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

Dashuai Lv, Weikang Gong, Yue Zhang, Yang Liu, Chunhua Li *

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.

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-parallel 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

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

http://dx.doi.org/10.1016/j.chemphys.2017.05.019
0301-0104/

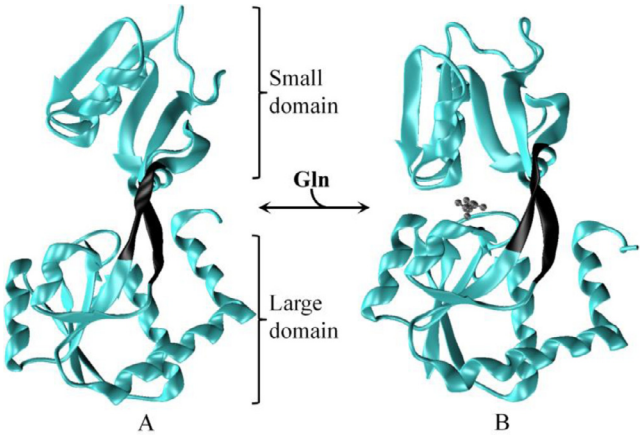

Fig. 1. The structures of GlnBP at the ligand-free open state (A) and the ligand-bound closed state (B). The hinge regions are highlighted with black color. The ligand glutamine is shown in the scaled ball and stick model.

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 coarse-grained 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 $R_O^{(0)}$ (open state) and $R_C^{(0)}$ (closed state) are used to generate the structures of the intermediate conformers, $R_O^{(k)}$ 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:

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

The conformations are represented as $3N$-dimensional vectors, corresponding to the coordinates of the C_α atoms. The distance vector $d^{(k)}$ between the pair of conformations produced at the kth 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 $R_O^{(0)}$ and $R_C^{(0)}$ structures.

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

The deformation vectors $v_O^{(k)}$ and $v_C^{(k)}$ used to produce the kth conformations are defined and calculated by the following formula as:

$$\begin{aligned} v_O^{(k)} &= R_O^{(k)} - R_O^{(k-1)} \\ v_C^{(k)} &= R_C^{(k)} - R_C^{(k-1)} \end{aligned} \tag{4}$$

The root mean square deviation (RMSD) at kth iteration can be written as:

$$RMSD(R_O^{(k)}, R_C^{(k)}) = \frac{|d^{(k)}|}{\sqrt{N}} \tag{5}$$

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 kth conformation $R_O^{(k)}$:

$$R_O^{(k)} = R_O^{(k-1)} + v_O^{(k)} = R_O^{(k-1)} + s_O^{(k)} \sum_{i=1}^{m_O^{(k)}} (d^{(k-1)} \cdot u_{iO}^{(k)}) u_{iO}^{(k)} \tag{6}$$

A similar expression holds for $R_C^{(k)}$. In this equation, $u_{iO}^{(k)}$ $(1 \leqslant i \leqslant m_O^{(k)})$ are the dominant, i.e., low frequency eigenvectors obtained by the ANM for the conformation $R_O^{(k-1)}$. The step sizes $s_O^{(k)}$ and $s_C^{(k)}$ are simultaneously determined at iteration k, as a fraction f (scaling factor) of those, $s_{Om}^{(k)}$ 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].