Analysis of Domain Movements in Glutamine-Binding Protein with Simple Models

Ji Guo Su,*† Xiong Jiao,* Ting Guang Sun,* Chun Hua Li,* Wei Zu Chen,* and Cun Xin Wang*
*College of Life Science and Bioengineering, Beijing University of Technology, Beijing, China; and †College of Science, Yanshan University, Qinhuangdao, China

ABSTRACT In this work, the mechanism of domain movements of glutamine-binding protein (GlnBP), especially the influence of the ligand on GlnBP dynamic behavior is investigated with the aid of a Gaussian network model (GNM) and an anisotropy elastic network model. The results show that the "open-closed" transition mainly appears as the large movement of the small domain, especially the top region including two α-helices and two β-strands. The slowest mode of each three forms of GlnBP—ligand-free open, ligand-bound closed, and ligand-free closed GlnBP—shows that the open-closed motion of the two domains has a common hinge axis centered on Lys-87 and Gln-183. Accompanying the conformational transition, the residues within both large and small domains move in a highly coupled way. The peaks of the fast modes correspond to residues that were thought, in the GNM, to be important for the stability of the protein, and these residues may be involved in the interactions with the membrane-bound components. With the contacts between the large domain and the small domain increasing, the ability of the "open-closed" motion is decreased. All the results agree well with those of molecular dynamics simulations, and it is thought that the open-closed conformation transition is the nature of the topology structure of GlnBP. Also, the influence of the ligand on GlnBP is studied with a modified GNM method. The results obtained show that the ligand does not influence the closed-to-open transition tendency.

INTRODUCTION

In Gram-negative bacteria, the cell envelope consists of an outer membrane and an inner membrane, and the periplasmic space between them contains a number of complex transport systems (1,2). The outer membrane is semipermeable and acts as a coarse molecular sieve. The inner membrane contains numerous active transporters for some molecular compounds. Periplasmic binding proteins can carry small ligands from the periplasmic space into the cytoplasmic space via the inner membrane. When a ligand enters the periplasmic region, it will bind to the corresponding periplasmic binding protein with a high affinity. Upon ligand binding, the periplasmic binding protein adopts conformational changes so that the complex formed is recognized by the transporter, which is another component of the system embedded in the inner membrane. Then, the ligand is released from the binding protein and translocated across the cytoplasmic membrane by the transporter (3). The x-ray structures of more than 14 different periplasmic binding proteins share a similar fold topology containing two distinct domains linked by two or three short hinges (4).

GlnBP from *Escherichia coli* is one of the representative structures of the periplasmic binding proteins. Binding of glutamine at the cleft between two domains causes a conformational change corresponding to a closure of two domains around the ligand. Besides the ligand-free open (open-apo) and ligand-bound closed (closed-ligand) forms, there may also be a closed form of GlnBP in the absence of ligand (closed-apo) (5–7). It is suggested that both forms (open-apo and closed-apo) are present in equilibrium with a low activation energy barrier between them (6) and the equilibrium will be shifted toward the closed form upon ligand binding (8).

It seems to be clear that the open-closed transition of GlnBP is of utmost functional significance for ligand binding. Also, the significance of slow and fast modes of other proteins has been addressed in some recent work (9,10). But several questions still need to be answered. Is the transition encoded in the structure itself? Is it an intrinsic property of the system and can it be predicted from the structure alone? There is a tendency for a conformational change from the closed-*apo* form to the open-*apo* form (6,8). Then, for the case with the ligand bound to GlnBP, can the ''closed-to-open'' transition also occur? If so, how does the ligand affect the tendency of the closed-ligand structure to open?

To answer the above questions, we have used molecular dynamics (MD) simulations to study the sequence of events involved in the large-scale conformational change for the open-apo and closed-ligand forms (11). Pang et al. have explored the ''open-closed'' domain motion using MD simulations and essential dynamics analysis (12). Although the presence of the expected domain motion was indicated by extensive mobility in their direction, the short-timescale simulation could not detect the complete conformation transition, and the substantial energy barriers could impede simulations of the complete transition pathway.

In this work, these issues are investigated from a novel point of view through two simple coarse-grained schemes—a Gaussian network model (GNM) and an anisotropy

Submitted April 3, 2006, and accepted for publication October 16, 2006.

J. G. Su and X. Jiao contributed equally to this work.

Address reprint requests to Prof. Cun Xin Wang, College of Life Science and Bioengineering, Beijing University of Technology, Beijing 100022, China. Tel.: 86-10-67392724; E-mail: cxwang@bjut.edu.cn.

0006-3495/07/02/1326/10 $2.00

doi: 10.1529/biophysj.106.086512

elastic network model (ANM) (13–15). The GNM method allows direct estimation of the conformation transition from the crystal structures without the high computational cost of MD. Information about the directions of this transition can be obtained from analysis with the ANM. The mechanism of the open-closed transition and the influence of the ligand on the transition will be explored by analyzing the large-scale domain motions around a hinge with the GNM. The GNM method has been proved in numerous application studies to be a simple yet useful tool for investigating large-scale conformational motions, domain motions, and collective dynamics of the biomolecular systems (16–20). The GNM also has been used to find kinetically hot residues and folding cores of proteins (21,22). Several previous studies have proved that the results of the GNM are in agreement with those of MD simulation (23,24).

Because structure information is available both for the open and the closed forms, it is possible to compare the results obtained from the GNM and ANM with those from the crystal structures for the closed and open forms. In this work, the information about crystal structure will be used to compare and verify the conclusion reached using the GNM and ANM methods.

MATERIALS AND METHODS

Crystal structure of the protein

Crystal structures have been solved in the open-apo form (Protein Data Bank (PDB) code 1GGG) (25) and the closed-ligand form (PDB code 1WDN) (8). The crystal structure of the open-apo form includes two monomers that have the same conformations. We chose only one monomer for our calculation in this work. The closed-apo structure was obtained by removing the bound glutamine from the closed-ligand structure. GlnBP contains a single polypeptide chain of 226 residues with a secondary structure of ~35% α-helices and 37% β-strands. The large domain, which contains both the N-terminal and C-terminal of the protein, is built from two separate peptide segments, residues 1–84 and 186–226. This domain includes five α-helices and eight β-strands. The small domain consists of residues 90–180, which form three α-helices and four parallel and one antiparallel β-strands connected by a large loop (residues 96–109). The two domains are connected by two β-strands (residues 85–89 and 181–185). (see Fig. 1). The first four residues and the first three residues are missing in the crystal structures for the open-apo and closed-ligand structures, respectively. For our crystal structure comparison, all three coordinate sets were truncated to the shortest protein (1GGG) with residues from Leu^5 to Glu^{224}.

The Gaussian network model and the anisotropy elastic network model

The Gaussian network model describes a three-dimensional structure of protein as an elastic network of C_α atoms connected by harmonic springs within a certain cutoff distance (7.3 Å is adopted in this work). The force constant is identical for all springs. Considering all contacting residues, the internal Hamiltonian of the system can be written as (20)

$$V = \frac{1}{2}\gamma[\Delta R^{\mathrm{T}}(\Gamma \otimes E)\Delta R], \tag{1}$$

where γ is the harmonic force constant; ΔR represents the $3N$-dimensional column vector of the X, Y, and Z components of the fluctuation vectors

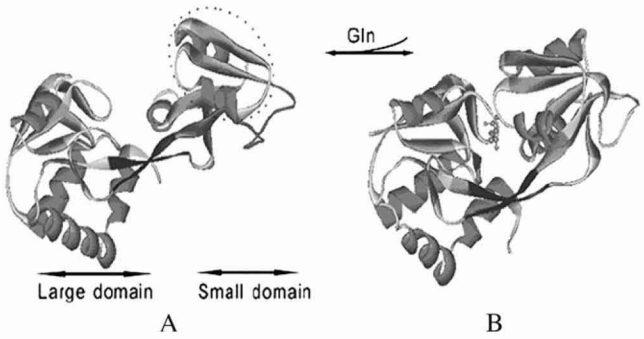

FIGURE 1 Open-apo (1GGG) and closed-ligand (1WDN) crystal conformations. (*A*) The open-apo conformation (1GGG). The location with large movement (residues 110–150, see text) in 1GGG compared to 1WDN is indicated by a dotted circle. (*B*) The closed-ligand conformation (1WDN). The bound glutamine is shown in the scaled ball and stick model.

$\Delta R_1, \Delta R_2, \ldots \Delta R_N$ of the C_α atoms, where N is the number of residues; the superscript T denotes the transposition; E is the third-order identity matrix; $\otimes$ is the direct product; and Γ is the $N \times N$ symmetric Kirchhoff matrix in which the elements are written as (13,20)

$$\Gamma = \begin{cases} -1 & \text{if } i \neq j \text{ and } R_{ij} \leq r_c \\ 0 & \text{if } i \neq j \text{ and } R_{ij} > r_c \\ -\sum_{i,i\neq j} \Gamma_{ij} & \text{if } i = j \end{cases}, \tag{2}$$

where R_{ij} is the distance between the ith and jth C_α atoms and r_c is the cutoff distance.

The mean-square fluctuations of each atom and the cross-correlation fluctuations between different atoms are in proportion to the diagonal and off-diagonal elements of the pseudoinverse of the Kirchhoff matrix. The inverse of the Kirchhoff matrix can be decomposed as

$$\Gamma^{-1} = U\Lambda^{-1}U^{\mathrm{T}}, \tag{3}$$

where U is an orthogonal matrix whose columns $u_i (1 \leq i \leq N)$ are the eigenvectors of Γ, and Λ is a diagonal matrix of eigenvalues λ_i of Γ. The cross-correlation fluctuations between the ith and jth residues are given by

$$\langle \Delta R_i \cdot \Delta R_j \rangle = \frac{3k_B T}{\gamma}[\Gamma^{-1}]_{ij}, \tag{4}$$

where k_B is the Boltzmann constant, T is the absolute temperature, and the meanings of γ and Γ are the same as in Eq. 1. When $i = j$, the mean-square fluctuations of the ith residue can be obtained. The Debye-Waller or B-factor, which is correlated to the mean-square fluctuation, can be calculated with the expression

$$B_i = 8\pi^2 \langle \Delta R_i \cdot \Delta R_i \rangle / 3 \tag{5}$$

The mean-square fluctuation of the ith residue associating with the kth mode is given by

$$\langle \Delta R_i \cdot \Delta R_i \rangle_k = \frac{3k_B T}{\gamma}\lambda_k^{-1}[u_k]_i[u_k]_i. \tag{6}$$

In the GNM, the cross correlation is normalized as

$$C_{ij} = \frac{\langle \Delta R_i \cdot \Delta R_j \rangle}{[\langle \Delta R_i^2 \rangle \times \langle \Delta R_j^2 \rangle]^{1/2}}. \tag{7}$$

The GNM model can provide the amplitudes of residue fluctuations but no information about the directions of the fluctuations. Then an ANM model

is introduced, by which information about the orientation of fluctuations is elicited. In ANM, the motion mode of a protein is determined by a Hessian matrix H.

$$H = \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1N} \\ h_{21} & h_{22} & \cdots & h_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ h_{N1} & h_{N2} & \cdots & h_{NN} \end{pmatrix} \quad (8)$$

The elements of H are submatrix with size 3×3. The ijth submatrix h_{ij} is:

$$h_{ij} = \begin{pmatrix} \frac{\partial^2 V}{\partial x_i \partial x_j} & \frac{\partial^2 V}{\partial x_i \partial y_j} & \frac{\partial^2 V}{\partial x_i \partial z_j} \\ \frac{\partial^2 V}{\partial y_i \partial x_j} & \frac{\partial^2 V}{\partial y_i \partial y_j} & \frac{\partial^2 V}{\partial y_i \partial z_j} \\ \frac{\partial^2 V}{\partial z_i \partial x_j} & \frac{\partial^2 V}{\partial z_i \partial y_j} & \frac{\partial^2 V}{\partial z_i \partial z_j} \end{pmatrix}. \quad (9)$$

when $i \neq j$, the analytic expression for the elements of h_{ij} is

$$\frac{\partial^2 V}{\partial x_i \partial y_j} = \left. \frac{-\gamma (x_j - x_i)(y_j - y_i)}{R_{ij}^2} \right|_{R_{ij}=R_{ij}^0}, \quad (10)$$

when $i = j$, the analytic expression for the elements of h_{ij} is

$$\frac{\partial^2 V}{\partial x_i \partial y_i} = \gamma \sum_{j \neq i} \left. \frac{(x_j - x_i)(y_j - y_i)}{R_{ij}^2} \right|_{R_{ij}=R_{ij}^0}. \quad (11)$$

The meanings of γ and R are the same as in Eq. 1. x, y, and z represent the coordinates of atoms.

CORRELATION COEFFICIENT AND OVERLAP

The linear correlation coefficient between the calculated and experimental B-factors (obtained from the PDB) is given by

$$\rho = \frac{\sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})}{\left[\sum_{i=1}^{N} (x_i - \bar{x})^2 \sum_{i=1}^{N} (y_i - \bar{y})^2\right]^{1/2}}, \quad (12)$$

where x_i and y_i are the calculated and experimental B-factor values of the ith C_α atom, $\bar{x}$ and $\bar{y}$ are the mean values of x_i and y_i, and N is the total number of C_α atoms of the protein.

The overlap is used to measure the similarity between the direction given by motion mode j and the direction of the conformational change (26).

$$I_j = \frac{\left| \sum_{i=1}^{3N} a_{ij} \Delta r_i \right|}{\left[\sum_{i=1}^{3N} a_{ij}^2 \sum_{i=1}^{3N} \Delta r_i^2\right]^{1/2}}. \quad (13)$$

In this expression, Δr_i is the amount of the ith atomic coordinate's change between the open and closed structures. A higher value of the overlap means that the direction gotten from the motion mode is more similar to the one gotten from Δr_i.

The correlation coefficient c_j is used to measure the similarity of the patterns of atomic displacement in the conformational change and those in the jth motion mode:

$$c_j = \frac{1}{N} \frac{\sum^{N} (A_{ij} - \overline{A_j})(\Delta R_i - \overline{\Delta R})}{\sigma(A_j)\sigma(\Delta R)}, \quad (14)$$

where A_{ij} is the magnitude of the displacement of atom i involved in the mode j and ΔR_i is the amplitude of the displacement in the conformation change. The corresponding average displacements are marked as, respectively, A_j and $\overline{\Delta R}$, and the corresponding root mean-square are $\sigma(A_j)$ and $\sigma(\Delta R)$.

RESULTS AND DISCUSSION

Comparison of predicted mean-square fluctuations with experimental B-factors

To evaluate the feasibility of applying the GNM method to study GlnBP, the B-factors are calculated with this method and then compared with the data from x-ray crystallography. According to Eq. 5, the B-factors of the open-apo, closed-apo, and closed-ligand forms are calculated. Fig. 2 shows the comparison between the calculated B-factor of C_α atoms (*dotted line*) and the experimental data from x-ray crystallography (*solid line*). It can be seen, from the theory described above, that the only adjustable parameter in this work is γ, which is determined by normalizing the theoretical distribution of the B-factors based on the experimental one. The resulting $k_B T/\gamma$ value used for the open-apo structure is 0.97 Å^2, that for the closed-apo structure is 1.22 Å^2, and that for the closed-ligand structure is 1.25 Å^2.

Fig. 2 shows the correlation between the experimental B-factor and the theoretical one, in which the correlation coefficients are 0.488, 0.568, and 0.581 for the open-apo, the closed-apo, and the closed-ligand structures, respectively. The results are similar to those of recent studies for other proteins (14,27). Because the closed-apo structure is generated by removing the bound glutamine from the ligand-bound structure, the theoretical B-factor of the former is compared with the experimental B-factor of the latter. It is found that the closed-ligand structure gives a higher correlation coefficient than the closed-apo structure, which reflects the effect of the ligand glutamine on the residue fluctuations of the protein. In this regard, the difference between the B-factor values of the two structures was calculated to figure out which residues were affected by the ligand (see Fig. 3 *A*). The difference is obtained by subtracting the B-factor value of the closed-ligand structure from that of the closed-apo structure, and the result shows that the ligand reduces the fluctuations of the receptor's residues in contact with it, such as Gly^{68}, Ile^{69}, Thr^{70}, Thr^{118}, Gly^{119}, Ser^{120}, and Asp^{157}, as well as the residues around them.

Surprisingly, the correlation coefficient of the open-apo structure is only 0.488, which is much lower than that of the

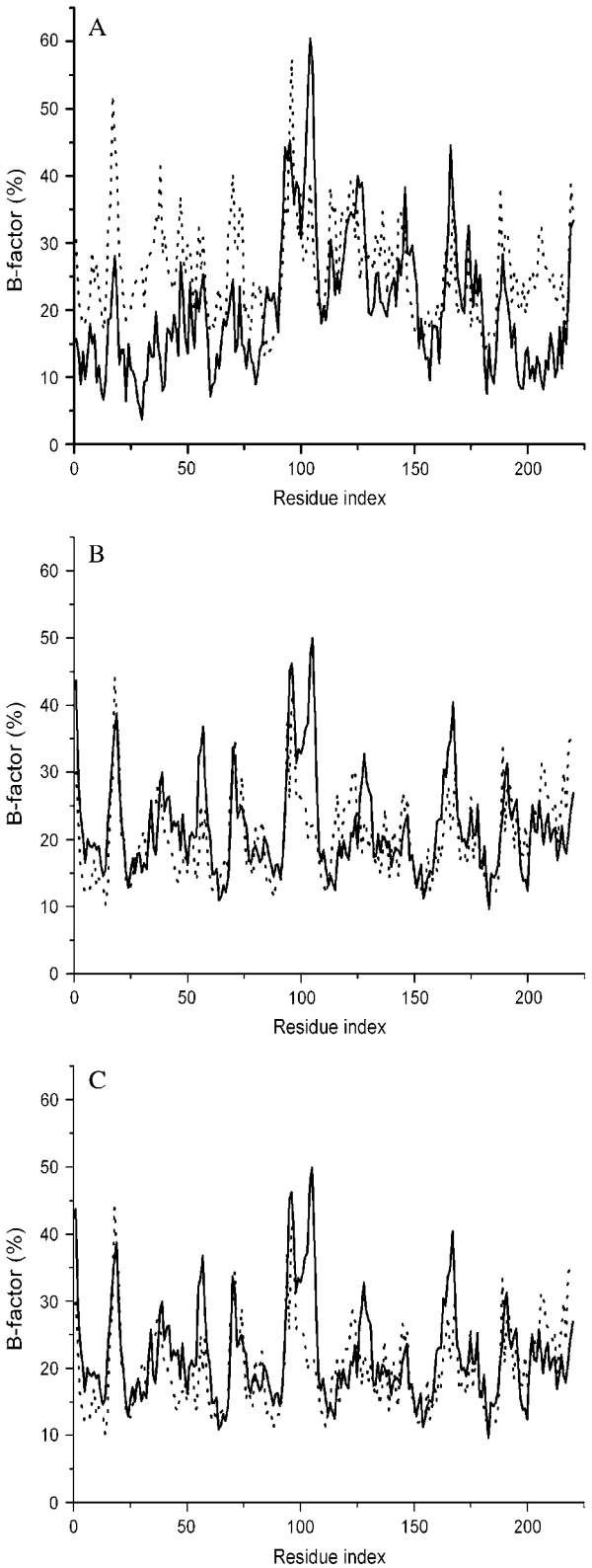

FIGURE 2 (*A*) Experimental (*solid line*) and calculated (*dotted line*) B-factor of the open-apo form of GlnBP (PDB code 1GGG). (*B*) Experimental (*solid line*) and calculated (*dotted line*) B-factor of the closed-apo structure of GlnBP. The experimental data are from the closed-ligand structure. (*C*) Experimental (*solid line*) and calculated (*dotted line*) B-factor of the closed-ligand structure.

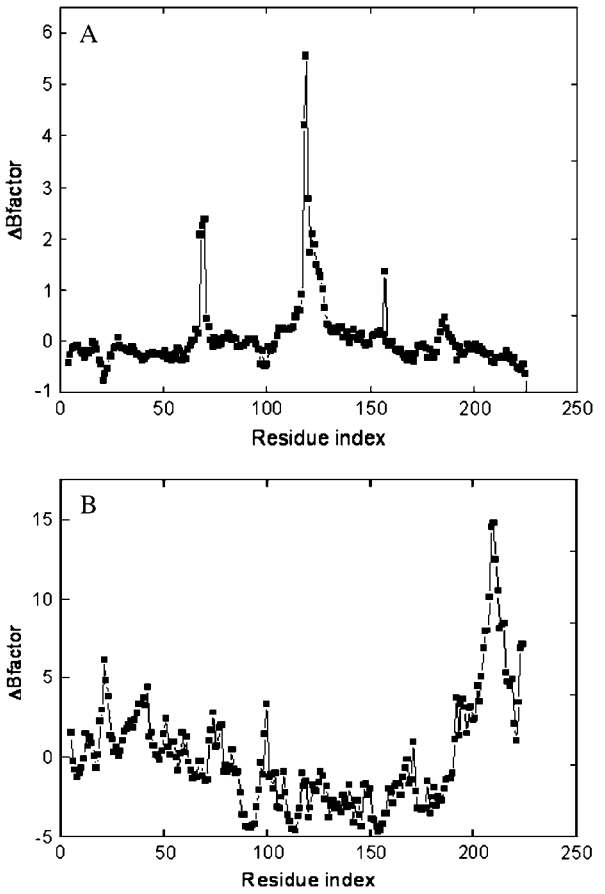

FIGURE 3 (*A*) Calculated B-factor difference between the closed-apo and closed-ligand structures. (*B*) The difference of B-factor of the open-apo structure calculated by the following two methods: using only one monomer to calculate the B-factor or using two monomers in the crystal structure together to calculate the B-factor, which takes into consideration the interaction between the two monomers.

other two structures. This may result from omitting the effect of the other protein nearby, since the open-apo structure for our theoretical B-factor calculation is selected from one of two identical monomers in contact with each other in the crystal structure. To confirm this, the B-factor of the crystal structure consisting of the two monomers is also calculated and then the correlation coefficient between the theoretical and experimental values is obtained, which really shows a remarkable improvement from 0.488 to 0.696 (27). As shown in Fig. 3 *B*, the obvious difference between B-factor values is mainly located at the region of the contact interface between the two monomers, which suggests that the fluctuations of the residues in this region greatly decrease.

The slow modes of the motions

The slow and long-wavelength collective modes represent functionally relevant motions of protein (28). Fig. 4 displays the first mode of each three structures calculated by the GNM. The ordinates in Fig. 4 show the normalized distribution of

squared fluctuations driven by the first slowest modes. From the figure, the two domains of the protein and the hinges between them are highly distinguished. All three structures have the common hinge axes located around Lys^{87} and Gln^{183}, with the fluctuation values approximate to zero. Around the hinge axis, the major domain movements occur. It can also be seen that the α-helix (residues Thr^{158}–Thr^{167}) near the hinges, which belongs to the small domain, has small fluctuation during the domain movements.

It is also seen in Fig. 4 that the fluctuation of the small domain is higher than that of the large domain, especially the block at the top position in the small domain, including two α-helices and two β-strands (residues 110–150). This indicates that the open-closed conformational transition mainly exhibits as the large movement of the small domain, particularly part of the top region of it. This is consistent with the results of MD simulation, in which the corresponding residues have a large root mean-square fluctuation (RMSF) (13,14). As shown in Fig. 4, residues 212–225 in both ligand-free and ligand-bound closed structures also exhibit large fluctuation values corresponding to the flexibility of the C-terminal of the protein other than in the open-closed transition. Regarding the fluctuation of the closed-ligand structure, the value of ligand glutamine is very low, so there is a sharp decrease at the end of the curve.

From Fig. 4, comparing the slowest mode of the closed (closed-apo and closed-ligand) structures with that of the open-apo structure, it can be seen that the fluctuations of some residues in the closed structures are reduced remarkably. Those are the residues around Ala^{12}, Asp^{49}, Gly^{117}, and Pro^{137}. The inset in Fig. 4 shows the ribbon diagram of the closed-ligand structure, with those residues whose fluctuations decreased remarkably marked in black. Evidently, these residues are located in the jaws of the ligand docking pocket, among which Asp^{10} and Lys^{115} act as a doorkeeper that locks the glutamine ligand tightly inside the binding pocket (8). Thus, these residues play an important role in the open-closed transition of the pocket. The fluctuation decrease of these residues implies that the binding pocket became more stable in the closed structure than in the open-apo structure. It became difficult to open the binding pocket of the closed structure due to the tight binding of the ligand. Obviously, it should be noted that the first mode of the closed-apo in Fig. 4 is extremely similar to that of the closed-ligand structures. It suggests that the ligand of glutamine has little effect on the domain motion.

GNM can only provide the magnitude of displacement of atoms from their equilibrium positions for large-scale motions. To ascertain the direction of motion, ANM is applied to the three structures. For the open-apo structure, the ANM calculation shows a hinge-bending motion and a twisting motion, as seen in Fig. 5. The first slowest mode corresponds to the hinge-binding motion. This mode is explored in its two opposite directions, resulting in two structures. The magnitude of the amplification is adjusted so that the two structures have an RMSD of ~2 Å, which makes the difference large enough and easily inspected visually. The two structures resulting from the first slowest mode are shown in Fig. 5 *A*. The hinge-bending motion can easily be identified and this motion results in the open-closed transition of the binding pocket. The second slowest mode can be described as a twisting motion involving the two lobes of the GlnBP, as

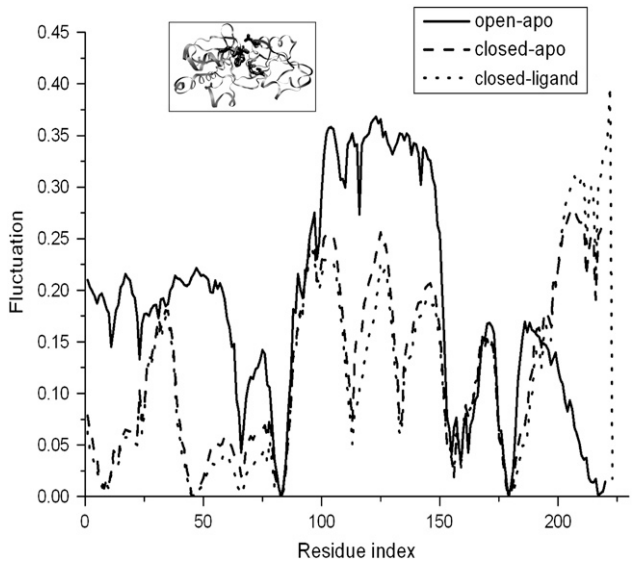

FIGURE 4 Comparison between the slowest mode shapes of three structures. (The unit of the mean-square fluctuation is $Å^2$). The inset is a ribbon diagrams of the closed-ligand structure, with those residues for which fluctuations decreased remarkably marked in black.

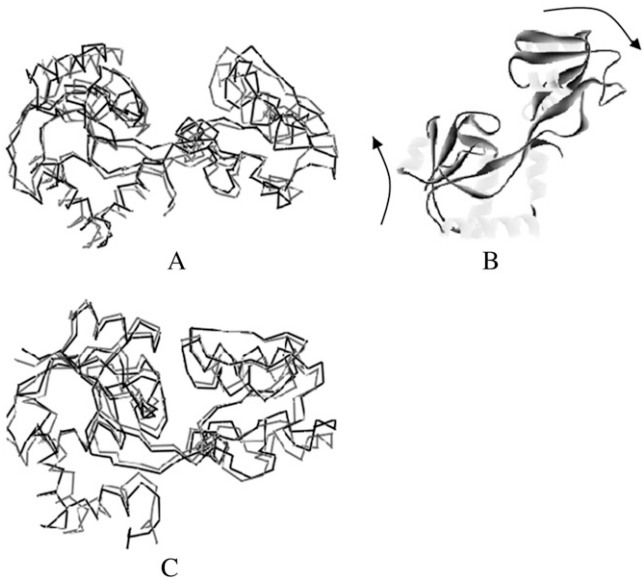

FIGURE 5 (*A*) Hinge-bending motion of the open-apo structure. (*B*) The twisting motion of the open-apo structure. (*C*) The hinge-bending motion of the closed-ligand structure.

shown in Fig. 5 *B*. Such a twisting motion of the periplasmic binding protein has been suggested previously (5).

For the closed-ligand structure, the ANM calculation also shows a twisting motion and a hinge-binding motion. The first slowest motion is the twisting motion and the second slowest motion is the hinge-bending motion. The two structures resulting from the hinge-bending motion, which have an RMSD of ~2Å, are shown in Fig. 5 *C*. From the figure, it is demonstrated that the binding pocket is stable in the closed-ligand structure. The hinge-bending motion can hardly result in the open-closed transition of the pocket. The most flexible regions for the hinge-bending motion are residues 18–26 of a β-loop in the large domain and residues 96–109 of the large loop in the small domain. It was proposed that the β-loop may be involved in interactions with the membrane-bound components of the glutamine transport system (8). Thus, the large fluctuation of this region may correlate to the binding of the closed-ligand complex with the membrane-bound components. The results of the closed-apo structure (data not shown) are the same as that of the closed-ligand structure.

Fig. 6 shows the reciprocal of eigenvalues for all the modes, which represent the contribution of individual modes to the observed dynamics. To show the differences between the slow motion modes clearly, a local amplificatory graph for the first 16 slow motion modes is inserted in Fig. 6. The curve in Fig. 6 indicates that the first two slowest modes make a significant contribution to the motions of the open-apo structure. These modes make a fractional contribution of 0.2 to the observed dynamics. For the closed-ligand structure, this value is 0.06, which means that the twisting motion and the hinge-bending motion make a smaller contribution to the dynamics than the open-apo structure. The overlap and correlation coefficients of each mode for the two structures are calculated using Eqs. 13 and 14, respectively. The expression in Eq.13 describes the similarity between the mode calculated by ANM and the conformational change vector, and the expression in Eq. 14 describes the magnitude of the correlation between the two vectors. For the open-apo structure, the overlap and correlation values are 0.472 and 0.563, respectively, with respect to the first slowest mode, and 0.173 and 0.089 with respect to the second slowest mode. For the closed-ligand structure, these values are 0.194 and 0.00161 with respect to the first slowest mode, and 0.261 and 0.312 with respect to the second slowest mode, which indicates a remarkable decrease compared to the corresponding values in the open-apo structure.

The fast modes of the motions

The fast modes correspond to geometric irregularity in the local structure and the fluctuations associated with fast modes are accompanied by a decrease in entropy larger than that for slow modes (28). Therefore, residues acting in the fast modes are thought to be kinetically hot residues and they are critically important for the stability of the tertiary fold (21,22,28). Fig. 7 shows the fastest six modes of the three structures. Surprisingly, the mode shapes of the three structures are found to be extremely similar, which implies that their local structures are substantially similar and the domains will keep their rigid structures during domain motions. That is to say, the residues within the two domains move in a highly coupled way, which will also be found from the cross-correlation maps discussed in the following section.

As shown in Fig. 7, there are several peaks representing the kinetically hot residues in the curves, which may play a key role in the stability of the protein. It was also found that these corresponding hot residues are tightly packed, and a large part of them are located at an exposed region, including

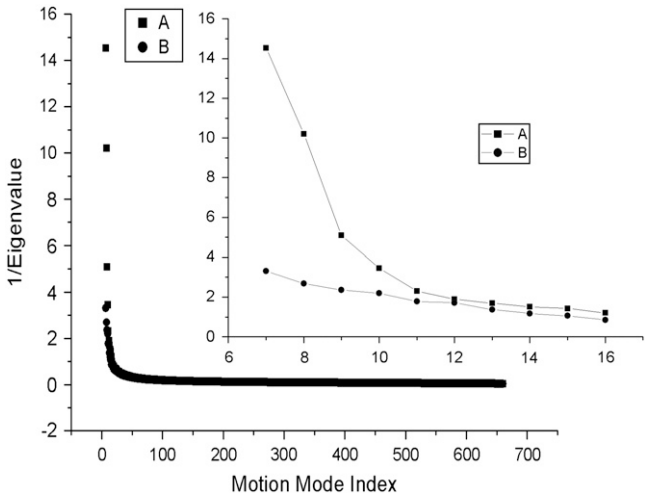

FIGURE 6 Reciprocal of eigenvalues for all the modes of the open-apo structure (serial *A*) and that of the closed-ligand structure (serial *B*). The inset is a local amplificatory graph for the slow motion mode.

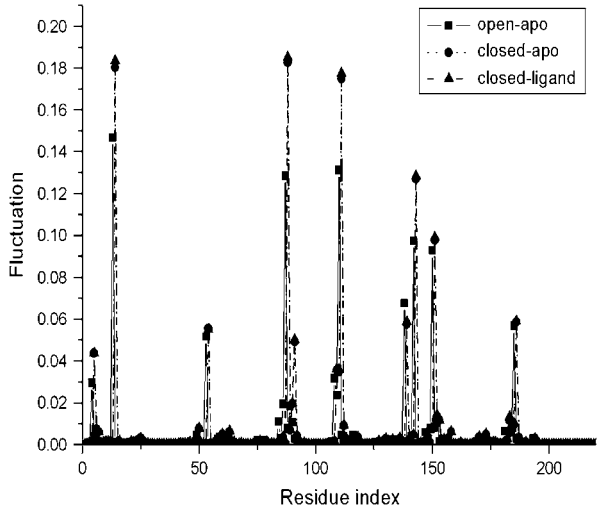

FIGURE 7 Fastest six mode shapes of the open-apo structure, the closed-apo structure, and the closed-ligand structure. There are several peaks marked in the curve that correspond to the kinetically hot residues.

some residues of an α-helix in the small domain and residues of a β-strand in the large domain. It has been proposed that this ''exposed region'' may be involved in the interactions with the membrane-bound components (8). In Fig. 7, this exposed region corresponds to the region around Glu^{17} and the region around Leu^{146} and Val^{154}.

Cross correlations between atomic fluctuations

The cross correlations between the fluctuations of C_α atoms are calculated using Eq. 7. Since the modes with low frequency correspond to functional motions and those with high frequency correspond to localized motions, here only low-frequency modes are used to improve the signal/noise ratio. We used the first 40 modes in our calculation. The results are shown in Fig. 8. The cross-correlation value ranges from -1 to 1, in which range the values are positive when the residues move in the same direction and negative when they move in the opposite direction. The higher the absolute cross-correlation value, the better the two residues are correlated (or anticorrelated). Also, uncorrelated fluctuations yield $C_{ij} = 0$. As shown in Fig. 8, there is negative correlation in the blue regions and positive correlation in the orange-red regions.

It is obvious that all the cross-correlation maps of these structures are divided into five orange-red and four blue portions. The four blue regions represent the negative correlation between the two domains, which move in the opposite direction. This corresponds to the open-closed conformation transition. The center region of the five orange-red regions represents the residues of the small domain moving as a whole, whereas the other four indicate the residues of the large domain moving in a coupled way. This demonstrates that the large and small domains will keep their structures stable during the open-closed transition. In the cross-correlation map, there is a region with high correlation values, marked out by a rectangle in Fig. 8. This region corresponds to the highly coupled movement of the top part of the small domain (residues 110–150). As mentioned above, some residues in this part have a large fluctuation decrease when the structure is changed from the open to the closed form. With this conformation transition occurring, more contacts will form between the residues in the large domain and those in the small domain, especially the residues in the jaws of the ligand docking pocket. Thus, the correlation between atoms in the small domain will be weakened, and this phenomenon will appear in the large domain at the same time. On the other hand, the correlation between the atoms in the large domain and those in the small domain is changed toward positive, especially for the atom pairs between which contact appears, such as Ala^{12}, Asp^{49}, Gly^{117}, and Pro^{137}. From the comparison between Fig. 8, *A* and *B*, this change of cross correlation is present clearly as some bright dot in the blue region and the color of the point around the diagonal is darkened.

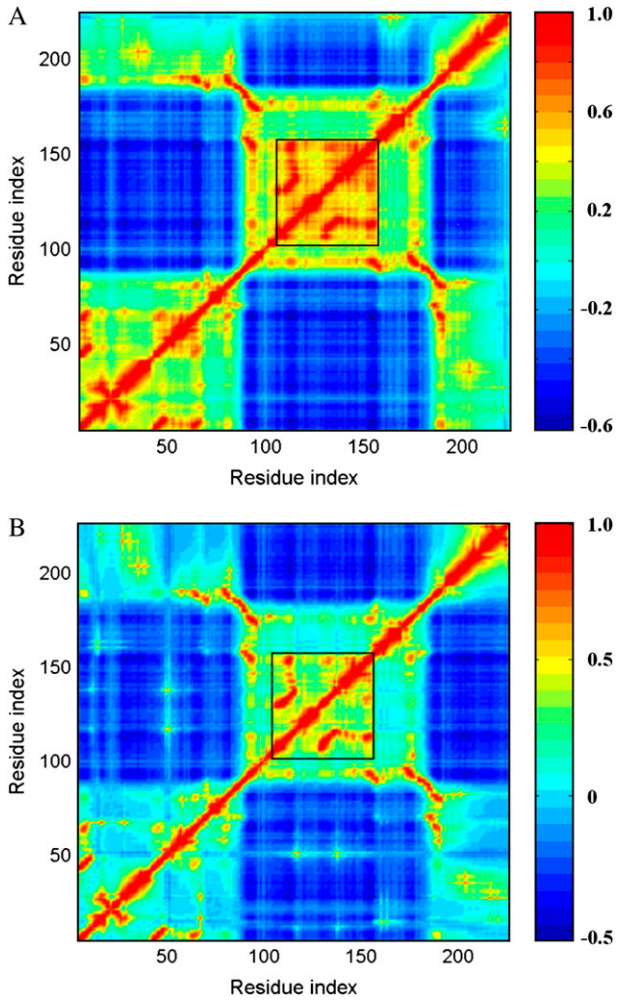

FIGURE 8 Cross-correlation maps calculated using the first 40 modes for two systems: the open-apo structure (*A*) and the closed-apo structure (*B*). Blue regions in the figure indicate negative correlation and orange-red regions indicate positive correlation, as shown in the color bar on the right. Both the *x* axis and *y* axis of the map are residue indices. The region in the rectangle (residues 110–150) indicates the highly coupled region.

Topology and interaction between domains determine domain motions

The directions of the displacements can be obtained from the ANM analysis. According to this information, three reconstructed models of the GlnBP were generated. The backbone of the protein is modified first, based on the information of direction at various coefficients, to make sure the RMSD between these three conformations and the open-apo structure is distinct. According to the open degree of the structure, the RMSD values are 2.9, 3.9, and 5.5 Å, respectively. The side chain is added. Then these resultant conformations are optimized by energy minimization followed by a 500-ps equilibrium dynamics with the GROMACS program. For each trajectory, an equilibrated conformation is picked out randomly to do the GNM analysis. These three structures are shown in Fig. 9 *A*. The slowest modes of motion of these

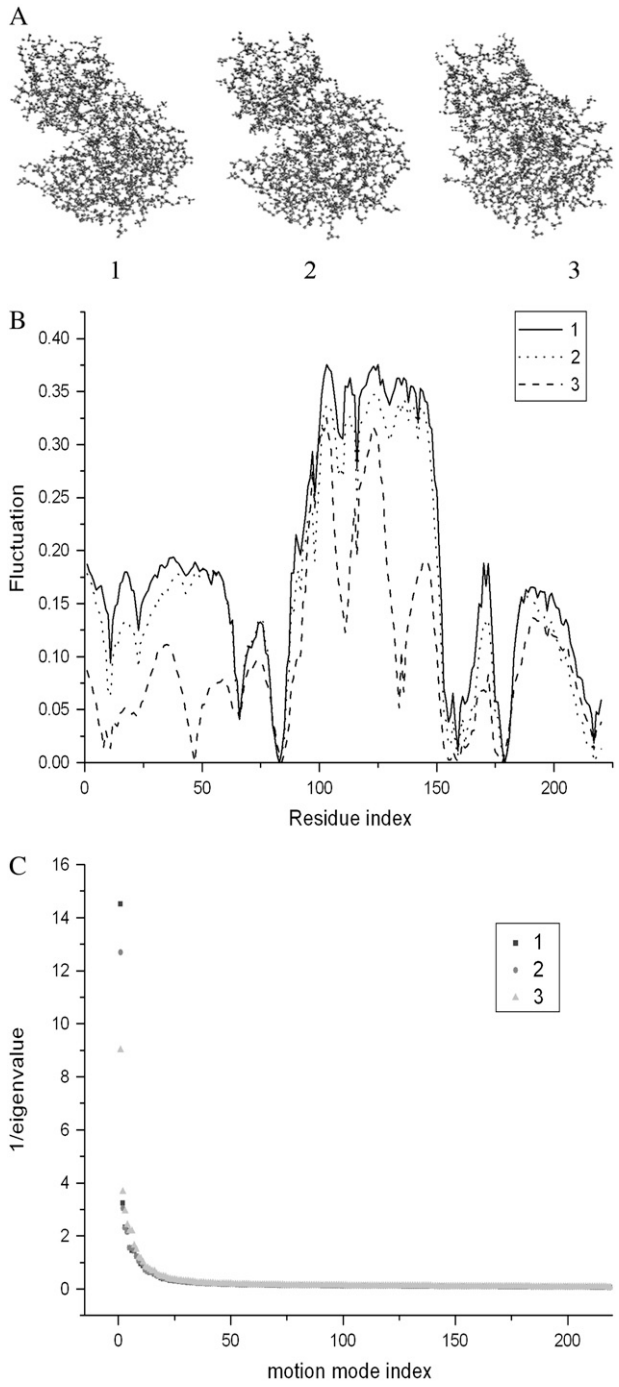

FIGURE 9 (*A*) Three structures at varying degrees of open-closed. (*B*) The slowest motion mode of three structures. (*C*) The reciprocal of eigenvalues for all motion modes of three structures.

structures are shown in Fig. 9 *B* and the reciprocals of their eigenvalues are shown in Fig. 9 *C*. Only for the first eigenvalue, corresponding to the slowest motion mode, is there a little difference between them. The other ranked eigenvalues of these three structures are very similar.

The degree of the open-closed motion is distinct, as shown in Fig. 9 *A*, and the motion modes differ correspondingly. When the conformation changed from the open form to the closed form, consulting with the eigenvalue, the motion modes changed remarkably. In particular, when the conformation changed from structure 2 to structure 3, the fluctuation of some residues decreased observably, as mentioned above.

In addition, the results of GNM analysis for several equilibrated structures selected randomly from a trajectory are very similar. It means that the partial change of the conformation does not affect the motion mode, which is considered from a holistic viewpoint.

From these analyses, it can be deduced that the open-closed transition is encoded in the structure, but the degree of this motion is related to the structure. When the GlnBP is free of the ligand, it has a remarkable open-closed motion and the extent of this motion is limited. Regarding the closed conformation of GlnBP with Gln bound, the trend of the open-closed transition is reduced.

Analysis of the influence of the ligand on the domain motions

The slowest mode shapes of the closed-apo and the closed-ligand structures are very similar as shown in Fig. 4. It means that the ligand has little influence on the tendency of the open-closed domain movement.

In the ligand-bound crystal structure, Gln is bound in a pocket formed between the two domains. The ligand is completely buried in the protein without any solvent accessibility. Gln is stabilized by hydrogen bonds and ionic interactions with Asp^{10}, Gly^{68}, Thr^{70}, Ala^{67}, Asp^{157}, Arg^{75}, Lys^{115}, Gly^{119}, and His^{156} (8). In the conventional GNM method, two residues will be connected by a harmonic spring when the distance between them is $<r_c$ (here 7.3 Å). With this method, Gln has harmonic force with Gly^{68}, Ile^{69}, Thr^{70}, Thr^{118}, Gly^{119}, Ser^{120}, and Asp^{157}, but has no harmonic force with Asp^{10}, Ala^{67}, Arg^{75}, Lys^{115}, Gly^{119}, and His^{156}, which have hydrogen bonds or static-electric interactions with the ligand in the crystal structure. To emphasize the influence of Gln, in our work we add springs connecting them with Gln, and then we calculate the modes of the motions based on this new network. The slowest mode, which corresponds to the open-closed transition, is shown in Fig. 10 *A*.

Furthermore, we doubled the spring constants between Gln and the connected residues and maintained the usual spring constant between other residues for highlighting the influence of Gln on the protein. The modes of the motions were calculated and the slowest mode was also shown in Fig. 10 *A*.

The slowest modes of the closed-ligand and the closed-apo structures obtained with the conventional GNM were also shown in Fig. 10 *A*. From the comparison between these results, it was shown that the shapes of these modes are extremely similar. The reciprocals of each eigenvalue corresponding to all motion modes for these structures are shown in Fig. 10 *B*. There is a small difference only between the first

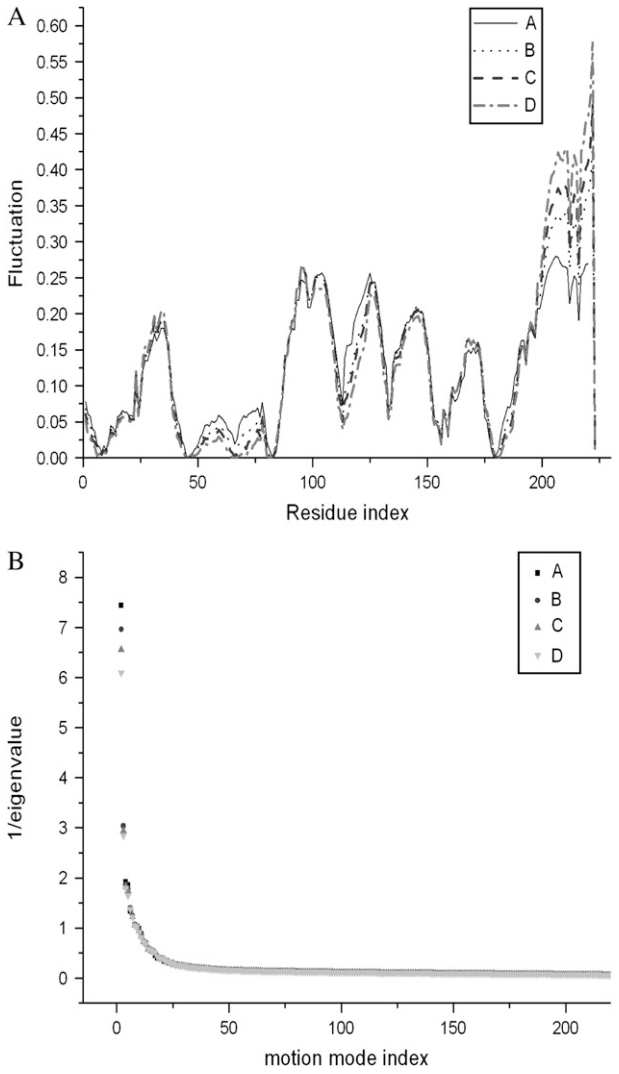

FIGURE 10 (*A*) Slowest mode shapes of the closed-ligand structure calculated by three methods: the conventional GNM method (series *B*); the method in which the residues are connected with the ligand by a spring, either within a distance cutoff or having a hydrogen-bond or static-electric interaction with the ligand (series *C*); and the method in which the spring constants between Gln and the residues connected are doubled (series *D*). All three methods are compared with the slowest mode of the closed-apo structure (series *A*). (*B*) The reciprocal of all eigenvalues for these four structures. (The definition of series *A–D* is the same as in Fig. 10 *A*.)

eigenvalues, which correspond to the slowest motion mode, and the maximal and minimal values are 7.44 and 6.09, respectively.

From the discussion above, a conclusion can be drawn that the ligand has little influence on the tendency of the domain transition. This result is consistent with the viewpoint of Lu and Ma (29), who found that the topology of a molecule plays a more dominant role in determining the low-frequency motions than the absolute values of strength and direction of local interactions. In our work, the ligand Gln only affects the local structure but does not change the motion nature of GlnBP. Therefore, the low-frequency motions of the closed-apo and the closed-ligand structures are fairly similar.

CONCLUSIONS

The open-closed conformation transition is important for GlnBP to transport Gln from the outer periplasmic space into the cytoplasmic space via cytomembrane. This work shows how simple coarse-grained methods can investigate the mechanism of the function motions. The same motion hinge axes for the open-apo, closed-apo, and closed-ligand forms are found. The open-closed motion mainly appears as the movement of the top region of the small domain. The first two slow modes of the closed-apo form and the closed-ligand form are completely similar, implying that the ligand has little influence on the domain motions. The open-closed conformation transition has a strong relation to the topological structure of the protein; in particular, the contact situation between the large domain and the small domain plays a key role in the conformation transition. This is consistent with the view that the general shape, instead of the local interactions, plays an important role in low-frequency motions. The fastest six modes of the three structures are also similar, indicating that their local structures are similar. This can be confirmed by the cross-correlation maps, which show that the residues within the domains move in a highly coupled way during the open-closed transition. The peaks in the fastest six modes are the most packed residues involved in the interactions with the membrane-bound components.

This work was supported in part by grants from the National Natural Science Foundation of China (10574009 and 30400087) and the Specialized Research Fund for the Doctoral Program of Higher Education (20040005013).

REFERENCES

1. Buchanan, S. K. 2001. Type I secretion and multidrug efflux: transport through the TolC channel-tunnel. *Trends Biochem. Sci.* 1:3–6.
2. Ferguson, A. D., and J. Deisenhofer. 2002. TonB-dependent receptors: structural perspectives. *Biochim. Biophys. Acta.* 1565:318–332.
3. Ames, G. F.-L. 1986. Bacterial periplasmic transport systems: structure, mechanism, and evolution. *Annu. Rev. Biochem.* 55:397–425.
4. Gerstein, M., A. M. Lesk, and C. Chothia. 1994. Structural mechanisms for domain movements. *Biochemistry.* 33:6739–6749.
5. Sharff, A. J., L. E. Rodseth, J. C. Spurlino, and F. A. Quiocho. 1992. Crystallographic evidence of a ligand-induced hinge-twist mutation between the two domains of the maltodextrin binding protein involved in active transport and chemotaxis. *Biochemistry.* 31:10657–10663.
6. Oh, B. H., J. Pandit, C. H. Kang, K. Nikaido, S. Gokcen, G. F.-L. Ames, and S. H. Kim. 1993. Three-dimensional structures of the periplasmic lysine/arginine/ornithine-binding protein with and without a ligand. *J. Biochem.* (*Tokyo*). 268:11348–11355.
7. Flocco, M. M., and S. L. Mowbray. 1994. The 1.9 Å X-ray structure of a closed unliganded form of the periplasmic glucose/galactose receptor from *Salmonella typhimurium*. *J. Biol. Chem.* 268:8931–8936.

8. Sun, Y. J., J. Rose, B. C. Wang, and C. D. Hsiao. 1998. The structure of glutamine-binding protein complexed with glutamine at 1.94 Å resolution: comparisons with other amino acid binding proteins. *J. Mol. Biol.* 278:219–229.
9. Rader, A. J., and I. Bahar. 2004. Folding core predictions from network models of proteins. *Polymers.* 45:659–668.
10. Yang, L. W., and I. Bahar. 2005. Coupling between catalytic site and collective dynamics: a requirement for mechanochemical activity of enzymes. *Structure.* 13:893–904.
11. Sun, T. G., J. P. Hu, C. H. Li, W. Z. Chen, and C. X. Wang. 2005. A molecular dynamics simulation study of glutamine-binding protein. *J. Mol. Struct. THEOCHEM.* 725:9–16.
12. Pang, A., Y. Arinaminpathy, M. S. P. Sansom, and P. C. Biggin. 2003. Interdomain dynamics and ligand binding: molecular dynamics simulations of glutamine binding protein. *FEBS Lett.* 550:168–174.
13. Haliloglu, T., I. Bahar, and B. Erman. 1997. Gaussian dynamics of folded proteins. *Phys. Rev. Lett.* 79:3090–3093.
14. Bahar, I., A. R. Atilgan, and B. Erman. 1997. Direct evaluation of thermal fluctuations in proteins using a single- parameter harmonic potential. *Fold. Des.* 2:173–181.
15. Atilgan, A. R., S. R. Durell., R. L. Jernigan, M. C. Demirel, O. Keskin, and I. Bahar. 2001. Anisotropy of fluctuation dynamics of protein with an elastic network model. *Biophys. J.* 80:505–515.
16. Kundu, S., and R. L. Jernigan. 2004. Molecular mechanism of domain swapping in proteins: an analysis of slower motions. *Biophys. J.* 86: 3846–3854.
17. Wang, Y., A. J. Rader, I. Bahar, and R. L. Jernigan. 2004. Global ribosome motions revealed with elastic network model. *J. Struct. Biol.* 147:302–314.
18. Keskin, O., S. R. Durell, I. Bahar, R. L. Jernigan, and D. G. Covell. 2002. Relating molecular flexibility to function: a case study of tubulin. *Biophys. J.* 83:663–680.
19. Keskin, O., R. L. Jernigan, and I. Bahar. 2000. Proteins with similar architecture exhibit similar large-scale dynamic behavior. *Biophys. J.* 78:2093–2106.
20. Jernigan, R. L., M. C. Demirel, and I. Bahar. 1999. Relating structure to function through the dominant slow modes of motion of DNA topoisomerase II. *Int. J. Quantum Chem.* 75:301–312.
21. Haliloglu, T., O. Keskin, B. Y. Ma, and R. Nussinov. 2005. How similar are protein folding and protein binding nuclei? Examination of vibrational motions of energy hot spots and conserved residues. *Biophys. J.* 88:1552–1559.
22. Demirel, M. C., A. R. Atilgan, R. L. Jernigan, B. Erman, and I. Bahar. 1998. Identification of kinetically hot residues in proteins. *Protein Sci.* 7:2522–2532.
23. Micheletti, C., P. Carloni, and A. Maritan. 2004. Accurate and efficient description of protein vibrational dynamics: comparing molecular dynamics and Gaussian models. *Proteins.* 55:635–646.
24. Doruker, P., A. R. Atilgan, and I. Bahar. 2000. Dynamics of proteins predicted by molecular dynamics simulations and analytical approaches: application to a-amylase inhibitor. *Proteins.* 40:512–524.
25. Hsiao, C. D., Y. J. Sun, J. Rose, and B. C. Wang. 1996. The crystal structure of glutamine-binding protein from *Escherichia coli*. *J. Mol. Biol.* 262:225–242.
26. Tamm, F., and Y.-H. Sanejouand. 2001. Conformational change of protein arising from normal mode calculations. *Protein Eng.* 14:1–6.
27. Kundu, S., J. S. Melton, D. C. Sorensen, and G. N. Phillips. 2002. Dynamics of proteins in crystals: comparison of experiment with simple models. *Biophys. J.* 83:723–732.
28. Bahar, I., A. R. Atilgan, M. C. Demirel, and B. Erman. 1998. Vibrational dynamics of proteins: significance of slow and fast modes in relation to function and stability. *Phys. Rev. Lett.* 80:2733–2736.
29. Lu, M. Y., and J. P. Ma. 2005. The role of shape in determining molecular motions. *Biophys. J.* 89:2395–2401.