



Article

# Domain Motions and Functionally-Key Residues of L-Alanine Dehydrogenase Revealed by an Elastic Network Model

Xing-Yuan Li, Jing-Chao Zhang \*, Yan-Ying Zhu and Ji-Guo Su \*

Received: 26 September 2015; Accepted: 2 December 2015; Published: 9 December 2015

Academic Editor: Christo Christov

College of Science, Yanshan University, Qinhuangdao 066004, China; lxy@ysu.edu.cn (X.-Y.L.); yywlxzyy@163.com (Y.-Y.Z.)

\* Correspondence: ysu-zjc@163.com (J.-C.Z.); jiguosu@ysu.edu.cn (J.-G.S.);  
Tel./Fax: +86-335-805-7027 (J.-C.Z. & J.-G.S.)

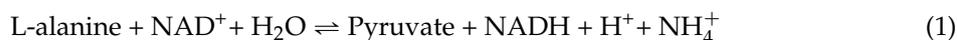
**Abstract:** *Mycobacterium tuberculosis* L-alanine dehydrogenase (L-MtAlaDH) plays an important role in catalyzing L-alanine to ammonia and pyruvate, which has been considered to be a potential target for tuberculosis treatment. In the present work, the functional domain motions encoded in the structure of L-MtAlaDH were investigated by using the Gaussian network model (GNM) and the anisotropy network model (ANM). The slowest modes for the open-apo and closed-holo structures of the enzyme show that the domain motions have a common hinge axis centered in residues Met133 and Met301. Accompanying the conformational transition, both the 1,4-dihydronicotinamide adenine dinucleotide (NAD)-binding domain (NBD) and the substrate-binding domain (SBD) move in a highly coupled way. The first three slowest modes of ANM exhibit the open-closed, rotation and twist motions of L-MtAlaDH, respectively. The calculation of the fast modes reveals the residues responsible for the stability of the protein, and some of them are involved in the interaction with the ligand. Then, the functionally-important residues relevant to the binding of the ligand were identified by using a thermodynamic method. Our computational results are consistent with the experimental data, which will help us to understand the physical mechanism for the function of L-MtAlaDH.

**Keywords:** *Mycobacterium tuberculosis* L-alanine dehydrogenase; domain motions; functionally-key residues; Gaussian network model; anisotropy network model; thermodynamic cycle method

## 1. Introduction

Tuberculosis is still a serious threat to human health worldwide, which is caused by the infection of the *Mycobacterium tuberculosis* (*Mt*) [1]. World Health Organization (WHO) has pointed out that at least one-third of the population in the world could be infected with latent *Mt* [2]. Up to now, the treatment of tuberculosis is still a difficult challenge due to the emergence of the drug-resistant strains of *Mt* [3]. At present, it is found that many anti-tubercular drugs become ineffective for the treatment of *Mt*, and thus, the development of new anti-tubercular drugs to treat this disease is urgent [4]. When the pathogens are under nutrient starvation regimes or anaerobic growth conditions, Rv2780 encoding L-alanine dehydrogenase (L-AlaDH) is overexpressed in persistent *Mt* [5]. The production and activity of L-AlaDH were improved when *Mt* was transferred from the aerobic to the anaerobic growth conditions [6]. L-AlaDH can catalyze the 1,4-dihydronicotinamide adenine dinucleotide

(NAD)-dependent conversion of L-alanine to ammonia and pyruvate or convert pyruvate to L-alanine in the reverse direction. The catalytic mechanism can be expressed as [7]:



Because the crystallographic structure of human L-AlaDH has not been determined, L-alanine dehydrogenase from *Mycobacterium tuberculosis* (L-MtAlaDH) is studied here. L-MtAlaDH is a secretory antigen, which is closely relevant to bacterial persistence during the persistent stage of infection [8]. Therefore, L-MtAlaDH is considered to be a potential target for the development of anti-tubercular drugs [9].

Two forms of structures for L-MtAlaDH have been determined in the Protein Data Bank (PDB), including the dimer (Figure 1C) and the hexamer (Figure 1D) states. The hexameric molecule is made of three-fold non-crystallographic axes, and two monomers nearby form the fold axis. As shown in Figure 1A, each monomer of L-MtAlaDH is composed of a substrate-binding domain (Residues 1–128 and 309–371) and a NAD-binding domain (Residues 129–308), which are connected by two  $\alpha$ -helices [9]. The NAD-binding domain includes a central seven-stranded  $\beta$ -sheet, as well as several surrounding  $\alpha$ -helices. The substrate-binding domain consists of an eight-stranded  $\beta$ -sheet and several  $\alpha$ -helices [9]. The crystal structure of the open-apo and closed-holo states of the protein have been determined by X-ray crystallography, which can be downloaded from the Protein Data Bank (PDB) with the access codes 2VOE and 2VHZ, respectively [8,9]. The experimental studies have demonstrated that upon the binding of the coenzyme, L-MtAlaDH undergoes a large-scale conformational transition from the open-apo to the closed-holo states, in which the substrate-binding domain (SBD) rotates by  $16^\circ$  toward the NAD-binding domain (NBD) [9]. Considering that the conformational transitions between open and closed structures are essential for the function of L-MtAlaDH, many experimental and theoretical studies have focused on the investigation of the functional domain motions and identification of the related key residues of the protein. The open-closed domain motions of L-MtAlaDH upon the binding of the coenzyme have been studied with molecular dynamics (MD) simulations at the atomic level [10]. The simulation results have shown that the conformational transition of L-MtAlaDH is achieved through the open-closed and twisting motions of the NBD and the SBD. Based on the simulation results, it is proposed that two loops (Residues 94–99 and 238–251) are responsible for the binding of NAD, and another loop (Residues 267–293) is crucially responsible for the binding of substrate [10]. The experiments of Agren *et al.* have shown that L-MtAlaDH completely lost its catalytic activity when the conserved active site residues His96 and Asp270 were mutated. These two residues play an essential role in the enzymatic reaction at the active site and exhibit strong stability in biological evolutionary process [9].

In this work, the domain motions and the related functionally-key residues encoded in the structure of L-MtAlaDH were investigated by using the coarse-grained Gaussian network model (GNM) and the anisotropy elastic network model (ANM) [11–13]. The GNM and ANM are simple yet effective methods to explore the intrinsic dynamics of proteins [14–16]. It has been proven in numerous application studies that the low-frequency motion modes calculated by GNM and ANM represent the large-scale collective motions usually relevant to the functions of the protein [17–19]; whereas the high-frequency motion modes reflect the geometric irregularity in the protein structure, and the residues motivated in these modes are considered to be crucial for the protein stability [20,21]. In the present work, the functional domain motions of L-MtAlaDH were studied through analyzing the slow motion modes of GNM and ANM, and the key residues responsible for the stability of the protein were identified by calculating the fast modes of GNM and ANM. Besides that, the functionally-important residues responsible for substrate binding of the protein were revealed by using a thermodynamic method proposed by Su *et al.* [22,23].



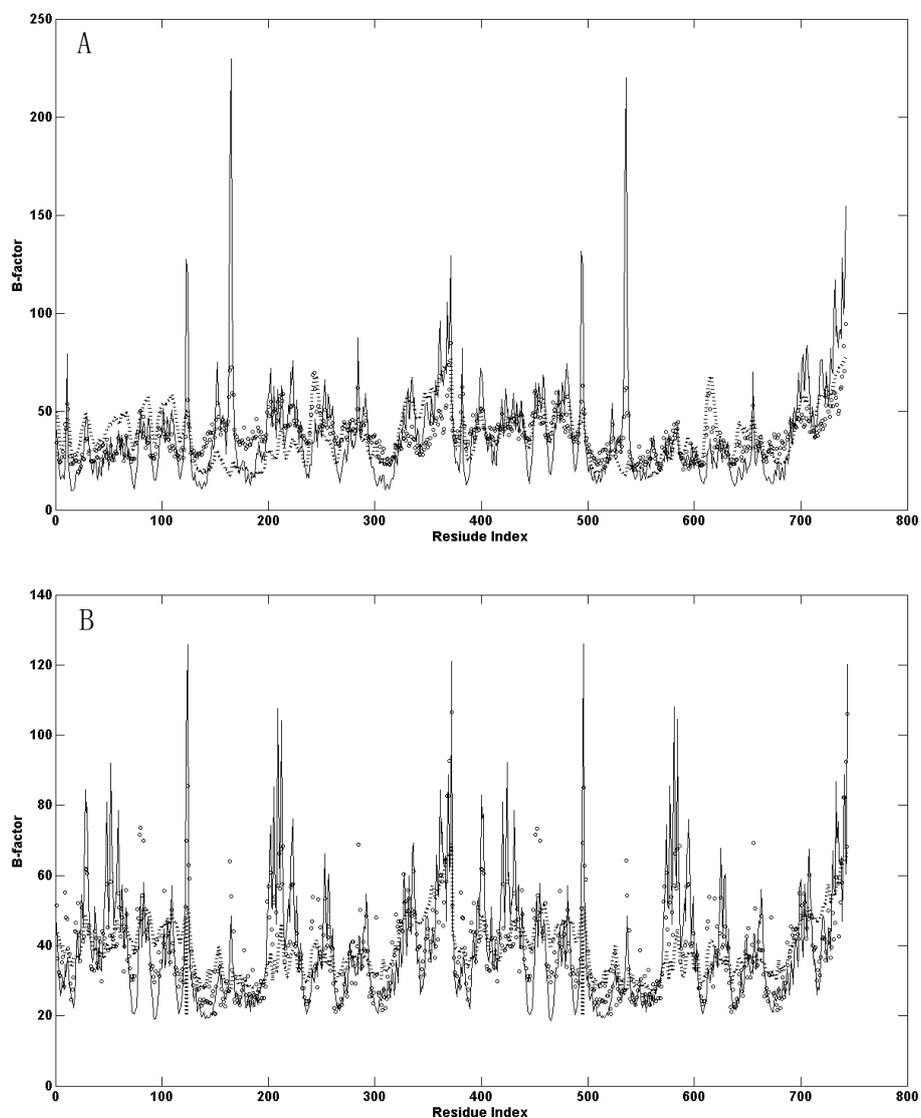
**Figure 1.** Open-apo (one monomer of 2VOE) (A) and closed-holo (one monomer of 2VHZ) (B) structures of *L-MtAlaDH*. The bound ligand is shown as a yellow ball-and-stick model in the closed-holo structure. The secondary substructures  $\alpha$ -helices and  $\beta$ -sheets of the protein are displayed in red and blue colors, respectively. The protein is divided into the substrate-binding domain and the 1,4-dihyronicotinamide adenine dinucleotide (NAD)-binding domain, which is marked in the structure; (C) The dimeric *L-MtAlaDH* molecule. One monomer is shown in the solid ribbon model, and the other is shown in gray; (D) The hexameric *L-MtAlaDH* molecule. One monomer is shown in the solid ribbon model, and others are shown in gray.

## 2. Results and Discussion

### 2.1. Comparison of the Predicted Mean-Square Fluctuations with the Experimental Temperature Factors (B-Factors)

According to Equation (7), the B-factors of the residues in the open-apo and closed-holo structures are calculated. In Equation (7), the only adjustable parameter is  $\gamma$ , which is determined by normalizing the computational B-factors to the experimental data. The resulting  $k_B T/\gamma$  values used for the open-apo and closed-holo structures in the GNM are 3.68 and 6.90  $\text{\AA}^2$ , respectively. The resulting  $k_B T/\gamma$  values used for the open-apo and closed-holo structures in the ANM are 0.45 and 11.64  $\text{\AA}^2$ , respectively. Figure 2 shows the comparison between the calculated and the experimental B-factors for the two studied systems. It is found that for GNM, the correlation coefficient between the calculated and experimental Temperature Factors (B-Factors) is 0.50 for the open-apo structure

and 0.63 for the closed-holo structure, respectively. For ANM, the correlation coefficients are 0.49 and 0.51 for the open-apo and closed-holo structures, respectively. The correlation coefficient for the two proteins studied in this work is similar to that for other protein [12,24], which provides the validity for the study of *MtAlaDH* by using the GNM and the ANM.

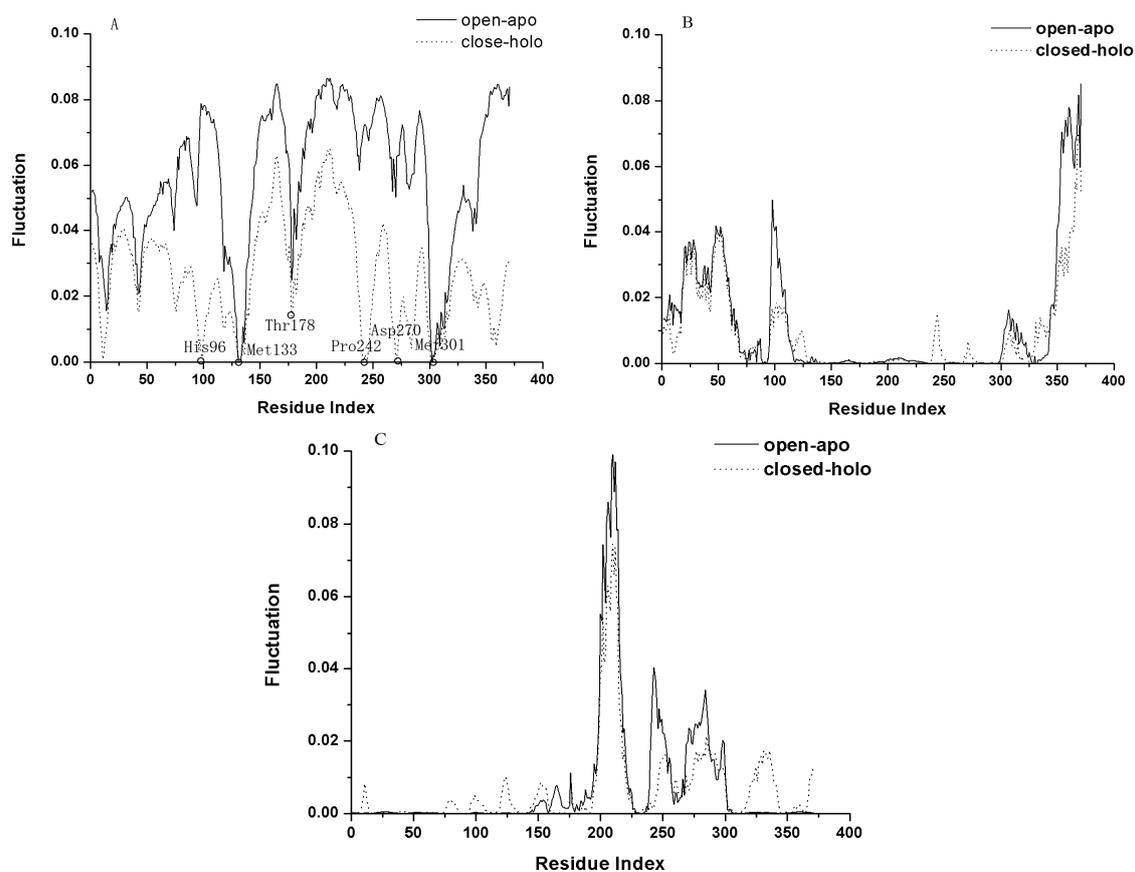


**Figure 2.** (A) The calculated B-factors based on the Gaussian network model (GNM) (circle curve) and the anisotropy elastic network model (ANM) (solid curve) compared to the experimental data (dotted curve) for the open-apo structure of *L-MtAlaDH* (PDB code: 2OVE); and (B) The calculated B-factors based on GNM (circle curve) and ANM (solid curve) compared to the experimental data (dotted curve) for the closed-holo structure of *L-MtAlaDH* (PDB code: 2VHZ).

## 2.2. The Slow Motion Modes

In the GNM, the slow motion modes represent the large-scale collective motions, which are usually relevant to the functions of the protein [25]. Figure 3A shows the first slowest mode calculated by GNM for the open-apo and closed-holo structures of *L-MtAlaDH*. From Figure 3A, it is found that the first modes for the two studied proteins have a common hinge axes located in Met133 and Met301, with their fluctuation values equal to zero. The fluctuation profiles in Figure 3A also exhibit that the whole structure of the protein is divided into two distinct domains with large fluctuations.

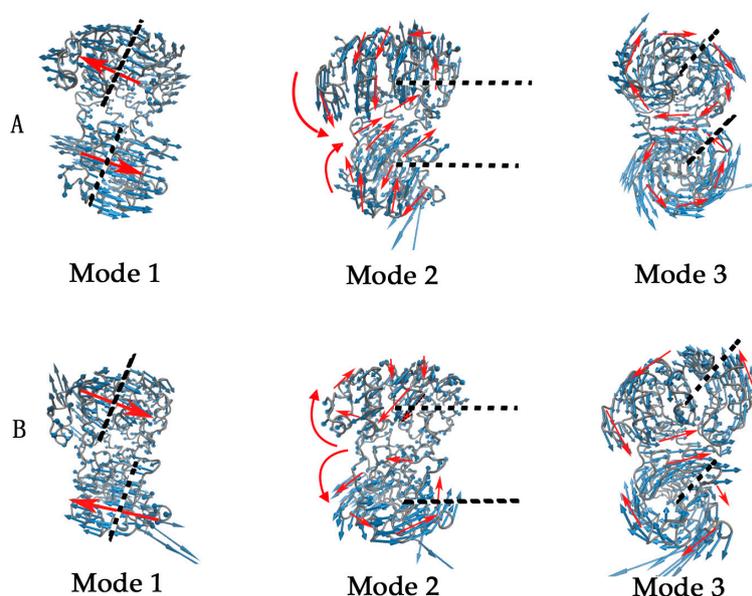
Comparing to the slowest mode of the open-apo structure, it can be found that the fluctuation of some residues around the residue His96 is reduced remarkably in the closed-holo structure. This is consistent with the simulation results that the root mean-square fluctuation (RMSF) of the residues around the residue His96 becomes small in the closed state of the protein [10]. It is also found that the fluctuation value of the residues around Asp270 and Pro242 is also approximate to zero, although those residues do not belong to the hinge region (Residues 126–133, 304–320) [10]. These two residues are located at the ligand-binding cleft. This result indicates that the ligand-binding cleft become stable in the closed structure of the protein. Besides the two residues, the residues Met133, Thr178 and Met301 of the NAD-binding domain in the closed-holo structure are located in the minima of the fluctuation curve in Figure 3A. These residues are the ligand-binding sites and are locate in the jaws of the ligand-binding pocket [8]. The crystal structure of the closed-holo state of the protein exhibits that the nicotinamide ring of the ligand is bound in the residues Met133, Ser134, Ala137, Ile267, Asp270 and Met301, and the pyrophosphate moiety forms contacts with the main-chain atoms of residues 178 and 179 [9]. The decrease of the fluctuation of these residues indicates that they are more stable in the closed-holo structure than the open-apo structure. Figure 3B,C show the second and third slowest modes calculated by the GNM for the open-apo and closed-holo structures of *L-MtAlaDH*, respectively. The overlap between the open-apo and closed-holo structure of the second and third slowest modes is better than the first mode. In Figure 3B,C, the amplitude of the open-apo and close-holo structure is similar. The large fluctuation of the second slowest mode lies on the C-terminal (Residues 342–371) and N-terminal (Residues 21–63) of the protein, and that of the third slowest mode lies in the middle domain of the protein.



**Figure 3.** The first three slowest modes of the open-apo (solid line) and the closed-holo (dotted line) structures of *L-MtAlaDH* are shown in (A–C), respectively. In (A), the important residues located in the minima of the fluctuation curve are marked by circles.

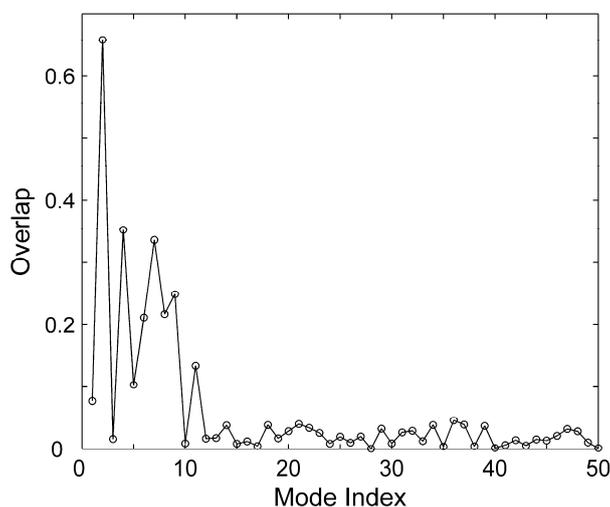
It should be noted that besides the structures studied in this work, there are seven other structures of this protein that have been deposited in PDB. The first three slowest modes of these structures of this protein have also been calculated by using the GNM (data not shown). The mode shapes of these structures are found to be similar to those of the open-apo or closed-holo structures studied in this work, which indicates that the functional motions are largely determined by the structural topology of the protein instead of the detailed local structures.

GNM can only describe the amplitude of residue fluctuations, but cannot give the direction of the fluctuations. In order to reveal the direction of the motions, ANM is used for the open-apo and closed-holo structures. The first three slowest normal modes calculated by ANM are shown in Figure 4A,B for the open-apo and the closed-holo structures of *L-MtAlaDH*, respectively. In Figure 4, the amplitude and direction of the motions are represented by the length and direction of the blue arrows, respectively. The red arrows and black dashed lines denote the motion direction and the motion axis, respectively. The first slowest mode (Figure 4 (Mode 1)) of the two structures exhibits a twist motion of the two domains in the proteins. In this motion, the NAD-binding domain (NBD) and the substrate-binding domain (SBD) rotate reversely around their respective axes. The rotation axes are displayed in Figure 4 (Mode 1). This motion mode might be important for the two domains in the proteins to adjust their relative position to bind or dissociate the ligand. The long blue arrows represent the amplitude of the motions. This is consistent with the first slowest mode of GNM. The second slowest mode describes a hinge-bending motion of the NBD and SBD for the two studied proteins, as shown in Figure 4 (Mode 2). The  $\alpha$ -helices connecting the NBD and SBD serve as the hinge of the bending motion. The motion results in the opening and closing of the ligand-binding pocket. The studies by MD simulations also showed that the SBD rotates about  $14.2^\circ$  around the hinge axis toward the NBD of the protein [10]. This is consistent with the second slowest motion mode of GNM, where the fluctuation of the C-terminus is larger than other residues. The third slowest mode (Figure 4 (Mode 3)) represents a reverse rotation of the NBD and SBD around their respective axes. Additionally, the rotation axes of these two domains are marked by black dashed lines in this figure.



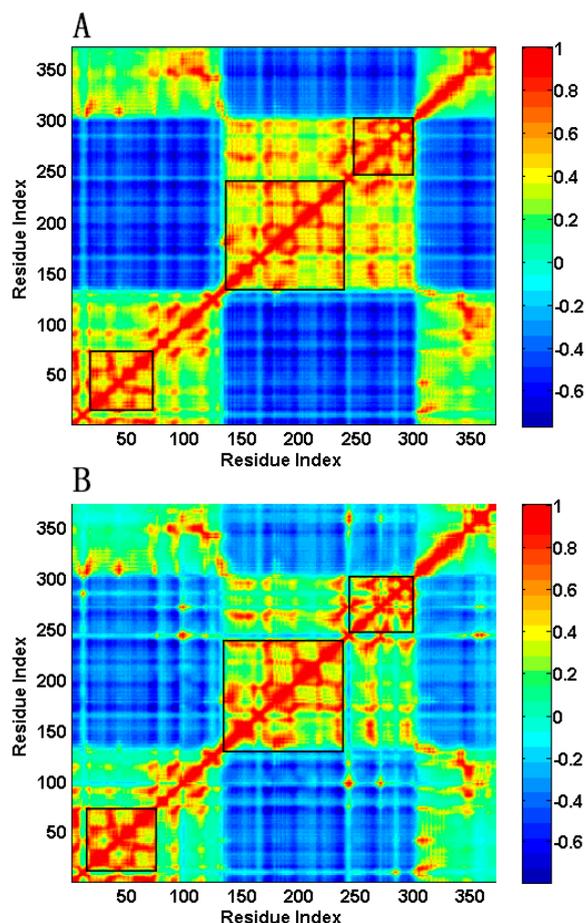
**Figure 4.** The first three slowest motion modes for the open-apo (**A**) and closed-holo (**B**) structure of *L-MtAlaDH*. These figures were generated by visual molecular dynamics (VMD) [26]. In this figure, the protein structures are displayed in gray tubes. The amplitude and direction of the motions are represented by the length and direction of the blue arrows, respectively. The red arrows and black dashed lines denote the motion direction and the motion axis, respectively.

In order to investigate whether the normal modes can reveal the conformational change observed by the experiment, the overlap between each mode and the experimental conformational change was calculated using Equation (17). Figure 5 shows the overlap for the first 50 modes. It is found that Mode 2 mostly contributes to the conformational change, where the value of the overlap is 0.658.



**Figure 5.** The overlap between the first 50 modes and the conformational change.

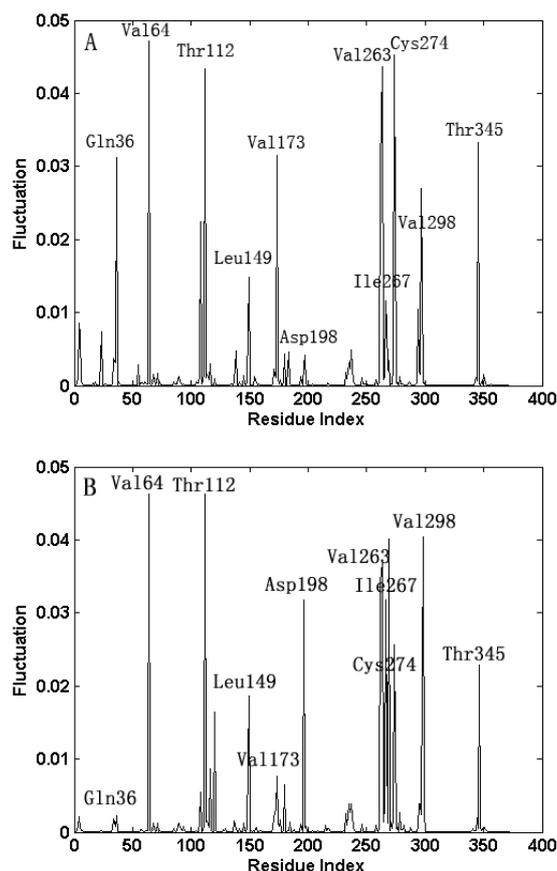
In order to analyze the coupling between the residues motions, the cross-correlation between residue fluctuations is calculated using Equation (10). Because the low-frequency normal modes are usually relevant to protein functions, the first 30 slowest modes are considered in the calculation. The calculation results are displayed in Figure 6. The value of the cross-correlation ranges from  $-1$  to  $1$ , where the positive value represents the residues moving in the same direction and the negative value means the residues moving in the opposite direction. In Figure 6, the regions in orange-red color represent positive correlation and the blue regions denote negative correlation. It is found that the cross-correlation map, both for the open-apo and closed-holo structures, is divided into five orange-red regions and four blue regions separated by the residues Met133 and Met301, which means that these two residues serve as the hinges for the correlated motions. This is consistent with the result obtained by the analysis of the slowest normal mode, in which these two residues act as the motion hinges with almost zero fluctuations. In Figure 6, the blue regions correspond to the negative correlation between the motions of NBD and SBD, illustrating the open-closed conformational transition of the protein. The orange-red region in the center of the maps (Residues 133–301) represents the residues of the NBD moving as a whole, whereas the other four orange-red regions indicate the residues of the SBD also moving in a coupled way. This result indicates that the two domains keep their rigid structure during the conformational transitions. Comparing the closed-holo structure to the open-apo structure of the protein, more contacts are formed for the residues in the jaws of the ligand-binding pocket. These residue contacts may reduce the correlation between domain motions. Therefore, the blue and orange-red colors are weakened in the correlation map for the closed-holo structure compared to that for the open-apo structure, as shown in Figure 6. From Figure 6B, it is found that the fluctuation correlations among the residues Arg15, Loops L1 (Residues 94–99), L2 (Residues 238–247) and L3 (Residues 267–270) are significantly improved in the closed-holo structure because of their interactions with the ligand. In Figure 6, there are several regions with high correlation values, marked by the rectangles. These regions correspond to the highly coupled movement of residues 134–241 and 243–300 of NBD and residues 16–74 of SBD.



**Figure 6.** The cross-correlation maps calculated by the first 30 slowest modes of GNM for the open-apo (A) and the closed-holo (B) structures of *L-MtAlaDH*. In this figure, blue regions indicate negative correlation and orange-red regions indicate positive correlation, as shown in the color bar on the right. The regions in the rectangle (Residues 16–74, 134–241, 243–300) indicate the highly coupled regions.

### 2.3. The Fast Motion Modes

In GNM, the fast motion modes reflect the geometric irregularity in the local structure, and the fluctuations in the fast motion modes are accompanied with the remarkable increases of the entropy. Then, the residues activated in the fast motion modes are regarded as the key residues that are important for the stability of the protein structure [21,25]. Figure 7 shows the eight fastest modes for the open-apo and closed-holo structures, respectively. It is found that the mode shapes for the open-apo and closed-holo structures are extremely similar, and the residues at the peaks of the fluctuation are almost the same, implying that the same set of residues is responsible for the stability of the open-apo and closed-holo structures of the protein. This result implies that the local structures are similar for these two proteins, and the domains of the protein keep their rigid structures during the open-closed motions. Figure 7 shows that the hot residues responsible for the stability of the proteins include Leu149, Val173, Asp198, Val263, Ile267, Cys274 and Val298 in NBD and Gln36, Val64, Thr112 and Thr345 in SBD. From the crystal structure of the proteins, it is found that these key residues are tightly packed, and most of them lie in the ligand-binding pocket, among which the residues Asp198 and Ile267 form polar interactions with the ligand NAD [8]. Our results imply that the residues crucial for protein stability also play important roles for the binding of the substrate to the protein.



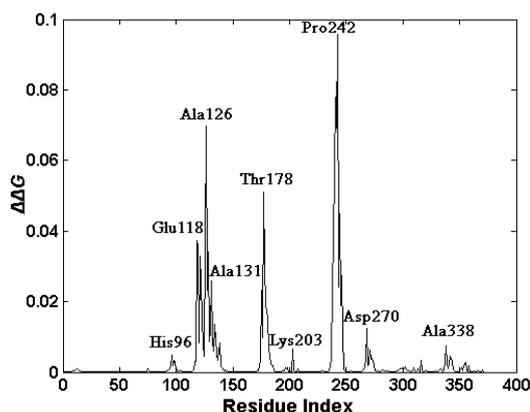
**Figure 7.** (A,B) The mode shape as the average of the eight fastest modes for the open-apo and closed-holo structure of *L-MtAlaDH*. The same key residue clusters identified by our method are marked by the residue name in the figure.

#### 2.4. The Functionally-Important Residues Identified with the Thermodynamic Method

In our work, the functionally-key residues responsible for the ligand binding were identified by using a thermodynamic method proposed by Su *et al.* [22,23,27]. The closed-holo structure of the protein is adopted as the complex structure, and the corresponding GNM was constructed. Then, the ligand NAD was removed from the closed-holo structure, and the corresponding GNM was also constructed. Based on these two constructed models, each residue of the protein was perturbed, and the  $\Delta\Delta G$  value in response to the perturbation was calculated according to the method proposed in the Methods Section. In our method, the free energy change  $\Delta\Delta G$  for each residue is the sum of the free energy changes for all of the spring perturbations involved in this residue. Residues with relatively high  $\Delta\Delta G$  values were regarded as the key residues involved in the ligand binding of the protein. The calculated result is shown in Figure 8. It was found that nine clusters of residues exhibit high  $\Delta\Delta G$  values, which are centered at the residues His96, Glu118, Ala126, Met133, Thr178, Lys203, Pro242, Asp270 and Ala338, respectively.

From Figure 8, it is found that almost all of these key residues are located at the ligand-binding site in the NBD. Tripathi *et al.* have proposed that NAD interacts extensively with enzyme through the interactions with Met301, Asp270, Ile267, Thr178, Lys203, Asp198, Met133, His96, Lys75 and Arg15 [8]. Most of our predicted residues belong to the residues directly interacting with NAD. The analysis of the slowest normal mode discussed in the previous section has identified Met133 to be the motion hinge. This residue plays an important role for the opening and closing of the ligand-binding pocket, and thus, it is crucial for the binding of the ligand to the protein. The residues His96 and Asp270 were identified as the conserved active site residues by Agren's group [9]. The residue Ala126

is also located in the hinge region between the two domains of the protein. In the slowest motion mode of the closed structure of the protein, the residues Thr178, Pro242 exhibit small fluctuations, which imply that these residues play important roles in the binding of the ligand, stabilizing the closed-holo structure of *L-MtAlaDH*. The functionally-key residues predicted by the thermodynamic method are consistent with the experimental observations and the results obtained by the analysis of the slow modes.



**Figure 8.**  $\Delta\Delta G$  values in response to residual perturbations. The residues of the clusters with peak values are marked.

### 3. Materials and Methods

#### 3.1. The Gaussian Network Model and the Anisotropy Elastic Network Model

In the Gaussian network model (GNM) [11,28], the tertiary structure of a protein is simplified as an elastic network. Each residue of the protein is represented by a point that is located at its  $C_\alpha$  atom, and all of the pairwise residues within a cutoff distance ( $8 \text{ \AA}$  is adopted here) are connected by a harmonic springs with a uniform force constant. Then, the internal energy of the protein can be expressed as [15]:

$$H = \frac{1}{2} \gamma \left[ \Delta R^T (\Gamma \otimes E) \Delta R \right] \quad (2)$$

where  $\gamma$  is the spring constant;  $\Delta R$  represents fluctuations of the residues;  $E$  is the identity matrix;  $\otimes$  is the direct product of the matrix;  $T$  represents the transpose of the matrix;  $\Gamma$  is the Kirchhoff matrix with  $N \times N$  elements that can be expressed as [11,15]:

$$\Gamma = \begin{cases} -1, & \text{if } i \neq j \text{ and } R_{ij} \leq r_c \\ 0, & \text{if } i \neq j \text{ and } R_{ij} > r_c \\ -\sum_{i,i \neq j} \Gamma_{ij}, & \text{if } i = j \end{cases} \quad (3)$$

where  $R_{ij}$  is the separation between the  $i$ -th and  $j$ -th  $C_\alpha$  atoms and  $r_c$  is the cutoff distance.

The cross-correlation between the fluctuations of the  $i$ -th and  $j$ -th residues is written as:

$$\langle \Delta R_i \cdot \Delta R_j \rangle = \frac{3k_B T}{\gamma} \left[ \Gamma^{-1} \right]_{ij} \quad (4)$$

When  $i = j$ , the mean-square fluctuation of the  $i$ -th residue can be written as:

$$\langle (\Delta R_i)^2 \rangle = \frac{3k_B T}{\gamma} \left[ \Gamma^{-1} \right]_{ii} \quad (5)$$

where  $k_B$  is the Boltzmann constant;  $T$  is absolute temperature.

The inverse of the Kirchhoff matrix can be calculated by diagonal decomposition as:

$$\Gamma^{-1} = U\Lambda^{-1}U^T \quad (6)$$

where  $U$  is an orthogonal matrix whose columns represent the eigenvectors of  $\Gamma$ ;  $\Lambda$  is a diagonal matrix, and the diagonal elements are the eigenvalues of  $\Gamma$ .

According to the Debye-Waller theory, the B-factor of the  $i$ -th residue can be calculated by:

$$B_i = 8\pi^2 \langle \Delta R_i \cdot \Delta R_i \rangle / 3 \quad (7)$$

The mean-square fluctuation of the  $i$ -th residue in the  $k$ -th mode is written by:

$$\langle \Delta R_i \cdot \Delta R_i \rangle_k = \frac{3k_B T}{\gamma} \lambda_k^{-1} [u_k]_i [u_k]_i \quad (8)$$

The cross-correlation between the fluctuations of different residues is written by:

$$\langle \Delta R_i \cdot \Delta R_j \rangle = \frac{3k_B T}{\gamma} \sum_{k=2}^N \lambda_k^{-1} [u_k]_i [u_k]_j \quad (9)$$

In the GNM, the cross-correlation is normalized as:

$$C_{ij} = \frac{\langle \Delta R_i \cdot \Delta R_j \rangle}{\left[ \langle \Delta R_i^2 \rangle \times \langle \Delta R_j^2 \rangle \right]^{1/2}} \quad (10)$$

For ANM, the motion mode of the protein is determined by the  $3N \times 3N$  Hessian matrix  $H$ .

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

The elements of  $H$  are a submatrix with a size of  $3 \times 3$ . The  $ij$ -th submatrix  $h_{ij}$  is:

$$h_{ij} = \begin{pmatrix} \frac{\partial^2 V}{\partial x_i \partial x_j} & \frac{\partial^2 V}{\partial x_i \partial y_j} & \frac{\partial^2 V}{\partial x_i \partial z_j} \\ \frac{\partial^2 V}{\partial y_i \partial x_j} & \frac{\partial^2 V}{\partial y_i \partial y_j} & \frac{\partial^2 V}{\partial y_i \partial z_j} \\ \frac{\partial^2 V}{\partial z_i \partial x_j} & \frac{\partial^2 V}{\partial z_i \partial y_j} & \frac{\partial^2 V}{\partial z_i \partial z_j} \end{pmatrix} \quad (12)$$

When  $i = j$ ,

$$h_{ij} = \begin{pmatrix} \frac{\partial^2 V}{\partial x_i^2} & \frac{\partial^2 V}{\partial x_i \partial y_i} & \frac{\partial^2 V}{\partial x_i \partial z_i} \\ \frac{\partial^2 V}{\partial y_i \partial x_i} & \frac{\partial^2 V}{\partial y_i^2} & \frac{\partial^2 V}{\partial y_i \partial z_i} \\ \frac{\partial^2 V}{\partial z_i \partial x_i} & \frac{\partial^2 V}{\partial z_i \partial y_i} & \frac{\partial^2 V}{\partial z_i^2} \end{pmatrix} \quad (13)$$

where the elements  $\frac{\partial^2 V}{\partial x_i \partial y_j}$ ,  $\frac{\partial^2 V}{\partial x_i \partial y_i}$  of the submatrix  $h_{ij}$  can be analytically expressed as:

$$\frac{\partial^2 V}{\partial x_i \partial y_j} = \frac{-\gamma(x_j - x_i)(y_j - y_i)}{r_{ij}^2} \Bigg|_{r_{ij}=r_{ij}^0} \quad (14)$$

$$\frac{\partial^2 V}{\partial x_i \partial y_i} = \gamma \sum_{j \neq i} \frac{(x_j - x_i)(y_j - y_i)}{r_{ij}^2} \Bigg|_{r_{ij}=r_{ij}^0} \quad (15)$$

where  $x$ ,  $y$  and  $z$  represent the coordinates of each atom.

### 3.2. Correlation Coefficient and Overlap

In our work, the correlation coefficient between the calculated and the experimental B-factors is written as:

$$\rho = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{[\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2]^{1/2}} \quad (16)$$

where  $x_i$  and  $y_i$  are the calculated and experimental values for the B-factor of the  $i$ -th  $C_\alpha$  atom,  $\bar{x}$  and  $\bar{y}$  are the mean values of  $x_i$  and  $y_i$  and  $N$  is the total number of  $C_\alpha$  atoms in the protein.

Because the open-apo and closed-holo structures of the protein have been determined by the X-ray crystallographic experiment, the conformational change can be calculated. Then, we investigated whether the slow motion modes obtained by ANM can account for the experimental conformational change, which can be evaluated by the overlap between the normal modes and the conformational change. The value of overlap can be calculated by:

$$overlap_i = \frac{\vec{\eta}_i \cdot \vec{\Delta R}_{exp}}{|\vec{\Delta R}_{exp}|} \quad (17)$$

where  $\vec{\eta}_i$  is the  $i$ -th normal mode,  $\vec{\Delta R}_{exp}$  is the experimental conformational change and  $|\vec{\Delta R}_{exp}|$  is the norm of  $\vec{\Delta R}_{exp}$ .

### 3.3. Thermodynamic Cycle Method

Base on the elastic network model (ENM), a thermodynamic method was developed by our group to reveal the functionally-important residues involved in the binding of ligand [22,23,27]. According to the theory of statistical physics, the vibrational entropy of ENM can be expressed as [25]:

$$S = \frac{\langle H \rangle - F}{T} = \frac{1}{T} \left[ \frac{3}{2} (N - 1) k_B T + k_B T \ln Z \right] = \frac{3}{2} (N - 1) k_B + k_B \ln Z \quad (18)$$

where  $\langle H \rangle$  is the average value of the Hamiltonian;  $F = -k_B T \ln Z$  is the Helmholtz free energy;  $T$  is the temperature;  $N$  is the number of residues in the protein;  $k_B$  is the Boltzmann constant;  $Z$  is the configurational integral part of the vibrational partition function that can be calculated by  $Z = \int \exp(-H/k_B T) d\{\Delta R\}$ ;  $\{\Delta R\}$  represents the 3N-dimensional column vector of fluctuations. When a perturbation of the spring between the  $i$ -th and  $j$ -th residues is introduced, the vibrational entropy of the protein will be changed as:

$$\Delta S = -\frac{\partial S}{\partial \gamma_{ij}} \Delta \gamma_{ij} \quad (19)$$

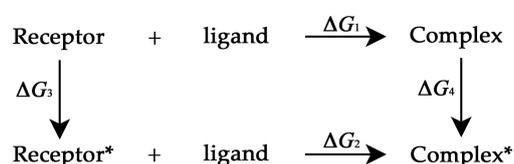
where  $\Delta\gamma_{ij}$  is the change of the force constant that accounts for the perturbation of the spring. According to Equation (18), the derivation of the entropy with respect to  $\gamma_{ij}$  can be expressed as [22,23]:

$$\frac{\partial S}{\partial \gamma_{ij}} = -\frac{1}{2T} \left( \langle (\Delta R_i)^2 \rangle + \langle (\Delta R_j)^2 \rangle - 2 \langle (\Delta R_i \cdot \Delta R_j)^2 \rangle \right) \quad (20)$$

In this method, all of the residues in the protein were perturbed one by one, and then, the change of the free energy for the protein-ligand binding was calculated.

In Figure 9,  $\Delta G_1$  denotes the ligand-binding free energy without the introduction of the perturbation, and  $\Delta G_2$  denotes the value of the binding free energy with a residue perturbation. In this work, the binding ligand is 1,4-dihydrnicotinamide adenine dinucleotide (NAD). The receptor\* and complex\* represent the systems where a residue perturbation is introduced. We want to calculate the change of the free energy resulting from the spring perturbation, *i.e.*,

$$\Delta\Delta G = \Delta G_2 - \Delta G_1 \quad (21)$$



**Figure 9.** The thermodynamic cycle,  $\Delta G_1$  denotes the ligand-binding free energy without the introduction of the perturbation, and  $\Delta G_2$  denotes the value of the binding free energy with a residue perturbation. The receptor\* and complex\* represent the systems where a residue perturbation is introduced. Two nonphysical processes were constructed, *i.e.*, Receptor  $\rightarrow$  Receptor\* and Complex  $\rightarrow$  Complex\*. The free energy changes associated with these two constructed processes are denoted as  $\Delta G_3$  and  $\Delta G_4$ .

The residues with relatively large  $\Delta\Delta G$  values are identified as the functionally-important residues, which are thermodynamically coupled with the binding of ligand to the protein. In order to calculate the free energy changes more easily, a thermodynamic cycle was proposed. Two nonphysical processes were constructed, *i.e.*, Receptor  $\rightarrow$  Receptor\* and Complex  $\rightarrow$  Complex\*. The free energy changes associated with these two constructed processes are denoted as  $\Delta G_3$  and  $\Delta G_4$ . Then, Equation (21) can be expressed as:

$$\Delta\Delta G = \Delta G_2 - \Delta G_1 = \Delta G_4 - \Delta G_3 \quad (22)$$

When a perturbation was introduced to a residue of the protein, the force constant of the corresponding spring is reduced, and the perturbation of  $\gamma_{ij}$  results in the same change of the potential energy of Receptor and Complex, *i.e.*,  $\Delta U_3 = \Delta U_4$ . Then:

$$\Delta\Delta G = (\Delta U_4 - T\Delta S_4) - (\Delta U_3 - T\Delta S_3) = -T(\Delta S_4 - \Delta S_3) \quad (23)$$

In our method, when a perturbation was introduced to a residue, all of the springs connecting to the residue should be perturbed in the same way. In this way, the  $\Delta\Delta G$  value resulting from the perturbation of a residue is calculated by the sum of the  $\Delta\Delta G$  values of all of the perturbed springs involved in this residue. It should be pointed out that only the first 30 slowest normal modes of the closed-holo structure were considered in our calculation, and for the protein-ligand complex, all of the modes were projected onto these slowest modes.

#### 4. Conclusions

The open-closed conformational changes of L-MtAlaDH induced by the coenzyme are important for the catalytic function of the protein. In this work, the domain motions and the functionally-key residues encoded in the structure of the protein were investigated by using the coarse-grained Gaussian network model (GNM) and the anisotropy network model (ANM). The calculated B-factors are consistent with the experimental data from the X-ray crystallographic experiments, which provide the validity for the application of the GNM to the studied system L-MtAlaDH. The analysis of the slowest normal mode for the open-apo and the closed-holo structures shows that the two domains in the proteins undergo large-scale motions, where the motion axes are centered at Met133 and Met301. The ANM calculation results indicate that both for the open-apo and the closed-holo structures, the first three slow modes respectively represent the twist, hinge-bending and rotation motions of the proteins. These motions are responsible for the conformational transition of the proteins and contribute to the opening and closing of the ligand binding pocket. The key residues responsible for the stability of the proteins were successfully identified by analyzing the high-frequency normal modes. It is found that the key residues are tightly packed, and most of them lie in the ligand-binding pocket. The results of cross-correlation analysis implied that the residues in two domains of the protein move in a highly coupled way. In addition, the key residues responsible for the ligand binding were identified by using the thermodynamic cycle method. The predicted functionally-key residues are consistent with the experimental observations. Our method is a simple and coarse-grained model, which considered the connections between C $\alpha$  atoms only. Our methods should be combined with other information to determine the functionally-important residues more exactly and should be combined with more novel methods to calculate the interaction of the side chain. We will construct a more detailed model, in which the side chains of residues are explicitly considered, in the next step to improve our method.

**Acknowledgments:** This work was supported in part by grants from the National Natural Science Foundation of China (No. 11204267), the Natural Science Foundation of Hebei Province (A2014203126) and the Support Program for the Top Young Talents of Hebei Province.

**Author Contributions:** Ji-Guo Su and Jing-Chao Zhang conceived of and designed the study. Xing-Yuan Li wrote the program code and performed the simulations. Xing-Yuan Li and Yan-Ying Zhu contributed to the data analysis and manuscript preparation with valuable discussions. Ji-Guo Su and Jing-Chao Zhang wrote the manuscript. All authors read and approved the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

#### References

1. Bloom, B.R.; McKinney, J.D. The death and resurrection of tuberculosis. *Nat. Med.* **1999**, *5*, 872–874. [[CrossRef](#)] [[PubMed](#)]
2. World Health Organization. *Global Tuberculosis Control: Surveillance and Financing*; WHO Report; WHO: Geneva, Switzerland, 2012.
3. Wayne, L.G.; Sramek, H.A. Metronidazole is bactericidal to dormant cells of *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* **1994**, *38*, 2054–2058. [[CrossRef](#)] [[PubMed](#)]
4. Schnell, R.; Schneider, G. Structural enzymology of sulphur metabolism in *Mycobacterium tuberculosis*. *Biochem. Biophys. Res. Commun.* **2010**, *396*, 33–38. [[CrossRef](#)] [[PubMed](#)]
5. Wayne, L.G.; Hayes, L.G. An *in vitro* model for sequential study of shutdown of *Mycobacterium tuberculosis* through two stages of nonreplicating persistence. *Infect. Immun.* **1996**, *64*, 2062–2069. [[PubMed](#)]
6. Usha, V.; Jayaraman, R.; Toro, J.C.; Hoffner, S.E.; Das, K.S. Glycine and alanine dehydrogenase activities are catalyzed by the same protein in *Mycobacterium smegmatis*: Upregulation of both activities under microaerophilic adaptation. *Can. J. Microbiol.* **2002**, *48*, 7–13. [[CrossRef](#)] [[PubMed](#)]
7. Grimshaw, C.E.; Cleland, W.W. Kinetic mechanism of *Bacillus subtilis* L-alanine dehydrogenase. *Biochemistry* **1981**, *20*, 5650–5655. [[CrossRef](#)] [[PubMed](#)]

8. Tripathi, S.M.; Ramachandran, R. Crystal structures of the *Mycobacterium tuberculosis* secretory antigen alanine dehydrogenase (Rv2780) in apo and ternary complex forms captures “open” and “closed” enzyme conformations. *Proteins* **2008**, *72*, 1089–1095. [[CrossRef](#)] [[PubMed](#)]
9. Agren, D.; Stehr, M.; Berthold, C.L.; Kapoor, S.; Oehlmann, W.; Singh, M.; Schneider, G. Three-dimensional structures of apo- and holo-L-alanine dehydrogenase from *Mycobacterium tuberculosis* reveal conformational changes upon coenzyme binding. *J. Mol. Biol.* **2008**, *377*, 1161–1173. [[CrossRef](#)] [[PubMed](#)]
10. Ling, B.; Sun, M.; Bi, S.; Jing, Z.; Liu, Y. Molecular dynamics simulations of the coenzyme induced conformational changes of *Mycobacterium tuberculosis* L-alanine dehydrogenase. *J. Mol. Graph. Model.* **2012**, *35*, 1–10. [[CrossRef](#)] [[PubMed](#)]
11. Haliloglu, T.; Bahar, I.; Erman, B. Gaussian dynamics of folded proteins. *Phys. Rev. Lett.* **1997**, *79*, 3090–3093. [[CrossRef](#)]
12. Bahar, I.; Atilgan, A.R.; Erman, B. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Fold. Des.* **1997**, *2*, 173–181. [[CrossRef](#)]
13. Atilgan, A.R.; Durell, S.R.; Jernigan, R.L.; Demirel, M.C.; Keskin, O.; Bahar, I. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys. J.* **2001**, *80*, 505–515. [[CrossRef](#)]
14. Kundu, S.; Jernigan, R.L. Molecular mechanism of domain swapping in proteins: An analysis of slower motions. *Biophys. J.* **2004**, *86*, 3846–3854. [[CrossRef](#)] [[PubMed](#)]
15. Jernigan, R.L.; Demirel, M.C.; Bahar, I. Relating structure to function through the dominant slow modes of motion of DNA topoisomerase II. *Int. J. Quantum Chem.* **1999**, *75*, 301–312. [[CrossRef](#)]
16. Wang, Y.; Rader, A.J.; Bahar, I.; Jernigan, R.L. Global ribosome motions revealed with elastic network model. *J. Struct. Biol.* **2004**, *147*, 302–314. [[CrossRef](#)] [[PubMed](#)]
17. Tirion, M.M. Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Phys. Rev. Lett.* **1996**, *77*, 1905–1908. [[CrossRef](#)] [[PubMed](#)]
18. Ma, J. Usefulness and limitations of normal mode analysis in modeling dynamics of biomolecular complexes. *Structure* **2005**, *13*, 373–380. [[CrossRef](#)] [[PubMed](#)]
19. Kim, M.K.; Chirikjian, G.S.; Jernigan, R.L. Elastic models of conformational transitions in macromolecules. *J. Mol. Graph. Model.* **2002**, *21*, 151–160. [[CrossRef](#)]
20. Demirel, M.C.; Atilgan, A.R.; Jernigan, R.L.; Erman, B.; Bahar, I. Identification of kinetically hot residues in proteins. *Protein Sci.* **1998**, *7*, 2522–2532. [[CrossRef](#)] [[PubMed](#)]
21. Haliloglu, T.; Keskin, O.; Ma, B.; Nussinov, R. How similar are protein folding and protein binding nuclei? Examination of vibrational motions of energy hot spots and conserved residues. *Biophys. J.* **2005**, *88*, 1552–1559. [[CrossRef](#)] [[PubMed](#)]
22. Su, J.G.; Xu, X.J.; Li, C.H.; Chen, W.Z.; Wang, C.X. Identification of key residues for protein conformational transition using elastic network model. *J. Chem. Phys.* **2011**, *135*, 174101. [[CrossRef](#)] [[PubMed](#)]
23. Su, J.G.; Du, H.J.; Hao, R.; Xu, X.J.; Li, C.H.; Chen, W.Z.; Wang, C.X. Identification of functionally key residues in AMPA receptor with a thermodynamic method. *J. Phys. Chem. B.* **2013**, *117*, 8689–8696. [[CrossRef](#)] [[PubMed](#)]
24. Kundu, S.; Melton, J.S.; Sorensen, D.C.; Phillips, G.N. Dynamics of proteins in crystals: Comparison of experiment with simple models. *Biophys. J.* **2002**, *83*, 723–732. [[CrossRef](#)]
25. Bahar, I.; Atilgan, A.R.; Demirel, M.C.; Erman, B. Vibrational dynamics of folded proteins: Significance of slow and fast motions in relation to function and stability. *Phys. Rev. Lett.* **1998**, *80*, 2733–2736. [[CrossRef](#)]
26. Humphrey, W.; Dalke, A.; Schulten, K. VMD-visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38. [[CrossRef](#)]
27. Li, X.Y.; Xie, F.; Zhang, J.C.; Su, J.G. Study, by use of coarse-grained models, of the functionally crucial residues and allosteric pathway of anesthetic regulation of the *Gloeobacter violaceus* ligand-gated ion channel. *Eur. Biophys. J.* **2014**, *43*, 623–630. [[CrossRef](#)] [[PubMed](#)]
28. Bahar, I.; Rader, A.J. Coarse-grained normal mode analysis in structural biology. *Curr. Opin. Struct. Biol.* **2005**, *15*, 586–592. [[CrossRef](#)] [[PubMed](#)]

