. In panels C–D, the red solid line shows the line fitting results.

has two hinges (one hinge for LID domain and one hinge for NMP domain) while the other 3 proteins (LIP1, LAOBP, and PhnD) only have one (see Figure S17). From the open to the closed state, ADK will go through a “relatively stable” intermediate state with one domain closed, which may be linked with the highest $\Lambda_{\text{oc}}$ of all the 5 proteins.

In simplified analytical models, $\Lambda$ or $\Lambda_{\text{fg}}$ or $\Lambda_{\text{global}}$, the thermodynamic stability against trapping of the global binding—folding can be determined by the whole intrinsic global binding—folding landscape topographic measure with $T_{\text{fg}}/T_{\text{g}} = \Lambda + (\Lambda^2 - 1)^{1/2}$. In the previous studies, $\Lambda_{\text{oc}}$ or $\Lambda_{\text{global}}$ has been shown to be correlated with the thermodynamic folding(binding)/stability against trapping temperature ($T_{\text{oc}}/T_{\text{g}}$ or $T_{\text{global}}/T_{\text{g}}$). Likewise, $\Lambda_{\text{global}}$ correlates with $T_{\text{fg}}/T_{\text{g}}$ or $T_{\text{global}}/T_{\text{g}}$ in our studies (as shown in Figure 4, $T_{\text{fg}}/T_{\text{g}}$ and $T_{\text{global}}/T_{\text{g}}$ are the folding transition temperature and glassy trapping temperature of global DOS). High $\Lambda_{\text{global}}$ and $T_{\text{fg}}/T_{\text{g}}$ or $T_{\text{global}}/T_{\text{g}}$ values suggest a funneled intrinsic global folding energy landscape and strong folding stability. Intriguingly, our results demonstrate the strong correlation between open—closed conformation energy topography measure $\Lambda_{\text{oc}}$ and conformation transition temperature against trapping $T_{\text{oc}}/T_{\text{g}}$. ($T_{\text{oc}}/T_{\text{g}}$ and $T_{\text{global}}/T_{\text{g}}$ are the open—closed transition temperature and glassy trapping temperature of open—closed DOS). The intrinsic energy landscape topographic measure $\Lambda$ is a valuable quantity in that it can reflect the effect of energy gap $\delta E$, roughness $\delta E$, and entropy $S$, as well as the effect of the transition temperature $T_{\text{trans}}$ and the glassy trapping temperature $T_g$. Combined as a whole, however, it is interesting to find out that the ratio of open—closed conformation and global folding $\Lambda_{\text{oc}}/\Lambda_{\text{global}}$ can be significantly correlated with the ratio of these characteristics (shown in Figure S7). As mentioned above, $\Lambda_{\text{oc}}$ and $\Lambda_{\text{global}}$ are correlated with $T_{\text{oc}}/T_{\text{fg}}$ and $T_{\text{global}}/T_{\text{fg}}$, respectively. Therefore, $\Lambda_{\text{oc}}/\Lambda_{\text{global}}$ is shown to have positive correlation with $T_{\text{oc}}/T_{\text{fg}}$ and $(T_{\text{oc}}/T_{\text{fg}})^{-1}$ (Figure S7A,B). Likewise, according to the equation $\Lambda = \delta E/\sqrt{2S}$, $\Lambda_{\text{oc}}/\Lambda_{\text{global}}$ is suggested to have a positive correlation with $\delta E_{\text{oc}}/\delta E_{\text{global}}$ and $((S_{\text{oc}})^{1/2}/2\delta E_{\text{oc}}/((S_{\text{global}})^{1/2}/2\delta E_{\text{global}}))^{-1}$ (Figure S7C,F). Though $\delta E_{\text{oc}}/\delta E_{\text{global}}$ and $((S_{\text{oc}})^{1/2}/2\delta E_{\text{oc}}/((S_{\text{global}})^{1/2}/2\delta E_{\text{global}}))$ have the opposite trends with $\Lambda_{\text{oc}}/\Lambda_{\text{global}}$, their product $((S_{\text{oc}})^{1/2}/2\delta E_{\text{oc}}/((S_{\text{global}})^{1/2}/2\delta E_{\text{global}}))$ has the reciprocal correlation with $\Lambda_{\text{oc}}/\Lambda_{\text{global}}$. The $\Lambda_{\text{oc}}/\Lambda_{\text{global}}$ shows the coupling effect between intrinsic open—closed and intrinsic folded—unfolded landscapes, but it has the same order with $\Lambda_{\text{oc}}$ (among the 5 proteins, $\Lambda_{\text{oc}}$ changes much greater than $\Lambda_{\text{global}}$). The results indicate that, of the 5 proteins in this study, ADK has relatively more of a funneled intrinsic open—closed conformation landscape and less of a funneled intrinsic folded—unfolded landscape (the difference between $T_{\text{oc}}/T_{\text{fg}}$ and $T_{\text{global}}/T_{\text{fg}}$ is the lowest), whereas DPO4 has relatively less of a funneled intrinsic open—closed conformation landscape and more of a funneled intrinsic folded—unfolded landscape (the difference between $T_{\text{oc}}/T_{\text{fg}}$ and $T_{\text{global}}/T_{\text{fg}}$ is the highest).
Average Local Frustration Reflects the Global Roughness. The localized frustration is connected with local conformational flexibility and large-scale conformational changes. The “frustratometer” introduced by Ferreiro et al. can provide the local frustration per residue. By using the online tool, the “frustratometer” (http://www.frustratometer.tk/), we analyzed the local conformational frustration of the open and closed forms of all the 5 proteins (see Figure S13). However, per residue frustration index cannot be used to compare with the characteristics of these proteins. We quantified the whole frustration by calculating the average highly local frustration (mean value of the entire protein) difference between open and closed forms $\Delta \text{frus} = \text{frus}^\text{oc} - \text{frus}^\text{ci}$. As shown in Figure S14, the results suggest that $\Delta \text{frus} = \text{frus}^\text{ci}$ correlates with the roughness $\Delta E^\text{oc}$ ($R^2 = 0.72$), $\frac{\Delta E^\text{oc}}{N}$ ($R^2 = 0.70$), and $\frac{(\Delta E^\text{oc})^2}{N}$ ($R^2 = 0.66$). This reveals that the average local frustration (quantified from frustratometer) can reflect the global roughness (roughness of the open–closed conformational landscape). In addition, the $\frac{(\Delta E^\text{oc})^2}{N}$ shows the “real” (intrinsic) roughness without the size effect. Here LIP1 is the system with the highest roughness and the lowest $\Delta \text{frus} = \text{frus}^\text{ci}$ of all; ADK is the one with the lowest roughness and the highest $\Delta \text{frus} = \text{frus}^\text{ci}$ of all. Consequently, the results indicate that the correlations between $\Delta \text{frus} = \text{frus}^\text{ci}$ and roughness of open–closed conformational landscape are independent with the protein size.

Intrinsic Conformation Energy Landscape Topography $\Lambda$ Determines the Kinetics of Protein Conformational Changes. In the previous studies, the intrinsic energy landscape topographic measure $\Lambda$ has been shown to be correlated to the kinetic folding/binding time $\tau_i$ during the processes of protein folding–binding–folding. Here in this study, we focus on the relationship between the topography measure of intrinsic local open–closed conformation energy landscape $\Lambda^\text{oc}$ and the rate of conformation switching from open to closed state. After 200 kinetic runs starting from the open configurations, the mean value of first passage time $\langle \tau_\varphi \rangle$ has been collected for each protein (see Table S1). Figure S12 illustrates the distribution of $\tau_\varphi$ for each system. Both $\langle \tau_\varphi \rangle$ and standard deviation of $\tau_\varphi$ of DPO4 are much higher than that of other proteins. As a result, DPO4 can be classified as GROUP I. The population of high $\tau_\varphi$ of DPO4 is extremely high. These can all be linked to the lowest $\Lambda^\text{oc}$ value with relatively low bias toward closed state and rougher conformation landscape. LIP1 and PhnD have similar medium $\langle \tau_\varphi \rangle$ and standard deviation of $\tau_\varphi$ values, which can be collected in GROUP II. In GROUP III, ADK and LAOBP have all relatively low $\langle \tau_\varphi \rangle$ and standard deviation of $\tau_\varphi$ values. In addition, the distributions of $\tau_\varphi$ in this group are toward the low $\tau_\varphi$ values. As shown in Figure S1, the protein with the lowest $\Lambda^\text{oc}$ (DPO4) has the largest open–closed conformation transition time. In $\langle \tau_\varphi \rangle$ can be fitted to $\Lambda^\text{oc}$ with the relationship of $y = a + be^{\Lambda^\text{oc}}$ (green dashed line) or $y = a + b/(x + c)$ (blue solid line) with similar fitting $R^2$. Except for the protein with extremely high $\Lambda^\text{oc}$ (ADK), other proteins with $\Lambda^\text{oc}$ lower than 2.0 seem to have a linear correlation between $\Lambda^\text{oc}$ and $\ln \langle \tau_\varphi \rangle$, indicating that low intrinsic landscape topography measure $\Lambda^\text{oc}$ may have a connection with low kinetic rate of switching from open and closed state, whereas high $\Lambda^\text{oc}$ does not correspond to significantly high kinetic rate. These results may have something to do with the protein structure as well as simulation model. In addition, logarithm of the standard deviation of $\tau_\varphi$ $\ln(\sqrt{\langle \tau_\varphi^2 \rangle - \langle \tau_\varphi \rangle^2})$ has the same behavior as $\ln \langle \tau_\varphi \rangle$. We also calculated the second-order moments of the open–closed transition time $\langle \tau_\varphi^2 \rangle / \langle \tau_\varphi \rangle^2$, which has a linear correlation with roughness $\Delta E^\text{oc}$ and configuration entropy $S$ of the intrinsic open–closed energy landscape. Above all, we demonstrate that a less biased, rougher, and greater sized conformation landscape will correspond to a lower conformational transition speed and lead to more significant fluctuations in conformation switching kinetics.

There are few experimental results of the open–closed transition kinetics ($k_\text{open}$ and $k_\text{closed}$) available. The open–closed kinetic rates of ADK have been reported by Hanson et al. via high-resolution single-molecule FRET ($k_\text{open} = 120 \pm 40$ s$^{-1}$, $k_\text{closed} = 220 \pm 70$ s$^{-1}$). However, these kinetic rates are obtained at experimental temperature, not the $\tau_\varphi$ in our kinetic simulations. Therefore, it is not appropriate to compare the kinetic rates directly. In DPO4, the open–closed conformational change is the rate-limited step of ligand binding. It has been reported that the mean binding time of DNA to DPO4 is about 1.3 s. Thus, the open–closed kinetic rate of DPO4 at experimental temperature is much lower than that of ADK. These results may have something to do with the topology of the open–closed conformational energy landscape.

Furthermore, we found that the $\ln \langle \tau_\varphi \rangle$ correlates highly with protein size $N$ (see the Supporting Information). This relationship may have something to do with the parameters of the model. Similar results have been reported in previous studies. This correlation may decrease or disappear if we continuously change the parameters of one protein system. However, the correlation between energy landscape topology and the crucial characteristics of proteins is independent with the parameters of the simulation model.

■ METHODS

The two-basin coarse-grained structure-based model (SBM) for each selected protein was constructed, on the basis of the open and closed reference structures of each protein in the Protein Data Bank (all the PDB IDs used in the simulations are listed in Table S1). The simulation details and analyses are introduced in the Supporting Information.

■ CONCLUSION

Intrinsic movements within a protein, such as large-scale domain–domain open–closed conformational transitions, are often essential for biomolecular functions. To uncover the underlying relationship between the intrinsic conformation energy landscapes and thermodynamics as well as kinetic rates of protein conformational changes, we quantify different kinds of intrinsic open–closed conformation energy landscapes and intrinsic global folding energy landscapes for 5 individual proteins with large-scale conformational changes. By applying the WHAM algorithm to transform the canonical ensemble to microcanonical ensemble, density of states, average energy, free energy, as well as other thermodynamic characteristics have been calculated for each protein. A dimensionless quantity, $\Lambda$, as a ratio of the energy gap or bias toward native state $\delta E$ versus energy landscape roughness $\delta E$ as well as configurational entropy $S(\Lambda = \delta E/(\Delta E\sqrt{2S}))$ is shown to quantify the topography of intrinsic conformation energy landscape.
For each protein, intrinsic conformation energy landscape was projected to a fraction of closed native contacts \((Q_{\text{closed}})\) and RMSD with respect to the native open structure \(RMSD_{\text{open}}\). The results suggest that there is only one basin on both intrinsic average energy and entropy (density of states) landscapes, pointing to the native closed state. In thermodynamics, the intrinsic conformation energy landscape topology measure \(\Lambda\) shows a strong linear correlation with the transition temperature against the glassy trapping temperature \((T_{\text{trans}}/T_g)\). In detail, intrinsic open–closed conformation topology measure \(\Lambda_{\text{oc}}\) correlates with conformation transition versus trapping temperature \(T_{\text{trans}}/T_g\), and intrinsic global folding landscape topology measure \(\Lambda_{\text{global}}\) correlates with folding transition versus trapping temperature \(T_{\text{global}}/T_g\). In addition, the ratio of open–closed conformation and global folding \(\Lambda (\Lambda_{\text{oc}}/\Lambda_{\text{global}})\) is found to be significantly correlated with the ratio of other thermodynamic characteristics, such as \(T_{\text{trans}}/T_{\text{global}}, \delta E_{\text{trans}}/\delta E_{\text{global}}, \Delta E_{\text{trans}}/\Delta E_{\text{global}}, (S_{\text{oc}}/S_{\text{global}})^{1/2}\), and \(((S_{\text{oc}})^{1/2}\Delta E_{\text{trans}})/(S_{\text{global}})^{1/2}\Delta E_{\text{global}})\). In addition, the average local frustration (quantified with \(\text{frus} - \text{frus}^+\)) is found to correlate with the global roughness (roughness of the open–closed conformational landscape), and this relationship is independent with the protein size. In kinetics, open–closed conformation transition time \(\tau_c\) can be linked with the topography measure \(\Lambda_{\text{oc}}\) as well as the structural model. Our investigations show that the intrinsic energy landscape topography can determine both the thermodynamics and the kinetics of conformation switching dynamics of proteins.

**REFERENCES**

(1) Anderson, C.; Zucker, F.; Steitz, T. Space-filling models of kinase clefts and conformation changes. Science 1979, 204, 375–380.
(2) Gerstein, M.; Lesk, A. M.; Chothia, C. Structural mechanisms for domain movements in proteins. Biochemistry 1994, 33, 6739–6749.
(3) Gerstein, M.; Krebs, W. A database of macromolecular motions. Nucleic Acids Res. 1998, 26, 4280–4290.
(4) Hayward, S.; Berendsen, H. J. Systematic analysis of domain motions in proteins from conformational change: New results on citrate synthase and T 4 lysozyme. Proteins: Struct., Funct., Genet. 1998, 30, 144–154.
(5) Wolf-Watz, M.; Thai, V.; Henzler-Wildman, K.; Hadjipavlou, G.; Eisenmesser, E. Z.; Kern, D. Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. Nat. Struct. Mol. Biol. 2004, 11, 945–949.
(6) Tama, F.; Sanejouand, Y.-H. Conformational change of proteins arising from normal mode calculations. Protein Eng., Des. Sel. 2001, 14, 1–6.
(7) McCammon, J. Protein dynamics. Rep. Prog. Phys. 1984, 47, 1–46.
(8) Henzler-Wildman, K. A.; Lei, M.; Thai, V.; Kerns, S. J.; Karplus, M.; Kern, D. A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. Nature 2007, 450, 913–916.
(9) Bryngelson, J. D.; Wolynes, P. G. Spin glasses and the statistical mechanics of protein folding. Proc. Natl. Acad. Sci. U. S. A. 1987, 84, 7524–7528.
(10) Brooks, C. L.; Gruebele, M.; Onuchic, J. N.; Wolynes, P. G. Chemical physics of protein folding. Proc. Natl. Acad. Sci. U. S. A. 1998, 95, 11037–11038.
(11) Shoemaker, B. A.; Wang, J.; Wolynes, P. G. Structural correlations in protein folding funnels. Proc. Natl. Acad. Sci. U. S. A. 1997, 94, 777–782.
(12) Clementi, C.; Nmeyer, H.; Onuchic, J. N. Topological and energetic factors: what determines the structural details of the transition state ensemble and “en-route” intermediates for protein folding? An investigation for small globular proteins. J. Mol. Biol. 2000, 298, 937–953.
(13) Turjanski, A. G.; Gutkind, J. S.; Best, R. B.; Hummer, G. Binding-induced folding of a natively unstructured transcription factor. PLoS Comput. Biol. 2008, 4, e1000600.
(14) Ganguly, D.; Chen, J. Topology-based modeling of intrinsically disordered proteins: Balancing intrinsic folding and intermolecular interactions. Proteins: Struct., Funct., Genet. 2011, 79, 1251–1266.
(15) Chu, J.-W.; Clarke, J.; Shammas, S. L.; Wang, J. Role of non-native electrostatic interactions in the coupled folding and binding of PUMA with McI-1. PLoS Comput. Biol. 2017, 13, e1005468.
(16) Lu, Q.; Lu, H. P.; Wang, J. Exploring the mechanism of flexible biomolecular recognition with single molecule dynamics. Phys. Rev. Lett. 2007, 98, 128105.
(17) Chu, J.-W.; Yoth, G. A. Coarse-grained free energy functions for studying protein conformational changes: a double-well network model. Biophys. J. 2007, 93, 3860–3871.
(18) Lu, Q.; Wang, J. Single molecule conformational dynamics of adenylate kinase: energy landscape, structural correlations, and transition state ensembles. J. Am. Chem. Soc. 2008, 130, 4772–4783.
(19) Wang, Y.; Gan, L.; Wang, E.; Wang, J. Exploring the dynamic functional landscape of adenylate kinase modulated by substrates. J. Chem. Theory Comput. 2013, 9, 84–95.
(20) Whittington, P. C.; Miyashita, O.; Levy, Y.; Onuchic, J. N. Conformational transitions of adenylate kinase: switching by cracking. J. Mol. Biol. 2007, 366, 1661–1671.
(21) Whittington, P. C.; Gosavi, S.; Onuchic, J. N. Conformational transitions in adenylate kinase allosteric communication reduces misfolding. J. Biol. Chem. 2008, 283, 2042–2048.
(22) Chu, X.; Liu, F.; Maxwell, B. A.; Wang, Y.; Suo, Z.; Wang, H.; Han, W.; Wang, J. Dynamic Conformational Change Regulates the Protein-DNA Recognition: An Investigation on Binding of a Y-Family Polymerase to Its Target DNA. PLoS Comput. Biol. 2014, 10, e1003804.
Proteins: Struct., Funct., Genet. 2008, 112, 15064−15073.

Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. J. Comput. Chem. 1992, 13, 1011−1021.

Ferrenberg, A. M.; Swendsen, R. H. Optimized monte carlo data analysis. Phys. Rev. Lett. 1989, 63, 1195−1198.

Bryngelson, J. D.; Wolynes, P. G. Intermediate barriers and random crossing in a random energy model (with applications to protein folding). J. Phys. Chem. 1989, 93, 6902−6915.

Wang, J. Diffusion and single molecule dynamics on biomolecular interface binding energy landscape. Chem. Phys. Lett. 2006, 418, 544−548.

Bryngelson, J. D.; Onuchic, J. N.; Socci, N. D.; Wolynes, P. G. Funnels pathways, and the energy landscape of protein folding: a synthesis. Proteins: Struct., Funct., Genet. 1995, 21, 167−195.

Onuchic, J. N.; Luther-Schulten, Z.; Wolynes, P. G. Theory of protein folding: the energy landscape perspective. Annu. Rev. Phys. Chem. 1997, 48, 545−600.

Plotkin, S. S.; Onuchic, J. N. Understanding protein folding with energy landscape theory part I: basic concepts. Q. Rev. Biophys. 2002, 35, 111−167.

Hanson, J. A.; Duderstadt, K.; Watkins, L. P.; Bhattacharyya, S.; Brokaw, J.; Chu, J.-W.; Yang, H. Illuminating the mechanistic roles of enzymes conformational dynamics. Proc. Natl. Acad. Sci. U. S. A. 2007, 104, 18055−18060.

Ferreiro, D. U.; Hegler, J. A.; Komives, E. A.; Wolynes, P. G. Localizing frustration in native proteins and protein assemblies. Proc. Natl. Acad. Sci. U. S. A. 2007, 104, 19819−19824.

Ferreiro, D. U.; Hegler, J. A.; Komives, E. A.; Wolynes, P. G. On the role of frustration in the energy landscapes of allosteric proteins. Proc. Natl. Acad. Sci. U. S. A. 2011, 108, 3499−3503.

Raper, A. T.; Gadkari, V. V.; Maxwell, B. A.; Suo, Z. Single-molecule investigation of response to oxidative DNA damage by a Y-family DNA polymerase. Biochemistry 2016, 55, 2187−2196.

Noel, J. K.; Whitford, P. C.; Sanbonmatsu, K. Y.; Onuchic, J. N. SMOG® cbp: simplified deployment of structure-based models in GROMACS. Nucleic Acids Res. 2010, 38, W657−W661.

Levy, Y.; Cho, S. S.; Onuchic, J. N.; Wolynes, P. G. A survey of flexible protein binding mechanisms and their transition states using native topology based energy landscapes. J. Mol. Biol. 2005, 346, 1121−1145.