Chemical Physics 493 (2017) 166-174

Contents lists available at ScienceDirect

Chemical Physics

journal homepage: www.elsevier.com/locate/chemphys

A coarse-grained method to predict the open-to-closed behavior of glutamine binding protein



College of Life Science and Bioengineering, Beijing University of Technology, Beijing 100124, China

ARTICLE INFO

Article history: Received 6 February 2017 In final form 24 May 2017 Available online 27 May 2017

Keywords: Glutamine-binding protein aANM Allosteric pathway Twisting motions Cooperative motion

ABSTRACT

The Glutamine-Binding Protein (GlnBP) of *Escherichia coli* is responsible for the first step in the active transport of glutamine across the cytoplasmic membrane. In present work, we explored the allosteric pathway of GlnBP from the open to closed states during the substrate binding process with the adaptive anisotropic network model (aANM). The results show that the allosteric transition proceeds in a coupled way and is more likely to be driven by the movement of hinge regions. The large-scale hinge-bending motion between the large and small domains occurs, accompanied by an interdomain twisting motion which proceeds mainly in the middle stage. The cooperative motion between the dominant hinge-bending motion and the twisting motion exerts a crucial role in the open-closed motion of GlnBP. These results are in close agreement with previous experimental and theoretical data, implying that the topology structure plays a crucial role in the allosteric transition process of GlnBP.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Complex macromolecular systems such as pumps, enzymes and ATP-Binding Cassette (ABC) transporters need to change their shapes and visit many conformational states in order to perform their functions. The Glutamine-Binding Protein (GlnBP) from *Escherichia coli*, a representative periplasmic binding protein in the ABC transporter superfamily, is in charge of the first step in the active transport of glutamine across the cytoplasmic membrane [1,2]. Many studies have proposed that the process of glutamine binding transport is accompanied by large-scale cooperative motions between the two domains of GlnBP [3,4]. Hence, the identification of its conformational transition pathway is important for our understanding of the physical mechanism for the functional motion of the GlnBP.

The crystal structures of GlnBP in two end-states, i.e., the ligand-bound closed and the ligand-free open forms, have been solved by X-ray crystallography [5,6]. GlnBP involves a single polypeptide chain of 226 residues that form a tertiary structure with two similar globular domains. The large domain contains both the C- and N-terminals of the protein, with two separate peptide segments, residues 1–84 and 186–226. This domain includes eight β -sheets and five α -helices. The small domain is comprised of residues 90–180, with three α -helices, four parallel and one anti-par-

* Corresponding author. *E-mail address:* chunhuali@bjut.edu.cn (C. Li).

http://dx.doi.org/10.1016/j.chemphys.2017.05.019 0301-0104/© 2017 Elsevier B.V. All rights reserved. allel β -sheets linked by a large loop (residues 96–109). These two domains are connected by two peptide hinges, residues 85-89 and 181–185 (Fig. 1, highlighted with black color). According to the previously proposed process [6-9], when the ligand glutamine enters the periplasmic region, some of them will bind to the ligand-free open GlnBP with a high affinity. The ligand binding induces the structural change of the hinge regions from a discontinuous β -sheet to a continuous one, which enhances softness of the hinges and modifies its direction to enable closing. This is the first step in the successful transport of glutamine across the cytoplasmic membrane. Although many studies have suggested that the interactions between hinge residues and the ligand may be the triggering mechanism that results in the closure of both domains, it is still not clear that how the conformational changes take place. The identification of the transition pathway can provide us an 'in-depth' understanding of the allosteric mechanism of the protein.

To understand allosteric transition mechanisms of macromolecules, one needs to go beyond the static information and determines how they change their conformations as a function of time. In practice, experimentally obtaining direct structural data about a transition pathway has been a challenge due to the transient nature of the high energy intermediates, and the multiplicity of pathways. However, computational methods can help generate physically plausible pathways for allosteric transitions which can then serve as "hypotheses" to be tested and refined experimentally [10–12]. Wriggers and Schulten [13] presented an algorithm based







Fig. 1. The structures of GlnBP at the ligand-free open state (A) and the ligandbound closed state (B). The hinge regions are highlighted with black color. The ligand glutamine is shown in the scaled ball and stick model.

on the crystallographic structures to identify and visualize the movements of rigid domains about common hinges in proteins. Given two conformations of a protein, Hayward and Lee [14] have developed a program to analyze the conformational change in terms of dynamic domains and hinge-bending regions. Additionally, some molecular dynamics simulation (MD) methods such as constrained MD [15], targeted and steered MD [16,17], as well as non-equilibrium MD [18] have been developed and applied to probe protein functional motions. However, the time-scale gap between the computational and experimental trajectories has made MD simulations too expensive and time-consuming for the study of the large-scale functional motions. Recently, the coarsegrained models have been proposed to solve this problem, among which the elastic network model (ENM) has been proved to be a very effective computational technique [19-22]. On the basis of the theories of ENM, Yang et al. [23] introduced a new approach, referred to adaptive anisotropic network model (aANM), to identify potential allosteric transition pathways between the known end conformations. This method utilizes the slow motion modes from the anisotropic network model (ANM) method to guide the conformational transitions along the directions intrinsically favored by the instantaneous inter-residue contact topology, and it has been successfully applied to the identification of the allosteric transition pathways of the maltose transporter and GroEL-GroES complex [23,24].

In the present work, we made use of the aANM method to explore the allosteric dynamics of GlnBP from the open to closed states, and the results present the entire allosteric process of the conformational transition. The movements between the large and small domains are studied based on the analyses of the changes in distances between some critical residues and cross-correlations among residues.

2. Materials and methods

2.1. Protein system

The crystal structures of the ligand-bound closed form and the ligand-free open form can be obtained from the Protein Data Bank (PDB codes: 1WDN (residues 4–226) and 1GGG (residues 5–224), respectively) [5,6]. We deleted one N- and two C-terminal residues from the structure 1WDN with the common parts (residues Leu5 to Glu224) of them remained, which has nearly no influence on the fluctuations of 1WDN. The adaptive anisotropic network model method was performed on the common parts of these two systems

to investigate the allosteric process. During aANM calculation, it should be noted that the closed state was built based on the structure with the ligand glutamine eliminated from the ligand-bound closed form.

2.2. Adaptive anisotropic network model

In aANM, a protein is considered as an elastic network, in which each residue is reduced to a node at its C_{α} atom, and the residue pairs within a cutoff distance are linked by elastic springs. The lowest frequency motion modes calculated from ANM are used to guide the conformation transition along the directions intrinsically favored by its instantaneous inter-residue contact topology, and the detailed information can be seen in the aANM paper by the Yang et al. [23].

During the simulation of the allosteric transition, the two end structures $R_0^{(0)}$ (open state) and $R_c^{(0)}$ (closed state) are used to generate the structures of the intermediate conformers, $R_0^{(k)}$ and $R_c^{(k)}$, where k is the kth iteration. After a total number of k_{tot} iterations, a series of conformations generated along the allosteric pathway can be represented as:

$$\{R_O^{(0)}, R_O^{(1)}, \cdots, R_O^{(k_{tot})}, R_C^{(k_{tot})}, \cdots, R_C^{(1)}, R_C^{(0)}\}$$
(1)

The conformations are represented as 3*N*-dimensional vectors, corresponding to the coordinates of the C_{α} atoms. The distance vector $d^{(k)}$ between the pair of conformations produced at the *k*th iteration is given by:

$$d^{(k)} = R_C^{(k)} - R_O^{(k)} \tag{2}$$

The original distance vector $d^{(0)}$ is computed after the optimal superimposition of the $R_{\alpha}^{(0)}$ and $R_{\alpha}^{(0)}$ structures.

$$d^{(0)} = R_C^{(0)} - R_O^{(0)}.$$
 (3)

The deformation vectors $v_o^{(k)}$ and $v_c^{(k)}$ used to produce the *k*th conformations are defined and calculated by the following formula as:

The root mean square deviation (RMSD) at *k*th iteration can be written as:

$$RMSD(R_{0}^{(k)}, R_{C}^{(k)}) = \frac{|d^{(k)}|}{\sqrt{N}}$$
(5)

The iteration procedure is described as follows:

i. Produce two sets of intermediate conformers starting from both ends. The following equation can be used to create the *k*th conformation $R_{\alpha}^{(k)}$:

$$\mathbf{R}_{O}^{(k)} = \mathbf{R}_{O}^{(k-1)} + \boldsymbol{\nu}_{O}^{(k)} = \mathbf{R}_{O}^{(k-1)} + \mathbf{s}_{O}^{(k)} \sum_{i=1}^{m_{O}^{(k)}} (\mathbf{d}^{(k-1)} \cdot \boldsymbol{u}_{iO}^{(k)}) \boldsymbol{u}_{iO}^{(k)}$$
(6)

A similar expression holds for $R_c^{(k)}$. In this equation, $u_{i0}^{(k)}$ $(1 \le i \le m_0^{(k)})$ are the dominant, i.e., low frequency eigenvectors obtained by the ANM for the conformation $R_0^{(k-1)}$. The step sizes $s_0^{(k)}$ and $s_c^{(k)}$ are simultaneously determined at iteration k, as a fraction f (scaling factor) of those, $s_{0m}^{(k)}$ and $s_{Cm}^{(k)}$, which minimize $d^{(k)}$. Usually, f = 0.2 (adopted here) can make an optimal balance between accuracy and efficiency [24,25].

ii. Determine the number of the selected modes $m_0^{(k)}$ used for producing the *k*th conformaton. First of all, the cumulative squared cosine is expressed as:

$$\left[C(m_{0}^{(k)})\right]^{2} = \sum_{i=1}^{m_{0}^{(k)}} \cos^{2}(d^{(k-1)}, u_{i0}^{(k)})$$
(7)

It has been confirmed that $C(m_0^{(k)})$ is identical to the correlation cosine between the instantaneous deformation vector $v_0^{(k)}$ and the distance vector targeted, i.e., $d^{(k-1)}$ [23,25].

$$C(m_0^{(k)}) = \cos(d^{(k-1)}, v_0^{(k)}).$$
(8)

Then, F_{\min} , as a threshold squared cosine, defines the maximal angular departure between $v_0^{(k)}$ and $d^{(k-1)}$. The minimal number of modes, starting from the low frequency end of the spectrum, that satisfy the inequality $[C(m_0^{(k)})]^2 \ge F_{\min}$ are selected to guide the structural transition. According to the previous researches, $F_{\min} = 0.5$ is adopted here [23].

iii. Determine the number of iteration steps. The process described above is repeated until the RMSD between the two intermediate conformers becomes sufficiently small. In this work, the iteration stops when the RMSD is less than 1.0 Å.

The cross-correlation between the displacements of the *i*th and *j*th residues at the *k*th iteration is given as:

$$C_{0,ij}^{(k)} = \cos(h_{0,i}^{(k)}, h_{0,j}^{(k)}), \tag{9}$$

where $h_{0,i}^{(k)}$ and $h_{0,j}^{(k)}$ are respectively the displacements of the *i*th and *j*th residues generated from the deformation vector $v_0^{(k)}$:

$$h_{0,i}^{(k)} = (\nu_{0,3i-2}^{(k)}, \nu_{0,3i-1}^{(k)}, \nu_{0,3i}^{(k)})$$
(10)

$$h_{0,j}^{(k)} = (v_{0,3j-2}^{(k)}, v_{0,3j-1}^{(k)}, v_{0,3j}^{(k)}),$$
(11)

where the deformation vector $v_0^{(k)}$ is from the Eq. (6). The cross-correlation value ranges from -1 to 1. The positive (negative) values represent that the residues move in the same (opposite) direction.

To calculate the energies of the conformations along the pathway, a double minima quadratic energy function is adopted in aANM [23]:

$$U = \frac{1}{2}(U_0 + U_C - \sqrt{(U_0 - U_C)^2 + 4\beta^2}), \qquad (12)$$

where the parameter β sets the height (and smoothness) of the energy barrier. The larger value of β makes the energy landscape smoother and the lower value makes the energy barrier higher. The potentials U_0 and U_c are defined as:

$$U_{0} = \frac{\gamma}{2} \sum_{r_{ij} < r_{c}} (|r_{ij}| - |r_{ij}^{0}|)^{2}$$

$$U_{C} = \frac{\gamma}{2} \sum_{r_{ij} < r_{c}} (|r_{ij}| - |r_{ij}^{C}|)^{2}$$
(13)

Here $|r_{ij}|$ is the instantaneous distance between residues *i* and *j* (based on C_{α} atoms), $|r_{ij}^{O}|$ and $|r_{ij}^{C}|$ are the equilibrium distances in the original conformations $R_{O}^{(0)}$ and $R_{C}^{(0)}$ respectively, and the summations are performed over all pairs of residues located within a cutoff distance r_{c} . The aANM trajectories are generated using Eq. (13) in each iteration by assuming that r_{ij}^{O} and r_{ij}^{C} are the final conformations of the previous iteration. In this work, we set the values $\beta = 10 \text{ kcal/mol}, \gamma = 0.7 \text{ kcal/(mol} \text{Å}^2)$ and $r_{c} = 7.3 \text{ Å}$ which are similar to the previous studies [3,23]. Due to the simplicity of the

energy function, the energies reported in the results provide information on the shape, rather than absolute values, of the energy profile.

3. Results and discussion

3.1. Operating modes

The lowest frequency ANM modes are used in the aANM to guide the structure transitions along the softest modes, or energetically favorable pathways, and thus selecting of lowest frequency modes is very crucial for aANM simulations [23,26]. Table 1 summarizes the $m_0^{(k)}$ and $m_c^{(k)}$ values for all steps along the forward and backward transitions. The original RMSD between the end points is 10.33 Å. After nineteen iterations, the RMSD between the instantaneous intermediate conformers reduces to lower than 1.0 Å. Notably, when the conformations shift away from the original states, the selected mode sets are gradually complemented by increasingly bigger subsets of higher frequency modes. As we know, the low frequency modes correspond to global or largeamplitude collective motions [27,28]. At the initial stages of the allosteric transition, there exist large and mainly global conformational differences between the two instantaneous endpoint conformations. Therefore, a small number of low frequency modes can meet the criterion $[C(m^{(k)})]^2 \ge F_{\min}$. And the structural changes driven by the slow modes produce substantial decrease in the RMSD between the instantaneous endpoints. From Table 1, during the first eight iterations, no more than 9 lowest frequency modes (1.38% of 654 nonzero eigenvectors) are selected for each iteration and there is a big decline in RMSD from 10.33 to 3.44 Å, making up 73.53% of the total drop. Then, the two instantaneous intermediate conformers get close to each other in conformational space as the iteration proceeds. At this moment, the structural changes involve some more local conformational adjustments which, we know, correspond to the higher frequency modes [27,28]. This is the reason why the broader numbers of higher frequency modes are recruited as the allosteric transition goes on. These imply that the slowest modes play a predominant role at the early steps of allosteric transition, and continue to exert a role throughout the entire trajectory, although they are gradually complemented by increasingly larger subsets of higher frequency modes. These results calculated by aANM agree well with previous studies about the allosteric process of GroEL and MsbA transporter [23,29].

Additionally, we want to know whether the modes recruited along the transition represent the most of the mean square fluctuations of the residues. In ANM, the contribution of the normal mode to the residue fluctuations scales with the reciprocal of its eigenvalue. Thus, we computed the percentage (called fluctuation percentage for simplicity) of the fluctuation contributed by the selected modes in residue fluctuations. It is identical to the ratio of $\sum_{i=7}^{m^{(k)}} 1/\lambda_i$ to $\sum_{i=7}^{3N} 1/\lambda_i$, where λ_i is the *i*th eigenvalue. The results are listed in Table 1 (see columns 3 and 5). It can be seen that at the first iteration, 3 (0.46% of 654 nonzero eigenvectors) and 5 (0.76%) modes contribute 58.18% and 41.65% to residue fluctuations, respectively. In the succeeding iterations, as the increasingly larger subsets of modes are adopted, the fluctuation percentage increases gradually. At the last iteration, 80 (12.23%) and 83 (12.69%) modes contribute 94.92% and 88.47% to residue fluctuations, respectively. These results indicate that the modes recruited along the transition are responsible for the most of the fluctuations in the system.

3.2. Allosteric sequence and energy profile during the transition

After nineteen iterations based on the aANM approach, thirtyeight intermediate conformers were generated, and six representative snapshots along the allosteric pathway are shown in Fig. 2. Before exploring the interdomain movement, we analyzed the interdomain movement for large and small domains. The average RMSDs of C_{α} atoms for the large (small) domain in the thirty-eight intermediate and final closed conformers from that in the initial open structure were calculated respectively after the former were superimposed on the latter using VMD [30]. The results show that the average RMSDs are 0.54 Å and 0.83 Å for large and small domains, implying that both domains move approximately as a rigid body respectively during the allosteric transition. To detect the interdomain movement, besides the six representative snapshots, the change of RMSDs of C_{α} atoms for the large (small) domain in the thirty-eight intermediate and final closed conformers relative to that in the initial open structure were also calculated respectively after their hinge regions superimposed along the

transition process (Fig. 3). From the upper snapshots in Fig. 2, it can be seen that both domains get close to each other around the hinge regions, indicating that allosteric transition is more likely to be driven by the movement of hinge regions. Many studies have reported that the allosteric transition of GlnBP is triggered by ligand binding which modifies the motion direction of the hinge region to enable closure correctly [9,31,32]. Additionally, from the change of the acute angle formed between the two black lines respectively connecting residue pairs Y86-K174 and Y86-W220 (Fig. 2), evidently for the upper snapshots the angle decreases gradually with the proceeding of the open-to-closed transition, while it seems to increase for the lower snapshots (obtained by rotating the corresponding upper ones 90° counterclockwise). This suggests that the open-to-closed transition of GlnBP is not a simple closing motion, and the movement in detail will be analyzed in the

Table 1

Numbers of ANM motion modes recruited at each iteration along the transition.

Iteration (k)	$m_0^{(k)}$	Fluctuation Percentage (%)	$m_{C}^{(k)}$	Fluctuation Percentage [*] (%)	RMSD (Å)
0	-	-	-	-	10.33
1	3	58.18	5	41.65	9.08
2	2	48.95	5	41.65	8.11
3	5	67.48	5	41.65	7.30
4	7	73.53	8	52.64	6.68
5	7	73.53	9	54.96	5.63
6	7	73.53	9	54.96	4.76
7	7	73.53	9	54.96	4.04
8	7	73.53	9	54.96	3.44
9	7	73.53	12	60.48	2.94
10	7	73.53	14	63.36	2.51
11	7	73.53	19	68.95	2.18
12	8	75.35	20	69.93	1.92
13	8	75.35	20	69.93	1.70
14	9	76.84	26	74.46	1.52
15	14	82.2	28	75.62	1.39
16	28	88.85	32	77.67	1.26
17	47	92.49	41	81.22	1.14
18	57	93.47	74	87.49	1.05
19	80	94.92	83	88.47	0.96

^{*} The percentages of the fluctuations contributed by the selected $m_{O}^{(k)}$ and $m_{C}^{(k)}$ modes in residue fluctuations, respectively.



Fig. 2. Representative conformations sampled along the allosteric pathway. The lower snapshots are obtained by rotating the corresponding upper ones 90° counterclockwise. The hinge region is highlighted. The acute angle between the two lines respectively connecting residue pairs Y86-K174 and Y86-W220 is used to illustrate that the open-to-closed transition of GlnBP is not a simple closing motion.



Fig. 3. Change of RMSD values of C_{α} atoms for the large (small) domain in the thirty-eight intermediate and final closed conformers relative to that in the initial open structure after their hinge regions superimposed along the allosteric transition of GlnBP.

following section. Furtherly from Fig. 3, comparing the RMSDs of the small and large domains, we note that the small domain has a much larger scale movement than the large domain during the transition, hinting that the movement of the small domain plays a dominant role for the allosteric transition process of GlnBP. Overall from Fig. 2, we can see the entire allosteric transition process of GlnBP from open to closed states, and the detailed transition information will be described below.

To evaluate if the resulting pathway obtained by aANM is energetically favorable, its energy profile calculated based on the double-well potential given by Eq. (12) is compared with that of the transition pathway obtained by the method of Cartesian interpolation (Fig. 4) [25]. The reaction coordinate is the projection of the cumulative displacement $v^{(n)} = R^{(n)} - R_0^{(0)}$ on the original distance vector d^{0} , that is, $x(n) = \frac{d^{(0)} \cdot p^{(n)}}{|d^{(0)}|^2}$, with the open and closed states representing the respective limits x(n) = 0 and 1. From Fig. 4, the energy peaks for the two methods are located in the later middle term of the allosteric transition (around x(n) = 0.60), and the energy barrier for aANM is much lower than that obtained by the interpolation method, which suggests that the aANM trajectory is



Fig. 4. Energy profiles of GlnBP along the reaction coordinate for the two transition pathways obtained by aANM and Cartesian interpolation methods. The reaction coordinate refers to the normalized projection of the cumulative displacement on the original distance vector.

energetically favorable. The linear interpolation corresponds to the rigid motion, and this kind of motion will undergo much higher barrier than that produced by aANM [25]. The results from previous studies on the intact Chaperon GroEL and fucose transporter are similar to ours [23,25].

3.3. Detailed conformational motions of GlnBP during the transition

During which stages of the allosteric transition do the main structural changes of different parts take place? How is the relationship among the orders of these structural changes? These are the important questions in allosteric transitions. To this end, the changes of distances between some critical residues as a function of the reaction coordinate are analyzed during the transition from the open to closed states (Fig. 5A) and these critical residues are labelled on the structure (Fig. 5B). Previous studies have shown that residues A12, V14, A67, T118 and D157 are important for ligand binding [33], and residues N160, I165, E181 and W220 are essential for ligand binding-induced conformational movements [9,34]. Thus in this part, we use the distances of residue pairs A12-T118 to reflect the movement of the doorkeeper region, the residue pairs V14-E181 (located in the binding pocket and hinge region respectively) to describe the hinge motion, the residue pairs A67-D157 (both situated in the binding pocket) to reflect the opento-closed transition of the binding pocket, and residue pairs V14-N160 (the first H-bond formed in the closing event) and I165-W220 (situated respectively at helix VI involving residues 158-167 and helix VIII composed of residues 212-220) to describe the closing motion between the small and large domains.

We know from the analyses on the allosteric sequence that the allosteric transition of GlnBP is not a simple open-close movement, but a complex one, in fact a hinge-bending movement accompanied with a twist one. Thus we also analyzed the twist motion between them. As both domains can be considered approximately as rigid bodies respectively, the change of the orientation deviation (orientation angle) of the line connecting a residue pair (here Lys125-Gly147 selected) in the small domain from that in the initial open structure during the transition process (the large domains superimposed together) can reflect the twist movement between the two domains (Figs. 5A, 6A). From Fig. 5A, this allosteric process can be divided into 3 stages according to the variation tendencies of the residue pair distances. For stage 1, we can see that the orien-



Fig. 5. Changes of the distances (A) of some critical residue pairs (B). (A) Also shows the change of the orientation deviation (orientation angle) of the line connecting the residue pair Lys125-Gly147 (Fig. 6A) in the small domain from that in the initial open structure during the transition process with the large domains superimposed together.

tation angle shows a little increase (1.3°) , and at the same time the four residue pair distances except for one pair I165-W220 have an evident descending trend, which indicates an obvious closing and a not big twist movements occur during this stage. For stage 2, the orientation angle displays a considerable increase (11.4°), which means that GlnBP is experiencing a large amplitude twist motion in this stage. And affected by the twist motion, the distance of residue pair V14-E181 has an increase to some extent, which cannot be caused by the closing movement. To depict the twist motion clearly, as a representative, the deformation vectors utilized to generate the twelfth iteration intermediate conformation from the eleventh are labelled on the eleventh iteration structure along the forward transition with black arrows (Fig. 6B, generated by VMD plugin Normal Mode Wizard [30]). From Fig. 6B, we can see an obvious twisting motion between two domains. Still for this stage, although the twist motion is obvious, its extent is weaker than that of the closing movement, which can be seen from the remarkable decline in the distances of residue pairs A12-T118 and A67-D157, indicating a large-scale closing motion for the doorkeeper region and binding pocket. For the stage 3, the distances of residue pairs and the orientation angle become gradually stable, which means the open-to-closed transition is almost over. During stage 3, it should be noted that there is an obvious increase in the distance between residues I165 of helix VI and W220 of helix VIII, which has been proved to be necessary to achieve closure between the two domains by MD simulations and principal component analysis (PCA) [9]. Overall, the allostatic transition of GlnBP



Fig. 6. Twist motion between the large and small domains. (A) The orientations of the line running through Lys125 and Gly147 in the small domain in the three intermediate conformations along the allosteric transition with the large domain superimposed together. (B) The deformation vectors utilized to generate the twelfth iteration intermediate conformation from the eleventh are labelled on the eleventh iteration structure along the forward transition with black arrows.

can be described as the dominant hinge-bending motion accompanied by a twist motion which occurs at the beginning and becomes more obvious in middle stage. Some studies have shown that along the open-to-closed conformational transition of GlnBP the dominant hinge-bending motion is accompanied by a twist motion [9,35–37].

3.4. Movement coupling between the small and large domains

To study the movement coupling between the two domains of GlnBP, the cross-correlations between residues were calculated based on formulas (9–11), respectively for the initial (Fig. 7A), final (Fig. 7C) and the 15th (from the closed state) (Fig. 7B) iteration steps along the allosteric transition. The recruited low frequency eigenvectors obtained by ANM for the three states are used to compute the cross-correlations. These eigenvectors reflect the inherent motion modes determined by the topological structure of GlnBP. Thus, the cross-correlation analyses below are closely related to the topological structures of GlnBP in the three states.

From the cross-correlation equation (9-11), its value ranges from -1 to 1; The positive values represent that the residues move in the same direction and the negative values represent that they move in the opposite direction; The higher the absolute value, the more correlatedly the two residues move. As shown in Fig. 7A, overall, the residues within the small (large) domain have positive correlations with each other, indicating the small (large) domain moves nearly as a whole. The region labelled with 'a' in small domain (residues 110-150) shows the highly consistent movements. Lai et al. have indicated that some residues in this region have a large fluctuation decrease during the open-to-closed transition [35]. In addition, the negative correlations exist between the two domains, indicating their opposite-direction movements and implying the open-to-closed allosteric transition. Many studies have confirmed that accompanying the conformational transition, the residues within large and small domains individually move

in a highly coupled way [3,38–40]. Compared with the cross-correlations in the initial iteration step, in the final iteration step (Fig. 7C), the correlations between some residues in large and small domains turn towards positive (rectangle regions c and f), especially for the residue pairs between Ala12, Asp49, Thr72, and Leu91, Gly117, Pro137, Thr167 (black circles), which is consistent with the observation by Lai et al. and D'Auria et al. that residue clusters Asp10-Arg75 in large domain and Lys115-Asn160 in small domain tend to form a contact and more compact rigid part with a lower extent of fluctuation during the open-to-closed transition [35,41]. On the other hand, the correlations between residues in the large domain become weak or change towards negative (regions b, d, e, g and h), which may be due to the fact that more residue contacts form between the two domains, especially between the residues in the jaws of the ligand binding pocket in the later stage of the allosteric transition [3,35], thus weakening the interactions between some residues in the large domain. Additionally, for the cross-correlation map from the 15th iteration in Fig. 7B, evidently, the residual cross-correlations are between those for the initial and final iterations. Overall, the cross-correlation analyses are in good agreement with the previous studies by ENM [3] and structure-based double-well model [35] methods.

It should be noted that the residues around Thr72 of the large domain have strong correlations with small domain residues, even in the initial state. From the structure, they are near the ligandbinding pocket and the back-door channel gate of GlnBP. The Xray crystallographic study has shown that residue Thr70 interacts with the glutamine directly, and is responsible for the tight binding of the ligand [6]. Using a perturbation method, we have shown that residues Arg75 and Gly119 (both in large domain) exert a key role in the regulation of the open-closed motion of the back-door



Fig. 7. Cross-correlations between residues respectively for the initial (A), 15th (from the closed state) (B) and final (C) iteration steps along the allosteric transition of GlnBP. Some correlation regions are highlighted by rectangles and circles.

channel gate [42]. Thus, we think that the strong correlations of the residues around Thr72 with the small domain residues are important for ligand binding and for the regulation of the open-closed motion of the back-door channel gate in the transition.

3.5. The effects of the parameters of aANM method on the results

The aANM calculations involve two parameters, F_{min} and f. The former controls the direction of motion, and the latter its size. Smaller F_{min} values permit us to proceed via lower energy ascent directions, at the cost of longer excursions. Smaller f implies smaller displacements at each iteration. Here we adopted their default values 0.5 and 0.2 respectively. In order to detect how the energy barrier depends on the step size of the iterations, the aANM calculations were repeated with f = 0.1, 0.2, 0.3, and the results show the values of the energy barrier are 281, 285 and 327 (arbitrary unit) respectively. Therefore, f = 0.2 is selected as a scaling factor that optimally balances between efficiency and accuracy. Additionally, we also explored how the pathways depend on the step size and the value of F_{min} through monitoring the changes of the C_{α} - C_{α} distances of some critical residue pairs along the transition pathways simulated with different values of F_{min} (=0.4, 0.5, 0.6) and f (=0.1, 0.2, 0.3) (see Fig. S1). From Fig. S1, the relative relationships among the distance changes of these residue pairs have almost no much change for different F_{min} and f value sets. The results indicate that the transition of GlnBP from open to closed states is of robustness to some extent over a relatively broad range of aANM parameters $(0.4 \le F_{min} \le 0.6 \text{ and } 0.1 \le f \le 0.3)$, which is consistent with the results obtained by aANM for bacterial chaperonin GroEL [23].

3.6. Discussion of the simulation method

In fact, for allosteric proteins, besides protein architecture, sequence specificities also play crucial roles in encoding motions [43,44]. Additionally, clearly nonlinearity should be of central importance when proteins pass through their transition states for rearrangement. In order to consider these issues, Li et al. developed a multi-basin model which combines an atomic interaction-based coarse-grained (AICG) model with sequence-specific and nonlinear interactions taken into account with the perfect funnel model to study protein allosteric transitions [45,46]. Thus, the multi-basin model developed by Li et al. should be more reasonable compared with aANM method [45,46]. However, the latter has also its advantages. Specifically, ANM model is of simplicity, robustness, low computational cost and relatively high accuracy [47] because its harmonic potential is very simple with only two parameters (distance cutoff and spring constant), ANM modes can be easily obtained analytically and more importantly quasi-harmonic fluctuations are encoded largely in the native three-dimensional architecture [45,46]; And besides in aANM method, the allosteric process is simulated iteratively with no need for molecular dynamics simulations. Theoretically, the low frequency modes correspond to large-amplitude collective or functional motions [48], and therefore the allosteric pathway obtained by aANM through guiding the structure transitions along the lowest, or softest modes, or energetically favorable pathways is of reasonability to some extent. aANM method may serve as a first approximation for exploring the transitions between not-too-distant pairs of protein functional states [23].

It should be noted that the elucidation of the allosteric transition mechanisms of macromolecules has been a challenge, both experimentally and computationally, because of the transition nature of the high energy conformers and the multiplicity of pathways [24,29]. The aANM method by definition aims at sampling those probable, structurally favored paths [23]. Actually, some of our calculation results about the allosteric pathway of the GlnBP are in good accordance with the experimental and theoretical data from other methods. However, the allosteric transition pathway predicted by aANM method may only approximate the real transition path. One reason is the modes calculated by the ANM are those exclusively based on inter-residue contact topology and no other specific interactions are taken into consideration during the iteration process [23]. Thus, these obtained results by aANM method should be further verified by more experimental and theoretical observations in the future.

4. Conclusions

In our present work, we have used the adaptive anisotropic network model to investigate the allosteric transition and dynamics of GlnBP. We have shown that aANM is a valuable tool in exploring functional global motions and allosteric dynamics of proteins. The calculations of cross-correlations between residues confirm the allosteric transition of GlnBP proceeds in a cooperative way. The changes of the orientation angle and the distances between some critical residues reveal that the allosteric transition can be described as the large-scale dominant hinge-bending motion accompanied by a twist motion which occurs at the beginning and becomes more obvious in the middle stage. Additionally, our results suggest that the topological structure plays an important role in the allosteric transition of GlnBP. Finally, although limitations exist, we envision that aANM represents a potentially useful tool for understanding of the allosteric mechanism of GlnBP.

Acknowledgments

This work was supported by grants from the Chinese Natural Science Foundation (11474013, 31171267) and the Beijing Natural Science Foundation (4152011).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.chemphys.2017. 05.019.

References

- [1] S. D'Auria, M. Staiano, A. Varriale, M. Gonnelli, A. Marabotti, M. Rossi, G.B. Strambini, The differences in the microenvironment of the two tryptophan residues of the glutamine-binding protein from *Escherichia coli* shed light on the binding properties and the structural dynamics of the protein, Proteins 71 (2008) 743–750.
- [2] S. Pistolesi, N. Tjandra, Temperature dependence of molecular interactions involved in defining stability of glutamine binding protein and its complex with L-glutamine, Biochemistry-US 51 (2012) 643–652.
- [3] J.G. Su, X. Jiao, T.G. Sun, C.H. Li, W.Z. Chen, C.X. Wang, Analysis of domain movements in glutamine-binding protein with simple models, Biophys. J. 92 (2006) 1326–1335.
- [4] G.A. Bermejo, M.P. Strub, C. Ho, N. Tjandra, Ligand-free open-closed transitions of periplasmic binding proteins: the case of glutamine-binding protein, Biochemistry-US 49 (2010) 1893–1902.
- [5] C.D. Hsiao, Y.J. Sun, J. Rose, B.C. Wang, The crystal structure of glutaminebinding protein from *Escherichia coli*, J. Mol. Biol. 262 (1996) 225–242.
- [6] Y.J. Sun, J. Rose, B.C. Wang, C.D. Hsiao, The structure of glutamine-binding protein complexed with glutamine at 1.94 Å resolution: comparisons with other amino acid binding proteins, J. Mol. Biol. 278 (1998) 219–229.
- [7] G.F. Ames, Bacterial periplasmic transport systems: structure, mechanism, and evolution, Annu. Rev. Biochem. 55 (1986) 397–425.
- [8] C. Chennubhotla, Z. Yang, I. Bahar, Coupling between global dynamics and signal transduction pathways: a mechanism of allostery for chaperonin GroEL, Mol. Biosyst 4 (2008) 287–292.
- [9] H.H. Loeffler, A. Kitao, Collective dynamics of periplasmic glutamine binding protein upon domain closure, Biophys. J. 97 (2009) 2541–2549.
- [10] A. Das, M. Gur, M.H. Cheng, S. Jo, I. Bahar, B. Roux, Exploring the conformational transitions of biomolecular systems using a simple two-state anisotropic network model, PLoS Comput. Biol. 10 (2014) e1003521.

- [11] J.J. Lacroix, S.A. Pless, L. Maragliano, F.V. Campos, J.D. Galpin, C.A. Ahern, B. Roux, F. Bezanilla, Intermediate state trapping of a voltage sensor, J. Gen. Physiol. 140 (2012) 635–652.
- [12] K.A. Henzler-Wildman, V. Thai, M. Lei, M. Ott, M. Wolf-Watz, T. Fenn, E. Pozharski, M.A. Wilson, G.A. Petsko, M. Karplus, C.G. Hubner, D. Kern, Intrinsic motions along an enzymatic reaction trajectory, Nature 450 (2007) 838–844.
- [13] W. Wriggers, K. Schulten, Protein domain movements: detection of rigid domains and visualization of hinges in comparisons of atomic coordinates, Proteins 29 (1997) 1–14.
- [14] S. Hayward, R.A. Lee, Improvements in the analysis of domain motions in proteins from conformational change: DynDom version 1.50, J. Mol. Graph. Model. 21 (2002) 181–183.
- [15] J. Ma, M. Karplus, Molecular switch in signal transduction: reaction paths of the conformational changes in ras p21, Proc. Natl. Acad. Sci. U.S.A. 94 (1997) 11905–11910.
- [16] A. Fornili, B. Giabbai, G. Garau, M. Degano, Energy landscapes associated with macromolecular conformational changes from endpoint structures, J. Am. Chem. Soc. 132 (2010) 17570–17577.
- [17] B. Isralewitz, J. Baudry, J. Gullingsrud, D. Kosztin, K. Schulten, Steered molecular dynamics investigations of protein function, J. Mol. Graph. Model. 19 (2001) 13–25.
- [18] K. Sharp, J.J. Skinner, Pump-probe molecular dynamics as a tool for studying protein motion and long range coupling, Proteins 65 (2006) 347–361.
- [19] O. Kurkcuoglu, R.L. Jernigan, P. Doruker, Collective dynamics of large proteins from mixed coarse-grained elastic network model, QSAR Combin. Sci. 24 (2005) 443–448.
- [20] C. Tuzmen, B. Erman, Identification of ligand binding sites of proteins using the gaussian network model, PLoS One 6 (2011) e16474.
- [21] H.Y. Li, Z.X. Cao, L.L. Zhao, J.H. Wang, Analysis of conformational motions and residue fluctuations for *Escherichia coli* ribose-binding protein revealed with elastic network models, Int. J. Mol. Sci. 14 (2013) 10552–10569.
- [22] J.G. Su, X.M. Han, X. Zhang, Y.X. Hou, J.Z. Zhu, Y.D. Wu, Analysis of conformational motions and related key residue interactions responsible for a specific function of proteins with elastic network model, J. Biomol. Struct. Dyn. 34 (2016) 560–571.
- [23] Z. Yang, P. Májek, I. Bahar, Allosteric transitions of supramolecular systems explored by network models: application to chaperonin GroEL, PLoS Comput. Biol. 5 (2009) e1000360.
- [24] C.H. Li, Y.X. Yang, J.G. Su, B. Liu, J.J. Tan, X.Y. Zhang, C.X. Wang, Allosteric transitions of the maltose transporter studied by an elastic network model, Biopolymers 101 (2014) 758–768.
- [25] S. Chang, K.S. Li, J.P. Hu, X. Jiao, X.H. Tian, Allosteric and transport behavior analyses of a fucose transporter with network models, Soft Mater. 7 (2011) 4661–4671.
- [26] A.R. Atilgan, M.C. Demirel, B. Erman, I. Bahar, Vibrational dynamics of folded proteins: significance of slow and fast motions in relation to function and stability, Phys. Rev. Lett. 80 (1998) 2733–2736.
- [27] N. Kurt, W.R. Scott, C.A. Schiffer, T. Haliloglu, Cooperative fluctuations of unliganded and substrate-bound HIV-1 protease: a structure-based analysis on a variety of conformations from crystallography and molecular dynamics simulations, Proteins 51 (2003) 409–422.
- [28] M.C. Demirel, A.R. Atilgan, R.L. Jernigan, B. Erman, I. Bahar, Identification of kinetically hot residues in proteins, Protein Sci. 7 (1998) 2522–2532.
- [29] X.L. Xie, C.H. Li, Y.X. Yang, L. Jin, J.J. Tan, X.Y. Zhang, J.G. Su, C.X. Wang, Allosteric transitions of ATP-binding cassette transporter MsbA studied by the adaptive anisotropic network model, Proteins 83 (2015) 1643–1653.

- [30] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (33–38) (1996) 27–28.
- [31] A. Kitao, Transform and relax sampling for highly anisotropic systems: application to protein domain motion and folding, J. Chem. Phys. 135 (2011) 119903.
- [32] C.A. Klug, K. Tasaki, N. Tjandra, C. Ho, J. Schaefer, Closed form of liganded glutamine-binding protein by rotational-echo double-resonance NMR, Biochemistry-US 36 (1997) 9405–9408.
- [33] J.P. Hu, T.G. Sun, W.Z. Chen, C.X. Wang, Molecular dynamics simulations and free energy calculation on glutamine-binding protein, Acta Chim. Sinica 64 (2006) 2079–2085.
- [34] S. Hayward, A. Kitao, Monte carlo sampling with linear inverse kinematics for simulation of protein flexible regions, J. Chem. Theory Comput. 11 (2015) 3895–3905.
- [35] Z.Z. Lai, Q. Lu, J. Wang, Exploring the thermodynamic landscape, kinetics, and structural evolution of a protein conformational transition with a microscopic double-well model, J. Phys. Chem. B 115 (2011) 4147–4159.
- [36] A. Pang, Y. Arinaminpathy, M.S. Sansom, P.C. Biggin, Comparative molecular dynamics-similar folds and similar motions?, Proteins 61 (2005) 809-822
- [37] A. Pang, Y. Arinaminpathy, M.S.P. Sansom, P.C. Biggin, Interdomain dynamics and ligand binding: molecular dynamics simulations of glutamine binding protein, FEBS Lett. 550 (2003) 168–174.
- [38] U. Emekli, D. Schneidman-Duhovny, H.J. Wolfson, R. Nussinov, T. Haliloglu, HingeProt: automated prediction of hinges in protein structures, Proteins 70 (2008) 1219–1227.
- [39] D.R. Livesay, D.H. Huynh, S. Dallakyan, D.J. Jacobs, Hydrogen bond networks determine emergent mechanical and thermodynamic properties across a protein family, Chem. Cent. J. 2 (2008) 17.
- [40] S.C. Flores, K.S. Keating, J. Painter, F. Morcos, K. Nguyen, E.A. Merritt, L.A. Kuhn, M.B. Gerstein, HingeMaster: normal mode hinge prediction approach and integration of complementary predictors, Proteins 73 (2008) 299–319.
- [41] S. D'Auria, A. Scirè, A. Varriale, V. Scognamiglio, M. Staiano, A. Ausili, A. Marabotti, M. Rossi, F. Tanfani, Binding of glutamine to glutamine-binding protein from *Escherichia coli* induces changes in protein structure and increases protein stability, Proteins 58 (2005) 80–87.
- [42] D.S. Lv, C.X. Wang, C.H. Li, J.J. Tan, X.Y. Zhang, An efficient perturbation method to predict the functionally key sites of glutamine binding protein, Comput. Biol. Chem. 67 (2017) 62–68.
- [43] S.R. Tzeng, C.G. Kalodimos, Dynamic activation of an allosteric regulatory protein, Nature 462 (2009) 368–372.
- [44] K.A. Henzler-Wildman, M. Lei, V. Thai, S.J. Kerns, M. Karplus, D. Kern, A hierarchy of timescales in protein dynamics is linked to enzyme catalysis, Nature 450 (2007) 913–916.
- [45] W. Li, P.G. Wolynes, S. Takada, Frustration, specific sequence dependence, and nonlinearity in large-amplitude fluctuations of allosteric proteins, Proc. Natl. Acad. Sci. U.S.A. 108 (2011) 3504–3509.
- [46] W. Li, W. Wang, S. Takada, Energy landscape views for interplays among folding, binding, and allostery of calmodulin domains, Proc. Natl. Acad. Sci. U.S. A. 111 (2014) 10550.
- [47] J.R. López-Blanco, P. Chacón, New generation of elastic network models, Curr. Opin. Struc. Biol. 37 (2016) 46–53.
- [48] I. Bahar, T.R. Lezon, A. Bakan, I.H. Shrivastava, Normal mode analysis of biomolecular structures: functional mechanisms of membrane proteins, Chem. Rev. 110 (2010) 1463–1497.