

Review Article

Comparative Study of Elastic Network Model and Protein Contact Network for Protein Complexes: The Hemoglobin Case

Guang Hu,¹ Luisa Di Paola,² Zhongjie Liang,¹ and Alessandro Giuliani³

¹Center for Systems Biology, School of Electronic and Information Engineering, Soochow University, Suzhou 215006, China

²Unit of Chemical-Physics Fundamentals in Chemical Engineering, Department of Engineering, Università Campus Bio-Medico di Roma, Rome, Italy

³Environment and Health Department, Istituto Superiore di Sanità, Viale Regina Elena 299, 00161 Roma, Italy

Correspondence should be addressed to Guang Hu; huguang@suda.edu.cn

Received 22 June 2016; Revised 17 November 2016; Accepted 20 December 2016; Published 22 January 2017

Academic Editor: Hesham H. Ali

Copyright © 2017 Guang Hu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The overall topology and interfacial interactions play key roles in understanding structural and functional principles of protein complexes. Elastic Network Model (ENM) and Protein Contact Network (PCN) are two widely used methods for high throughput investigation of structures and interactions within protein complexes. In this work, the comparative analysis of ENM and PCN relative to hemoglobin (Hb) was taken as case study. We examine four types of structural and dynamical paradigms, namely, conformational change between different states of Hbs, modular analysis, allosteric mechanisms studies, and interface characterization of an Hb. The comparative study shows that ENM has an advantage in studying dynamical properties and protein-protein interfaces, while PCN is better for describing protein structures quantitatively both from local and from global levels. We suggest that the integration of ENM and PCN would give a potential but powerful tool in structural systems biology.

1. Introduction

Proteins rarely act alone: in the great majority of cases they perform a vast array of biological functions by forming functional complexes [1, 2]. The study of protein complexes not only elucidates the molecular mechanism of many diseases [3] but also provides structural information of protein-protein interactions [4]. With the increasing number of structural data, a lot of regularities have been found for protein complexes based on their topological structures [5]. However, the structural and assembly principles underlying protein complexes organization are not yet fully understood, which poses a great challenge in structural systems biology [6]. A well-studied example of protein complex is hemoglobin (Hb) tetramer, which contains two α and two β subunits as a dimer of dimer [7]. Hbs exist in three quaternary conformations: the low-affinity (deoxy, *T*) state and the high-affinity (oxy, *R*; carbonmonoxy, *R2*) states. Hbs are never present in cells as monomers. Therefore, Hbs were considered as a sort of ‘obliged’ allosteric protein complexes and, even thanks to

the great amount of both structural and physiological data, attracted a lot of attentions [8–10].

Network theory has become a versatile method to study structures and dynamics of biological systems [11–13]. As a dynamical-based method introduced by Tirion [14], Elastic Network Model (ENM) allows performing normal mode analysis at C_α network level. Two mostly used ENM methods, Gaussian Network Model (GNM) and Anisotropic Network Model (ANM), were further proposed by Bahar and coworkers [15, 16]. ENM is an efficient computational tool to describe the essential vibrational dynamics encoded in the molecular topology [17–20]. It has been proved that the low-frequency modes of ENM are critical of collective motions [21], while the high-frequency modes can identify hot spots for protein-protein interactions [22].

The approach of Protein Contact Network (PCN) was proposed by Kannan and Vishveshwara [23] and now has become a new paradigm in protein ontology [24–28]. In a PCN, nodes correspond to C_α , while edges exist if two amino acid residues (nodes) are close to each other under different

cutoffs [29]. Based on this graphical representation, different topological parameters have been developed to describe protein structures and functions from both the global and the local prospective [30–32].

Both ENM and PCN offer computationally efficient tools to study the structure and function of protein complexes [33, 34], from predicting functionally important residues [35, 36], to characterize protein-protein interactions [37, 38] and allosteric communication paths [39, 40]. Of course, both models have strengths and weaknesses and their comparative study is needed.

In this paper, we have analyzed and compared four applications of ENM and PCN on Hb structures: conformational change characterization, modular analysis, allosteric mechanisms investigation, and interface characterization. Although there are several works reported on the ENM [41–43] and PCN [44, 45] studies of Hb independently, this work revisits Hb as case study and mainly focuses on the methodology comparison of ENM (specifically GNM and ANM) and PCN.

2. Materials and Methods

2.1. Data Sets. Hemoglobins (Hbs) have three states [7]. We select their structures for the ENM and PCN analysis, which are listed as follows: *T*-Hb (PDB code: 2dn2), *R*-Hb (PDB code: 2dn1), and *R2*-Hb (PDB code: 2dn3).

2.2. Gaussian Network Model and Anisotropic Network Model. GNM [15] describes a protein as a network of C_α connected by springs of uniform force constant γ if they are located within a cutoff distance r_c (7 Å in this study). In GNM, the interaction potential for a protein of N residues is [46]

$$V_{\text{GNM}} = -\frac{\gamma}{2} \left[\sum_{i=1}^{N-1} \sum_{j=i+1}^N (R_{ij} - R_{ij}^0) \cdot (R_{ij} - R_{ij}^0) \Gamma_{ij} \right], \quad (1)$$

where R_{ij} and R_{ij}^0 are the equilibrium and instantaneous distance between residues i and j , and Γ is $N \times N$ Kirchhoff matrix, which is written as follows:

$$\Gamma_{ij} = \begin{cases} -1 & i \neq j, R_{ij} \leq r_c \\ 0 & i \neq j, R_{ij} > r_c \\ -\sum_{i,i \neq j} \Gamma_{ij} & i = j. \end{cases} \quad (2)$$

Then, square fluctuations are given by

$$\begin{aligned} \langle (\Delta R_i)^2 \rangle &= \left(\frac{3kT}{\gamma} \right) \cdot [\Gamma^{-1}]_{ii}, \\ \langle \Delta R_i \cdot \Delta R_j \rangle &= \left(\frac{3kT}{\gamma} \right) \cdot [\Gamma^{-1}]_{ij}. \end{aligned} \quad (3)$$

The normal modes are extracted by eigenvalue decomposition: $\Gamma = U\Lambda U^T$, where U is the orthogonal matrix whose k th column u_k is k th mode eigenvector. Λ is the diagonal matrix

of eigenvalues, λ_k . $\langle \Delta R_i \cdot \Delta R_j \rangle$ can be written in terms of the sum of the contribution of each mode as follows:

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

Thus, the cross-correlation can be calculated by

$$C_{ij} = \frac{\langle \Delta R_i \cdot \Delta R_j \rangle}{\left[\langle \Delta R_i \rangle^2 \cdot \langle \Delta R_j \rangle^2 \right]^{1/2}}. \quad (5)$$

The cross-correlation value ranges from -1 to 1 : positive values mean that two residues have correlated motions, while the negative values mean that they have anticorrelated motions.

In ANM [16], the interaction potential for a protein of N residues is [46]

$$V_{\text{ANM}} = -\frac{\gamma}{2} \left[\sum_{i=1}^{N-1} \sum_{j=i+1}^N (R_{ij} - R_{ij}^0)^2 \Gamma_{ij} \right]. \quad (6)$$

The motion of the ANM mode of proteins is determined by $3N \times 3N$ Hessian matrix H , whose generic element is given as follows:

$$H_{ij} = \begin{bmatrix} \frac{\partial^2 V}{\partial X_i \partial X_j} & \frac{\partial^2 V}{\partial X_i \partial Y_j} & \frac{\partial^2 V}{\partial X_i \partial Z_j} \\ \frac{\partial^2 V}{\partial Y_i \partial X_j} & \frac{\partial^2 V}{\partial Y_i \partial Y_j} & \frac{\partial^2 V}{\partial Y_i \partial Z_j} \\ \frac{\partial^2 V}{\partial Z_i \partial X_j} & \frac{\partial^2 V}{\partial Z_i \partial Y_j} & \frac{\partial^2 V}{\partial Z_i \partial Z_j} \end{bmatrix}, \quad (7)$$

where X_i , Y_i , and Z_i represent the Cartesian components of residues i and V is the potential energy of the system. r_c used here is 13 Å. Accordingly, ANMs provide the information not only about the amplitudes but also about the direction of residue fluctuations.

The similarity between two ANM modes, u_k and v_l , evaluated for proteins with two different conformations can be quantified in terms of inner product of their eigenvectors [39]; that is,

$$O(u_k, v_l) = u_k \cdot v_l. \quad (8)$$

The degree of overlap between k th ANM modes u_k and the experimentally observed conformation change Δr of Hbs among different states is quantified by $(\Delta r \cdot u_k)/|\Delta r|$. Therefore, the cumulative overlap $\text{CO}(m)$ between Δr and the directions spanned by subsets of m ANM modes is calculated as follows:

$$\text{CO}(m) = \sqrt{\sum_{k=1}^m \left(\Delta r \cdot \frac{u_k}{|\Delta r|} \right)^2}. \quad (9)$$

The Markov model coupled with GNM was used for exploring the signal transductions of perturbations in proteins [47, 48]. The affinity matrix A describes the interactions

between residue pairs connected in GNM; its generic element a_{ij} is defined as follows:

$$a_{ij} = \frac{N_{ij}}{\sqrt{N_i N_j}}, \quad (10)$$

where N_{ij} is the number of atom-atom contacts between residues i and j based on a cutoff distance of 4 Å and N_i is the number of side-chain atoms in residue i . The density of contacts at each node i is given by

$$d_i = \sum_{j=1}^N a_{ij}. \quad (11)$$

The Markov transition matrix M , whose element $m_{ij} = d_j^{-1} a_{ij}$, determines the conditional probability of transmitting a signal from residue j to residue i in one time step. Accordingly, the hitting time for the transfer of a signal from residue j to i is given by [47]

$$H(i, j) = \sum_{k=1}^N \left\{ [\Gamma^{-1}]_{kj} - [\Gamma^{-1}]_{ij} - [\Gamma^{-1}]_{ki} - [\Gamma^{-1}]_{ii} \right\} \cdot d_k, \quad (12)$$

where Γ is Kirchhoff matrix obtained by GNM. The average hit time for i th residue $\langle H(i) \rangle$ is the average of $H(i, j)$ over all starting points i . The commute time is defined by the sum of the hitting times in both directions; that is,

$$C(i, j) = H(i, j) + H(j, i). \quad (13)$$

$C(i, j)$ was defined as the corresponding distance, as the weight of the edge between node i and j in the network.

2.3. Protein Contact Networks (PCNs). Protein Contact Networks (PCNs) provide a coarse-grained representation of protein structure [49], based on $C\alpha$ coordinates from PDB files: network nodes are the residues, while links exist between nodes whose Euclidean distance (computed with respect to α -carbons) is within 4 to 8 Å, in order to account only for significant noncovalent intramolecular interactions [24, 50, 51].

After building up the network, it is possible to quantify its features through the adjacency matrix A_d , whose generic element Ad_{ij} is 1 if i th and j th nodes are connected by a link; otherwise it is 0.

The most basic descriptor is the *node degree*, defined for each node as the number of links involving the node itself:

$$k_i = \sum_j Ad_{ij}. \quad (14)$$

Given a set of vertices V , the *shortest path* $sp_{u,v}$ between two nodes $u, v \in V$ is the minimum number of edges connecting them (Figure 1). Its role is crucial since it has been demonstrated that the lower the network *average shortest path* (or *characteristic length*, computed as the average value over

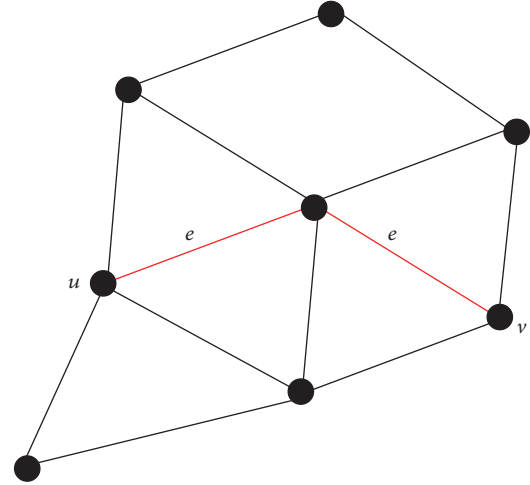


FIGURE 1: Example of a graph with 8 vertices and 13 edges, while the red line shows a path from vertices u to v .

the whole number of node pairs), the higher the efficiency of signal transmission through the network [52]. In PCNs the average shortest path describes the protein attitude to allosteric regulation.

The *betweenness centrality* of a node describes the number of shortest paths passing by it. Given a set of vertices V , the betweenness centrality of node $s \in V$ is defined as follows:

$$\text{betw}(i) = \sum_{v \in V, v \neq i} \sum_{u \in V, u \neq i} \frac{\sigma_{v,u}(i)}{\sigma_{v,u}}, \quad (15)$$

where $\sigma_{v,u}$ is the total number of the shortest paths connecting two nodes $u, v \in V$, whereas $\sigma_{v,u}(i)$ represents the number of shortest paths connecting the nodes u and v passing on i as well. Therefore, high betweenness centrality nodes take part in many shortest paths, so their removal is likely to be noxious for the whole network connectivity. We computed the betweenness centrality by means of the algorithm described in [53].

Closeness centrality describes the general closeness of a node to all other nodes, in terms of length of shortest paths:

$$\text{close}(i) = \frac{1}{\sum_{u \in V, u \neq i} sp(u, i)}. \quad (16)$$

Closeness centrality of residues in PCNs has been demonstrated to describe conformational transitions occurring in protein response to environmental stimuli through cooperative processes [54]: residues in the active site of enzymes show both high degree and closeness centrality; however, it does not provide any clue about allosteric regulation in the enzyme-substrate binding.

The Guimerà-Amaral cartography [55] provides a framework to classify nodes according to their topological role in the network. It is based on network clustering into nodes groups (clusters). We applied a spectral clustering procedure, previously demonstrated to catch functional modules in protein structures [56].

The spectral clustering algorithm [57] applies to the Laplacian matrix L defined as the difference between the adjacency matrix A and the degree matrix D (a diagonal matrix whose generic element D_{ii} is i th node degree). We applied the eigenvalue decomposition to L : the spectral clustering decomposition refers to the eigenvector v_2 corresponding to the second minor eigenvalue.

The procedure applies iteratively to get the final desired number of clusters (set by defining the number of iterations); nodes are parted according to the sign of corresponding v_2 components. So, for instance, if it is required to part the network into four clusters, the first partition produces two clusters, whose v_2 components have opposite signs and, successively, both clusters undergo the same procedure, applied to single cluster nodes subset.

We represented the clustering partition in two ways: first, we reported on the ribbon representation residues pertaining to different clusters in different colors, to identify at once clusters on the three-dimensional structure representation. Second, we reported the clustering color map, a matrix whose generic element is colored not in blue if residues corresponding to indices pertain to the same cluster and in blue, background, if corresponding residues do not belong to the same cluster. This representation helps understanding the distribution of clusters along sequence.

After clustering partition, it is possible to compute for each node (residue) the *participation coefficient* P , defined as follows:

$$P_i = 1 - \left(\frac{k_{si}}{k_i} \right)^2. \quad (17)$$

k_i is the overall degree of the node, k_{si} is the node degree in its own cluster (number of links the node is involved into with nodes pertaining to its own cluster).

A complementary descriptor is the intramodule connectivity z -score z , defined as follows:

$$z_i = \frac{k_{si} - \bar{k}_{si}}{SD_{si}}, \quad (18)$$

where \bar{k} and SD are the average value and the standard deviation of the degree k extended to the whole network. The descriptor z catches the attitude of nodes to preferentially connect with nodes in their own clusters; z strongly correlates with node degree, so high z residues are mostly responsible for global protein stability.

The participation coefficient P has been previously demonstrated of a crucial importance in identifying key residues in protein structure with a functional role [43, 56, 58]; residues with P values higher than 0.75 are mostly devoted to the communication between modules (clusters), since they spend more than half of their links with residues pertaining to clusters other than theirs. In other words, signaling pathways between clusters pass by them.

P - z maps show a peculiar shape (“dentist’s chair”) for PCNs [58]: high P residues show low z values, meaning the role of nodes (communication, high P , and z) are well separated. We previously reported [35] that in protein-ligand

TABLE 1: PCN descriptors and their structural and biological relevance.

| PCN descriptor | Structural and biological role |
|--|---|
| Node degree k | Local stability [24] |
| Betweenness centrality (<i>betw</i>) | Signal transmission throughout the structure [26] |
| Closeness centrality (<i>close</i>) | Residues located in the active site of enzymes [26] |
| Participation Coefficient P | Signal transmission through modules (domains) [27] |
| Intramodule Connectivity z | Intramodule connectivity and communication [27] |

binding P shifts from nonnull to null values for residues close to an active site in allosteric proteins.

We computed for each structure P profile and P - z maps. Then, for the two pairs apo-holo forms we report the heat maps of P variation on the ribbon structure, so to highlight regions in the protein structure undergoing changes upon ligand binding.

The analysis was performed by means of a purposed software implemented in Matlab environment v 2014a, including functions from Bioinformatics Toolbox. Heat maps of P variation (comparison between holo and apo forms), Guimerà-Amaral cartography and clusters onto the protein ribbon representation, have been produced by means of a purposed Python script compiled in the embedded Python environment; for further details and application of the method see [37].

Table 1 sums up PCN descriptors, along with their structural and biological relevance.

3. Results and Discussion

3.1. ENM Results. GNM and ANM are simple yet effective methods [33]. GNM can only describe the amplitude of residue fluctuations, but ANM can give the direction of the motions. In this section, ANM was used to investigate conformational change between T -Hb and R -Hb, and describe the dynamical properties of protein-protein interfaces. GNM was employed for the modular analysis of Hbs, which was coupled with Markovian stochastic analysis to study the allosteric mechanisms of Hbs. Expecting for the conformational change, we only chose T -Hb to exhibit these investigations. ENM results for other two states of Hbs show similar results, as shown in the supporting information.

3.1.1. Conformational Change. ENM results for the transition of tetrameric Hb between T -state (PDB code: 2dn2) and R -state (PDB code: 2dn1) are shown in Figure 2, while the results of the conformational change between T -state and $R2$ -state (PDB code: 2dn3) and R -state and $R2$ -state are shown in Supplementary Figure S1 in Supplementary Material available online at <https://doi.org/10.1155/2017/2483264>. First, the overlap map between the ten ANM slowest modes was calculated to compare the global dynamics of T - and R -Hbs. Along the diagonal in Figure 2(a), only the fifth and sixth modes are

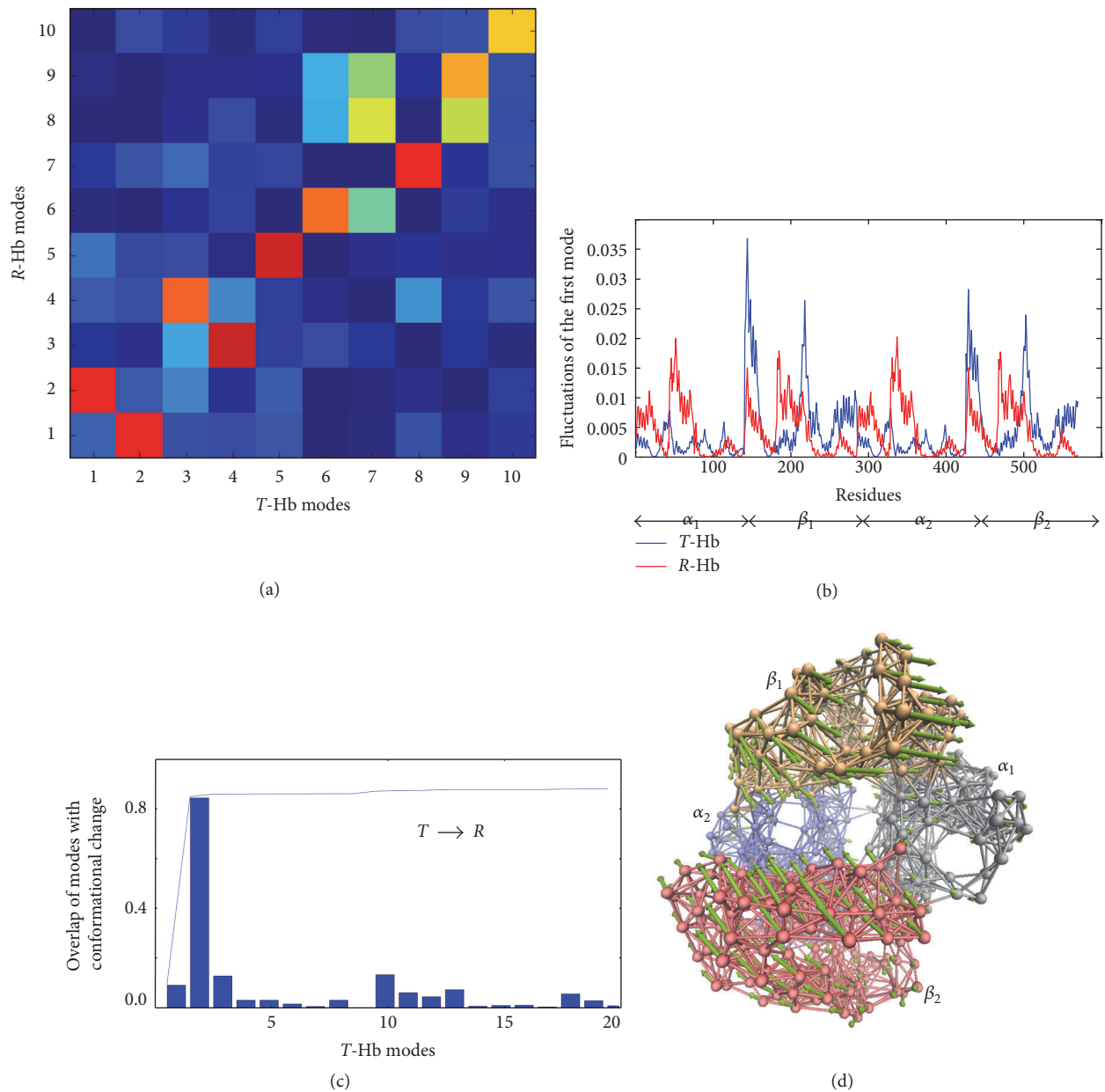


FIGURE 2: ENM results for $T \rightarrow R$ transition of tetrameric Hb. (a) Overlaps between the ten slowest ANM modes of T- and R-Hbs. (b) Distribution of mean-square fluctuations obtained by the first ANM mode of T- and R-Hbs. The residue index of the four chains is 1-140 (α_1), 141-285 (β_1), 286-425 (α_2), and 426-570 (β_2). (c) Overlaps of individual T-Hb ANM modes with the conformational change within $T \rightarrow R$ transition. (d) The motion of the second ANM mode of T-Hb; here the protein is represented as a network.

maintained, with the overlap values of 0.92 and 0.79. For other global modes, there are weaker correlations between two conformations. For example, the reordering of the first two modes was found, which means that the motion of the first mode of T-Hb is similar to the motion of the second mode of R-Hb, while the first mode of R-Hb shifts to the second mode of T-Hb. This result shows that global dynamics greatly changes between the two different states, even for the lowest mode. Then, the difference of two states was further investigated by the distribution of mean-square fluctuations driven by their global ANM modes, as shown in Figure 2(b).

For the first mode T-Hb, the two α chains exhibit different dynamical behavior with two β chains, but two dimers of $\alpha_1\beta_1$ and $\alpha_2\beta_2$ show similar global dynamics (the blue line). For the first mode of R-Hb, the mean-square fluctuations profile of α -chain is very similar to that of the β -chain (the red line). Comparing these two structures the fact that α chains are more stable and β chains are less stable in T state than in R state emerges.

Overlaps of each ANM mode with the structural difference between T and R conformations were calculated to detect which individual mode contributes significantly to the

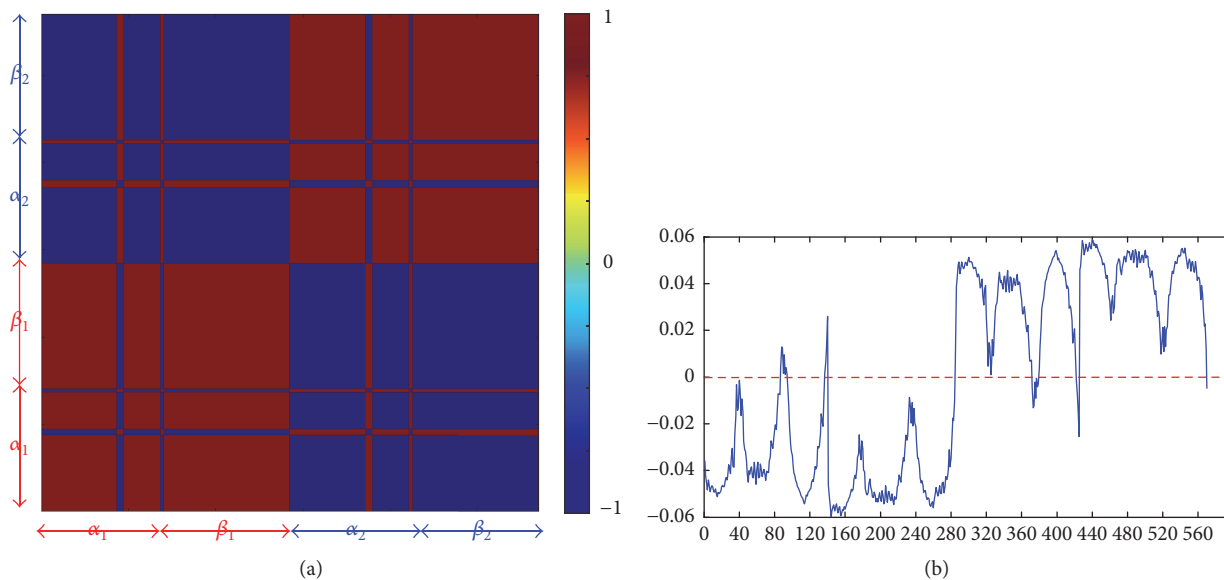


FIGURE 3: Modular analysis of T -Hb based on GNM. (a) The correlation map corresponding to the first mode divides the Hb into two modules. Red regions correspond to collective residue motions and blue-colored regions correspond to uncorrelated motions. (b) The shape of the first mode, which not only shows two modules but also predicts hinge sites.

structural differences between the results from experimental study and are calculated by (9). Figure 2(c) shows that the transformation from T into R is favored by the second mode of T -Hb with the highest overlap (0.84). In this mode, the global motion involves quaternary changes of two dimers, namely, $\alpha_1\beta_1$ dimer, exhibiting a torsional rotation in an opposite direction with $\alpha_2\beta_2$ dimer (Figure 2(d)). Furthermore, this mode is also coordinated by hinge sites at α_1 - β_1 and α_2 - β_2 interface. Tekpinar and Zheng [42] have previously performed the ENM study of conformation changes from T to R_2 structures, in which they found the first two modes contribute significantly to the conformational change. Our revisiting is in accordance with their results, because mode 2 observed herein seems like the combination motion of their two modes.

3.1.2. Modular Analysis. In their recent work, Li et al. [59] developed a new method based on GNM and ANM for dividing a protein into intrinsic dynamics modular analysis. Here, we adopted a much simpler way, just based on the analysis of the GNM lowest mode. Correlation maps for cross-correlation not only describe collective motion but also reflect the symmetry of proteins [36]. To our aim, the correlation map for the first GNM mode was used for the modular analysis of Hb [60]. In the map, red indicates the highly correlated motions, blue represents the anticorrelated motions, and green is for the uncorrelated motions. As shown in Figure 3(a), the correlation map shows that T -Hb tetramer is divided into two modules, which correspond to $\alpha_1\beta_1$ dimer and $\alpha_2\beta_2$ dimer. Two red blocks indicate that α_1 and α_2 move in the same direction with β_1 and β_2 , respectively. Blue blocks indicate that opposite motions are observed between these dimers.

Although the first GNM mode can only generate two modules, it can provide more dynamical information. After diagonalizing the Kirchhoff matrix, the first eigenvector corresponding to the highest eigenvalue can be derived and interpreted to represent the shape of a mode [61]. Figure 3(b) demonstrates that the shape corresponds to GNM mode of the Hb tetramer. It is easy to see that the shape of $\alpha_1\beta_1$ dimer distributes under zero and $\alpha_2\beta_2$ dimer above zero. Thus, the eigenvectors also partition the structure into two modules. In addition, some hinge sites were predicted at near zero positions, which are Thr41, Ala88, and Pro95 in Chain A, His146 in Chain B, Phe98, Leu105, and Ser138 in Chain C, and His2 and His146 in Chain D. Note that these hinge sites always locate at α_1 - β_2 and α_2 - β_1 interfaces. ENM results for modular analysis of R -Hb and R_2 -Hb are shown in Supplementary Figure S2.

3.1.3. Allosteric Mechanisms. Communication inside protein complexes is implicit in collective motions which are inherent to the network topology [62]. Based on this idea, the signal-processing properties of residues can be investigated by Markovian stochastic analysis coupled with GNM [63, 64]. The commute time, $C(i, j)$, a function of Markov transition probabilities, was used to measure the communication abilities of residue pairs. Figure 4(a) displays the commute time map of T -Hb, while the blue and red regions correspond to short and long hit times. Furthermore, we calculated the average values of each row or column of the commute time map to evaluate the communication abilities of each residue. As shown in Figure 4(b), the minima of the average commute time $\langle C(i) \rangle$ indicate the key residues for T -Hb allostery. The profiles of average commute times for α_1 chain and β_1 chain indicate that Val10, Leu29, Arg31, Thr39, Cys104, Val107,

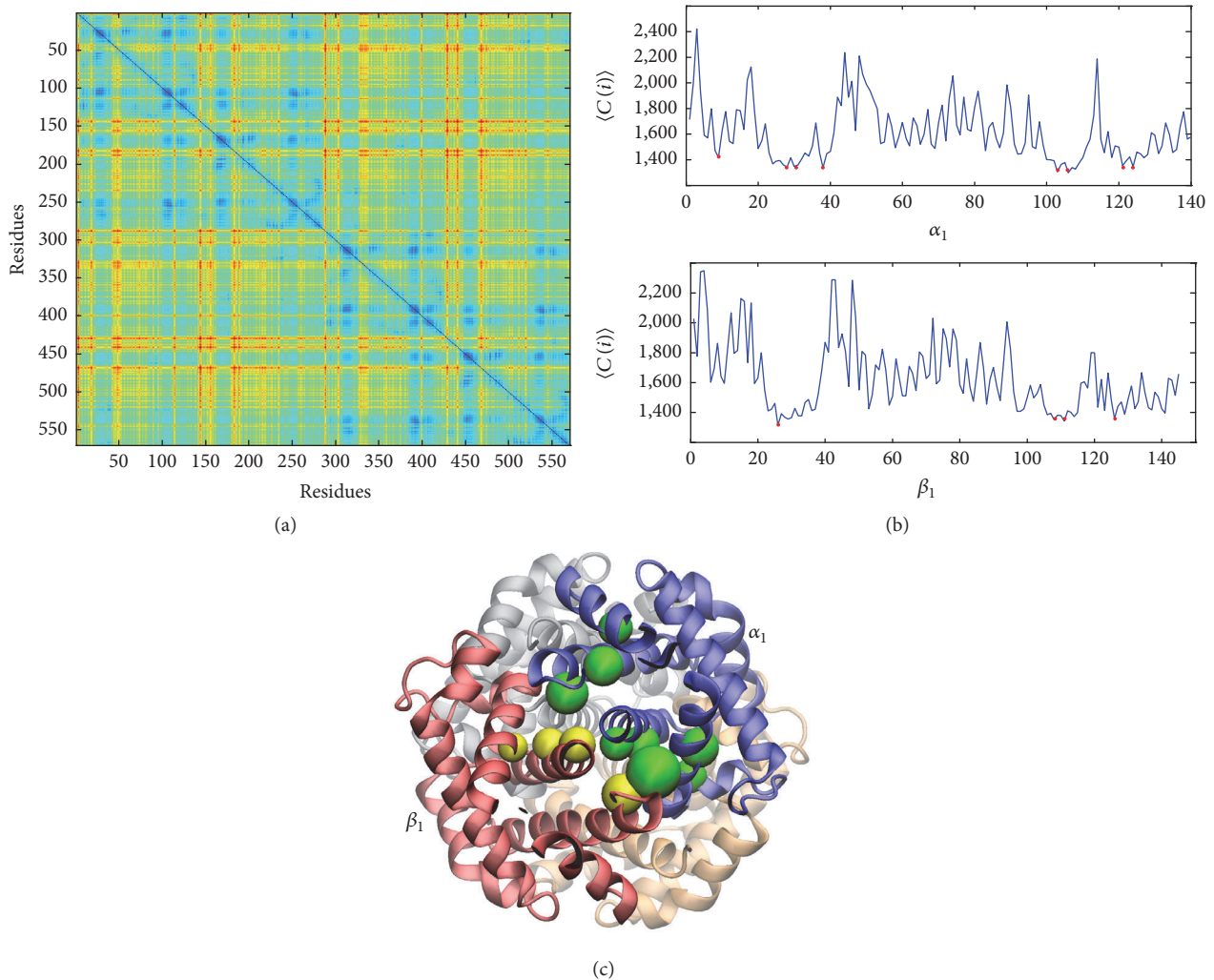


FIGURE 4: Signal propagation of residues for *T*-Hb. (a) The commute time map of *T*-Hb. Minimal of average commute time profiles (red circles) for α_1 chain and β_1 chain (b) indicates that most of residues with highest communication abilities (green beads in α_1 chain and yellow beads in β_1 chain) distribute at α_1 - β_1 interface (c). Results are presented for 2dn1 (*T*), and 2dn2 (*R*) and 2dn3 (*R*₂) showed similar behavior.

His122, and Leu125 in α_1 chain and Ala27, Val109, Cys112, and Gln127 in β_1 chain are residues with highest communication abilities. It is worth mentioning that the two α chains and two β chains have the same profile shapes. The distributions of these residues in α_1 chain and β_1 chain are also displayed in the three-dimensional representation (Figure 4(c)). It was found that Arg31, Cys104, Val107, and His122 in α_1 chain and Cys112 and Gln127 in β_1 chain are located at α_1 - β_1 interface. Likely, the same region was also found at α_2 - β_2 interface. ENM results for modular analysis of *R*-Hb and *R*₂-Hb are shown in Supplementary Figure S3.

3.1.4. Interface Characterization. Protein interfaces are the sites where proteins or subunits physically interact. Identification and characterization of protein interfaces are not only important to understand the structures of protein complexes and protein-protein interactions, but also disease phenotypes [65]. Both GNM and ANM have been used to investigate

protein-protein interfaces. Kantarci et al. [66] firstly applied GNM to classify interfaces of p53 core domain into the dimerization interface and crystal interface on the base of interfacial dynamics. Zen et al. [67] extended this method to study the interface of 22 representative dimers. More recently, Soner et al. [68] developed a web server to discriminate obligatory and nonobligatory protein complexes. Although GNM is the most used method to study protein-protein interfaces, we have showed here that ANM is also powerful to explore interfacial dynamics of Hbs.

Two kinds of interfaces have been classified in the Hb tetramer: allosteric sites located at α_1 - β_1 and α_2 - β_2 interfaces, which could be intended as allosteric interfaces. Hinge sites are detected always at α_1 - β_2 and α_2 - β_1 interfaces, providing structural interfaces. The analyses are in accordance with the results in Tekpinar and Zheng [42], which showed that allosteric interfaces are dynamically variable regions but not necessarily structural interfaces. In this section, square

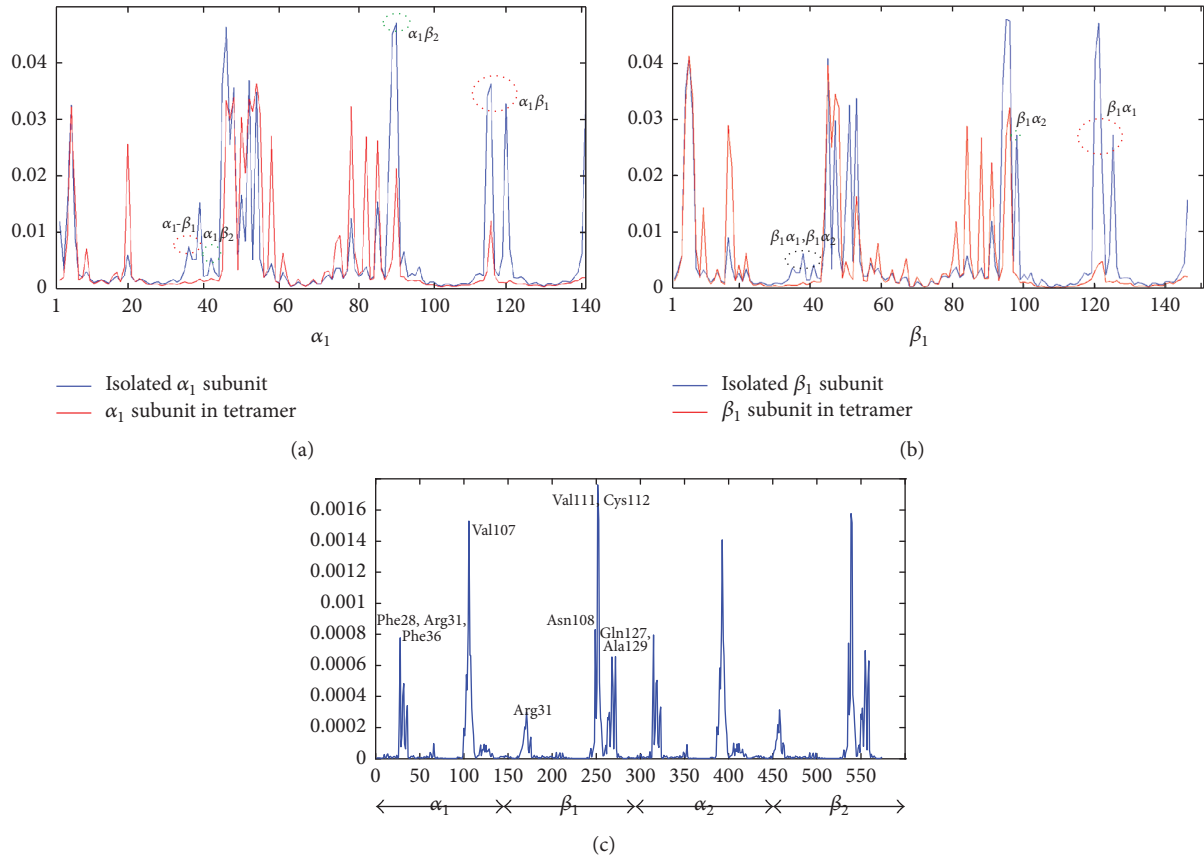


FIGURE 5: Square fluctuations of Hb monomers and tetramers. (a) Square fluctuations of α_1 subunits in isolated and tetrameric states based on the first 20 ANM modes. The differences of mobilities at $\alpha_1\text{-}\beta_1$ and $\alpha_1\text{-}\beta_2$ interfaces are indicated by red and green circles. (b) Square fluctuations of β_1 subunits in isolated and tetrameric states based on the first 20 ANM modes. Red, green, and black circles indicate the differences of mobilities at $\alpha_1\text{-}\beta_1$ and $\alpha_2\text{-}\beta_1$ interfaces and a common region of these two interfaces. (c) Square fluctuations of T -Hb tetramers based on the highest two modes. Hot spots are predicted by the peaks in the profile, while $\alpha_1\beta_1$ and $\alpha_2\beta_2$ dimer show the same prediction result.

fluctuations of both monomeric and oligomeric proteins based on a large set of slow modes and the highest modes are compared for a deeper analysis of interfaces.

Figures 5(a) and 5(b) show square fluctuations of α_1 and β_1 subunits in isolated and tetrameric states based on the first 20 ANM modes, while α_2 and β_2 subunits show similar behavior. Although α and β subunits are structurally identical, they are different in length and sequence. Accordingly, α_1 and β_1 subunits show different types of fluctuations, which have also been predicted by the previous molecular dynamics (MD) study [69]. The mobility of $\alpha_1\text{-}\beta_1$ and $\alpha_1\text{-}\beta_2$ interfacial residues of α_1 subunit is reduced in the tetramer. The same happens for the mobilities of $\alpha_1\text{-}\beta_1$ and $\alpha_2\text{-}\beta_1$ interfacial residues of β_1 subunit. Therefore, the flexibilities of residues located at both kinds of interfaces in bound states are lower than in unbound monomers. This kind of dynamical property of interfacial residues has also been detected by the MD simulation [70].

In addition, a similar region (residues 45–57) with high mobility in both isolated states was found, which corresponds to a long loop distributing between two adjacent subunits. The mobility of this region in β_1 subunit was reduced, while

no reduction was observed in α_1 subunit. This may suggest that this long loop in β subunits is an allosteric region controlled by interfacial residues.

Among the interface residues, hotspots are defined as residues that have the greatest contribution to the binding energy. The prediction of hotspots is helpful not only to guide drug design but also to understand disease mutations [71]. Based on ENM results, Chennubhotla et al. [72] revealed that hot spots residues show a moderate-high flexibility at global modes. On the other hand, hot spots correlated very well with the residues with high mean-square fluctuations in the highest frequency modes in both GNM [73, 74] and ANM [22]. Ozbek et al. [74] have found that hot spots predictions based on the highest, the second and third highest, and the average three and five highest GNM modes show similar accuracies. Our calculation demonstrates that the square fluctuation based on two highest ANM modes is enough to predict the distribution of of Hb-tetramer hot spots. The result for T -Hb is shown in Figure 5(c). It is surprising that hot spots have been predicted only at $\alpha_1\text{-}\beta_1$ and $\alpha_2\text{-}\beta_2$ interfaces: Phe28, Arg31, Phe36, and Val107 in α subunits and Arg31, Asn108, Val111, Cys112, Gln127, and Ala129 in

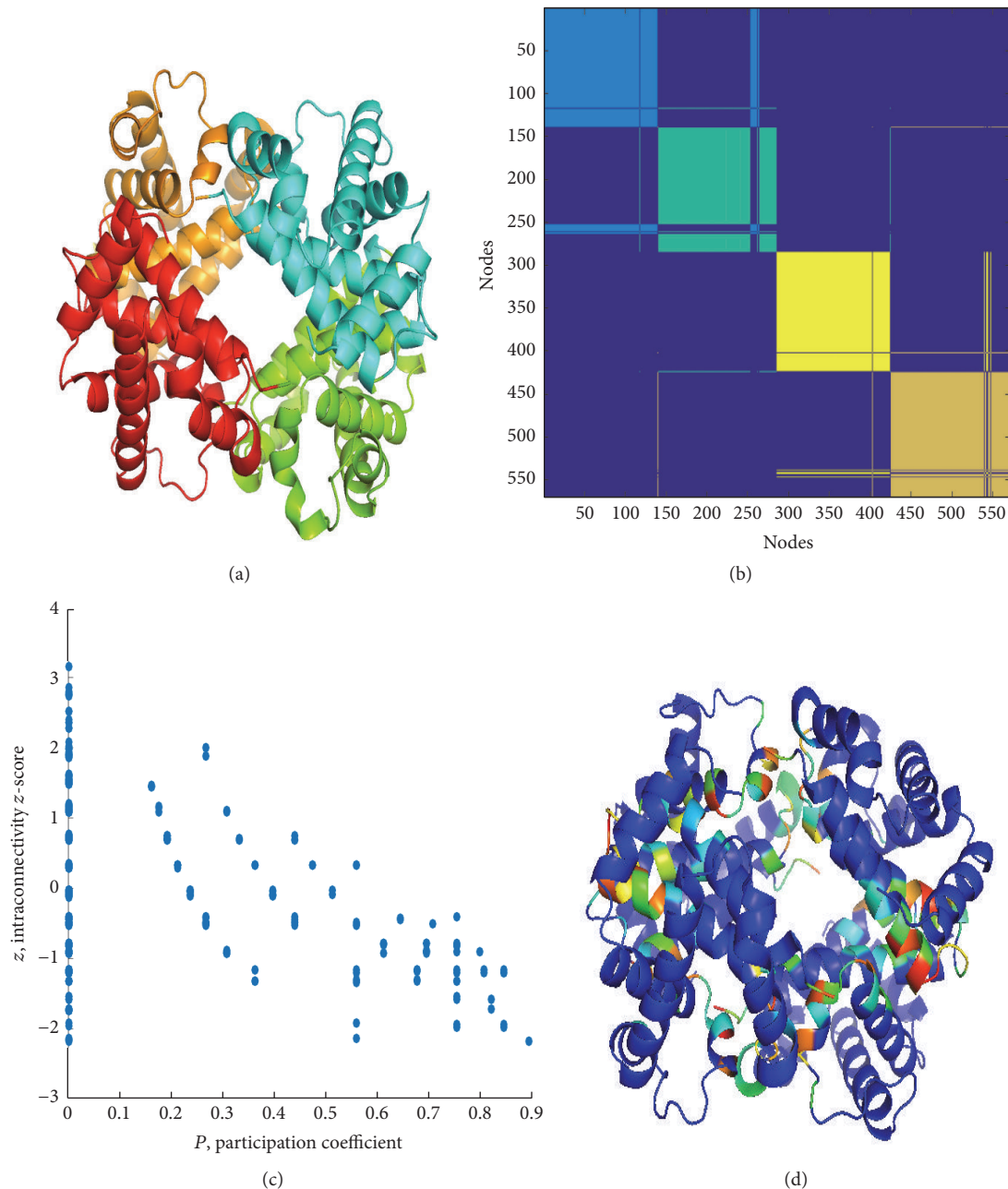


FIGURE 6: PCN results for *T*-Hb: (a) cluster partition, (b) cluster color map, (c) P - z map, and (d) P over ribbon. Cluster partition satisfactorily matches with chains, and high P residues are mostly located at the chain interfaces.

β subunits. Note that Arg31 and Val107 in α subunits and Arg31, Cys112, and Gln127 in β subunits are overlapped with allosteric sites. It also proved that allosteric interfaces rather than structural interfaces take part in the complex formation. The hotspots predicted for *R*- and *R2*-Hbs show small differences but still located at the same interfaces (Supplementary Figure S4).

3.2. PCN Results. In this section, results from the application of PCN method and spectral clustering are reported for the three structures under enquiry. Figures 6 and 7 and Supplementary Figure S5 clearly show that cluster partition

satisfactorily matches with chains, yet with some divergences (region pertaining to a chain falling in a cluster mainly composed of residues belonging to a second chain, “whiskers” in the clustering color map). In comparison with ENM, PCN results of three states of Hbs exhibit quite high similarity, even between *R*- and *R2*-states, as emerged mainly from the distribution of P along the ribbon structures (see Figures 8 and 9).

Pz maps show the typical profile for PCNs (and not for other real world networks), with most residues having $P = 0$ (only contacts with residues belonging to their own clusters). Residues with $P > 0$ are mostly interesting for protein functionality, since they account for signaling transmission

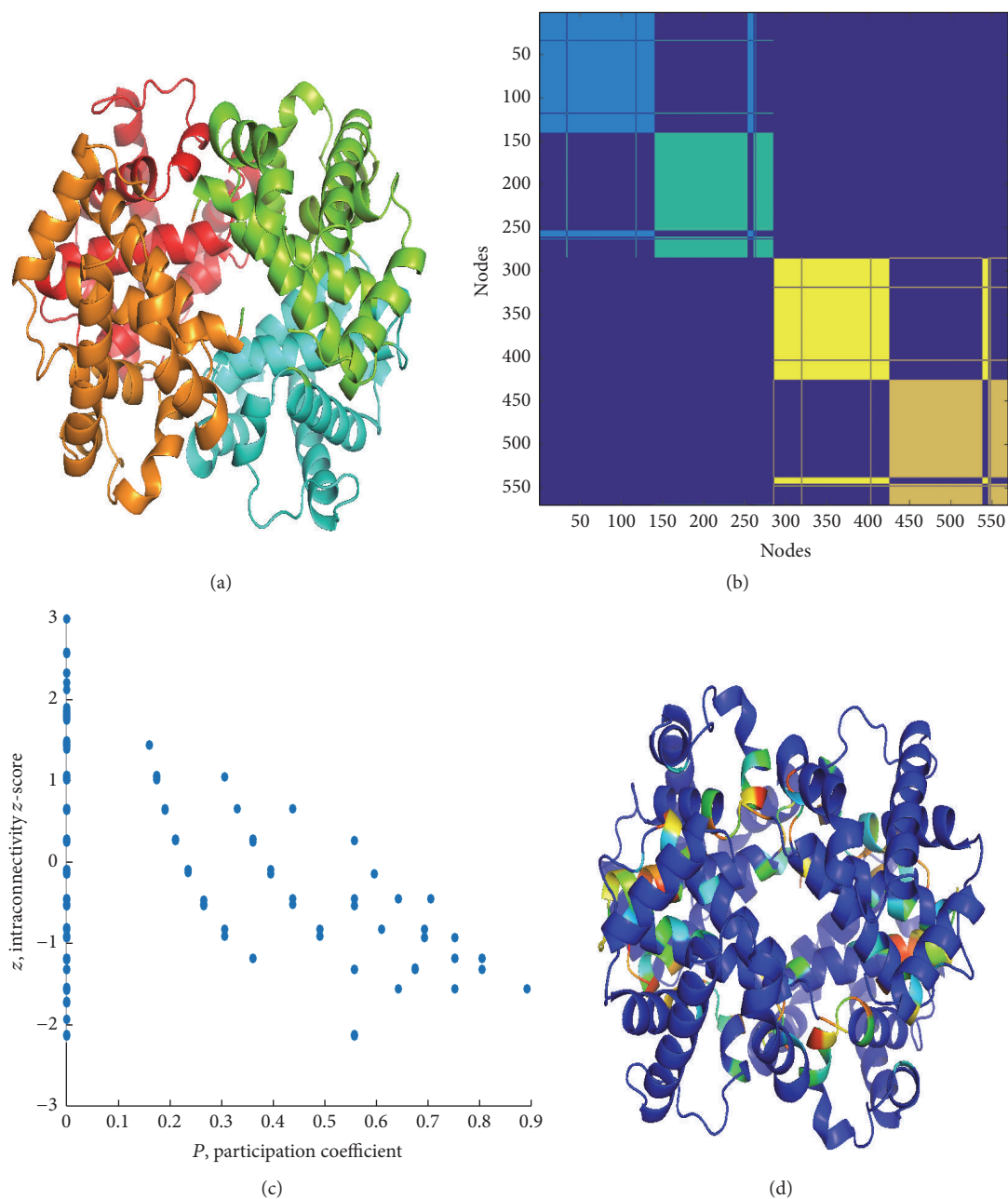


FIGURE 7: PCN results for *R*-Hb: (a) cluster partition, (b) cluster color map, (c) P - z map, and (d) P over ribbon. Again, clusters catch almost perfectly the chains and high P residues are located at the interfaces between chains. Similar results are obtained for *R2*-Hb; see Supplementary Figure S5.

through the protein structure (global property of protein structure).

High P residues are spotted in the structure and mostly (but not necessarily) placed in the interchain region. In previous works [35, 37, 51, 75], we demonstrated that the participation coefficient P addresses the functional role of residues in protein binding and, in general, identifies residues with a key role in protein structural and functional features.

Pz maps instruct a cartography, addressing a specific role to residues, as reported in Table 2. Hubs are nodes with $z > 2.5$, while P values address the role of nodes to connect

different clusters. The Guimerà-Amara cartography of the three Hbs is reported in Figures 8 and 9 and Supplementary Figure S6, as original form, on Pz maps, and projection on ribbon structures.

Noticeably, in PCNs $R6$ and $R7$ nodes are absent and only few $R5$ nodes are present, all at $P = 0$. In other words, high z nodes correspond to residues in charge for protein stability, while nonhub connector nodes are responsible for interdomain (intercluster) communication. Lys60 in α_1 , Glu26, His63, Lys66 in β_1 , and Leu28, Lys65, Leu68 in β_2 in *T*-Hb, whereas Gly24, Lys61, and Leu141 in two β chains in

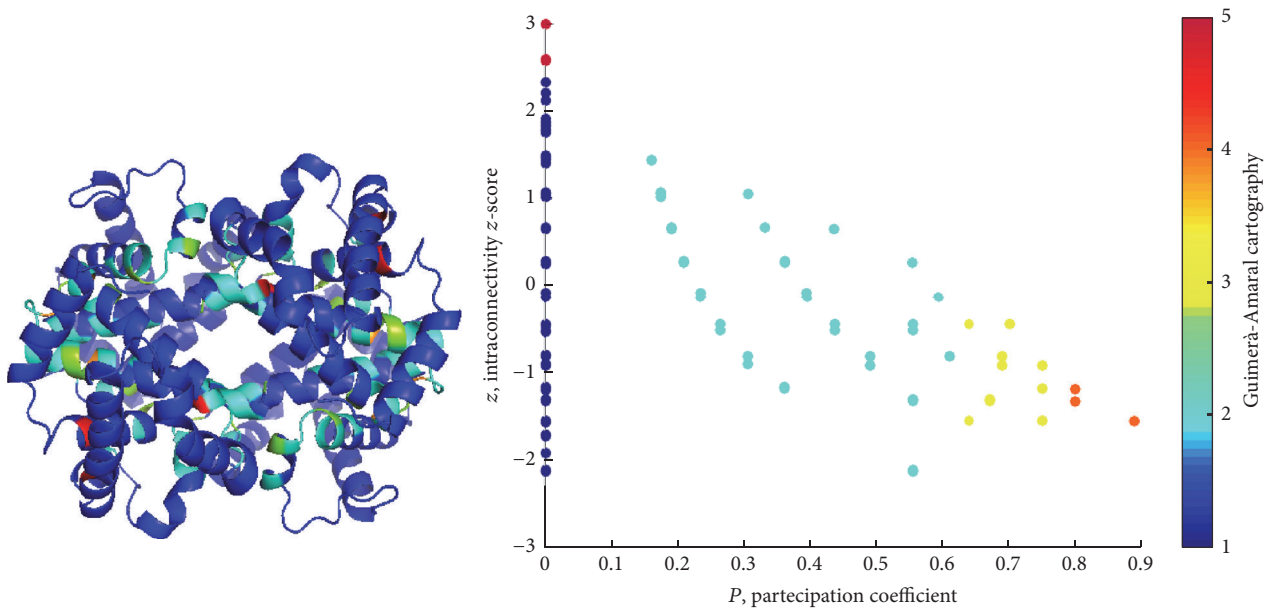


FIGURE 8: Guimerà-Amaral Cartography for *R*-Hb: only very few nodes are classified as hubs (*R5*) and located close to the active site. Non-hub kinless nodes (*R4*) are located in turn on the interface between chains and play a key role in the concerted motion underlying the allosteric regulation of hemoglobin.

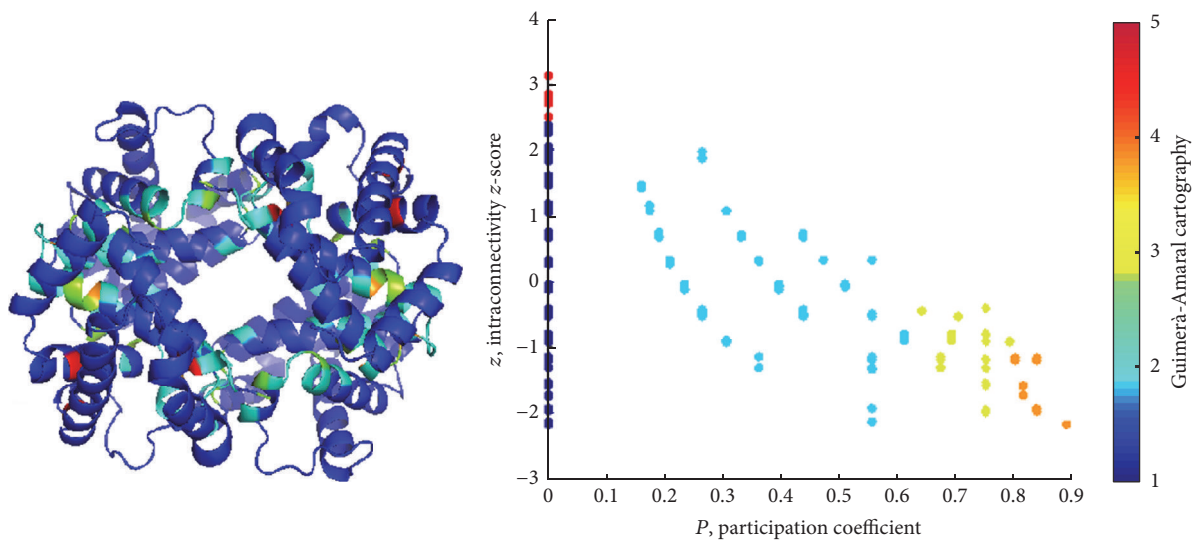


FIGURE 9: Guimerà-Amaral Cartography for *T*-Hb: very few nodes are classified as hubs (*R5*), but more than for *R/O*₂ complex, again close to the active site. Non-hub kinless nodes (*R4*) lie on the interface between chains. Similar results hold for *R2*-Hb; see Supplementary Figure S6.

TABLE 2: Guimerà-Amaral cartography.

| | Regions | <i>z</i> | <i>P</i> |
|-----------------------|----------------------------------|-----------|--------------------|
| <i>Module nonhubs</i> | <i>R1</i> : ultraperipheral node | $z < 2.5$ | $P < 0.05$ |
| | <i>R2</i> : peripheral node | $z < 2.5$ | $0.05 < P < 0.625$ |
| | <i>R3</i> : nonhub connectors | $z < 2.5$ | $0.625 < P < 0.8$ |
| | <i>R4</i> : nonhub kinless nodes | $z < 2.5$ | $P > 0.8$ |
| <i>Module hubs</i> | <i>R5</i> : provincial hubs | $z > 2.5$ | $P < 0.3$ |
| | <i>R6</i> : connector hubs | $z > 2.5$ | $0.3 < P < 0.75$ |
| | <i>R7</i> : kinless hubs | $z > 2.5$ | $P > 0.75$ |

TABLE 3: Pearson correlation coefficients of network descriptors with mean fluctuations (MF) and average commute times (CT).

| | Closeness/MF | Betweenness/MF | P/MF | P/CT |
|------|--------------|----------------|---------|---------|
| 2DN2 | -0.3741 | -0.1828 | -0.0861 | -0.2677 |
| 2DN1 | -0.4320 | -0.1226 | -0.1878 | -0.2904 |
| 2DN3 | -0.7331 | -0.1419 | -0.3649 | -0.2740 |

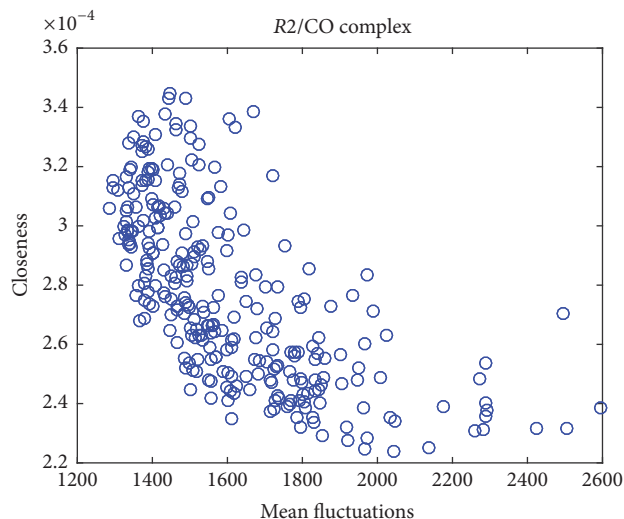


FIGURE 10: Closeness versus mean fluctuations: in this complex, there is an hyperbolic variation of closeness along with mean fluctuations: the closeness is thus a local rigidity descriptor.

R-Hb belong to R5 nodes. It is easy to note that R5 nodes distribute at β chains within the protein interior.

As previously stated [58], R4 nodes (nonhub kinless nodes) are crucial for the allosteric signal propagation: their kinless nature poses in the gray zone where residues acting at a global level play, so their role in the protein functionality is central. It was found that Leu91 and Arg92 in α_1 , His2, His116, and Ala129 in β_1 , His89 in α_2 , and Thr38, His116, and Phe118 in T-Hb and Trp37, Cys112, Phe122, and Pro124 in R-Hb belong to R5 nodes. Except His2, R4 residues were found at all four interfacial regions.

3.3. Comparison between ENM and PCN Results. We finally explicitly superpose the ENMs and PCNs results, in order to better specify key residues and features in allosteric regulation of Hb. Average commute times predict allosteric sites at both protein interior and two interfaces (α_1 - β_1 and α_2 - β_2 interfaces). In PCN, R4 nodes include allosteric sites at interfaces and R5 nodes include allosteric sites at protein interior. Combined with modular analysis and hot spots prediction, the use of ENM has advantage to classify protein-protein interfaces.

On the other hand, PCNs analysis relies upon a set of network descriptors to approach the study of protein structures quantitatively. Table 3 reports the Pearson correlation coefficients between mean fluctuations and network descriptors, closeness centrality, betweenness centrality, and participation coefficient.

Betweenness centrality poorly correlates (negatively) with mean fluctuations, while closeness anticorrelates more strongly with mean fluctuations, especially in the more rigid structure (R2/complex, Figure 10). The hyperbolic shape of the distribution confirms closeness is a general stiffness descriptor for protein structure. This property may indicate that closeness in PCN could provide an additional evidence to detect hinge sites. There is a relatively poor one-to-one correspondence of functional sites obtained between ENM and PCN, and thus the combination of these two approaches would improve the prediction.

4. Conclusion and Perspectives

ENM and PCN are light yet effective computational methods which simply require the three-dimensional coordinates of atoms in protein structures. In this work, the combination of the ENM and PCN methodologies has provided a plenty of information regarding the dynamic behavior of Hbs. It is noteworthy that the two classes of methods are able to catch the same features without a common, interexchange ground. In comparison with PCN, ENM can find the dominate motion for the conformational change of proteins and detect the dynamics of protein-protein interfaces observed by MD. Except for the topological parameters used in our work, there are more local and global network parameters that can be calculated in PCN to describe protein structures quantitatively. For example, residue centrality as a local network parameter was proposed to identify allosteric sites [76], and coefficient of assortativity as a global network parameter is related to the rates of protein folding [77]. In addition, we have found some correlations between ENM and PCN results. In previous studies [78], the average path lengths are highly correlated with residue fluctuations. Here, we show an additional positive correlation between residue fluctuations predicted by ENM and closeness centrality calculated by PCN. Although the general relationship between dynamical properties and more network parameters is needed to be established, we can conclude that ANM and GNM have advantages in studying dynamical properties and protein-protein interfaces, while PCN is better for describing structures quantitatively from both local and global levels.

In future, the combined description by means of these methods will largely contribute to understanding the dynamic behavior of complexes without heavy computational approaches, such as molecular dynamics (MD). Evidently, MD will anyway provide a very complete and fine description of dynamics, but the combination of lighter methods, such as ENM and PCN, will, for instance, guide MD simulations with well-grounded preliminary results, as preliminary approached in our previous works [79]. On the

other hand, the two methods may help understanding the relationship between local fluctuation of residues and protein stability and functionality, being a primer for identifying key residues, responsible for lethal mutations. For example, the first attempt to combine ENM and PCN has been reported to investigate allosteric communication pathways [80]. In our work, we only indicate that ENM and PCN can be applied to four types of structural and dynamical paradigms. More detailed analysis for each case is needed. Although the integration of these two methods is just at the beginning, it would give a potential but powerful tool in structural systems biology.

Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Authors' Contributions

Guang Hu and Luisa Di Paola are equal contributors.

Acknowledgments

This work was supported by the National Nature Science Foundation of China (21203131 and 81302700), the Natural Science Foundation of the Jiangsu Higher Education Institutions (12KJB180014 and 13KJB520022), and the China Postdoctoral Science Foundation (2013M531406 and 2016M590495). Guang Hu also thanks Professor Chakra Chennubhotla for providing the program of calculating hitting time and commute time.

References

- [1] B. J. G. E. Pieters, M. B. Van Eldijk, R. J. M. Nolte, and J. Mecnović, "Natural supramolecular protein assemblies," *Chemical Society Reviews*, vol. 45, no. 1, pp. 24–39, 2016.
- [2] T. Perica, J. A. Marsh, F. L. Sousa et al., "The emergence of protein complexes: quaternary structure, dynamics and allostery," *Biochemical Society Transactions*, vol. 40, no. 3, pp. 475–491, 2012.
- [3] T. L. Nero, C. J. Morton, J. K. Holien, J. Wielens, and M. W. Parker, "Oncogenic protein interfaces: small molecules, big challenges," *Nature Reviews Cancer*, vol. 14, no. 4, pp. 248–262, 2014.
- [4] O. Keskin, N. Tuncbag, and A. Gursoy, "Predicting protein-protein interactions from the molecular to the proteome level," *Chemical Reviews*, vol. 116, no. 8, pp. 4884–4909, 2016.
- [5] S. E. Ahnert, J. A. Marsh, H. Hernández, C. V. Robinson, and S. A. Teichmann, "Principles of assembly reveal a periodic table of protein complexes," *Science*, vol. 350, no. 6266, article no. 1331, 2015.
- [6] P. Aloy and R. B. Russell, "Structural systems biology: modelling protein interactions," *Nature Reviews Molecular Cell Biology*, vol. 7, no. 3, pp. 188–197, 2006.
- [7] S.-Y. Park, T. Yokoyama, N. Shibayama, Y. Shiro, and J. R. H. Tame, "1.25 Å resolution crystal structures of human haemoglobin in the oxy, deoxy and carbonmonoxy forms," *Journal of Molecular Biology*, vol. 360, no. 3, pp. 690–701, 2006.
- [8] M. D. Vesper and B. L. de Groot, "Collective dynamics underlying allosteric transitions in hemoglobin," *PLoS Computational Biology*, vol. 9, no. 9, 2013.
- [9] Y. Yuan, M. F. Tam, V. Simplaceanu, and C. Ho, "New look at hemoglobin allostery," *Chemical Reviews*, vol. 115, no. 4, pp. 1702–1724, 2015.
- [10] J. S. Hub, M. B. Kubitzki, and B. L. de Groot, "Spontaneous quaternary and tertiary T-R transitions of human hemoglobin in molecular dynamics simulation," *PLoS Computational Biology*, vol. 6, no. 5, pp. 1–11, 2010.
- [11] A.-L. Barabási and Z. N. Oltvai, "Network biology: understanding the cell's functional organization," *Nature Reviews Genetics*, vol. 5, pp. 101–113, 2004.
- [12] P. Csermely, "Creative elements: network-based predictions of active centres in proteins and cellular and social networks," *Trends in Biochemical Sciences*, vol. 33, no. 12, pp. 569–576, 2008.
- [13] C. Böde, I. A. Kovács, M. S. Szalay, R. Palotai, T. Korcsmáros, and P. Csermely, "Network analysis of protein dynamics," *FEBS Letters*, vol. 581, no. 15, pp. 2776–2782, 2007.
- [14] M. M. Tirion, "Large amplitude elastic motions in proteins from a single-parameter, atomic analysis," *Physical Review Letters*, vol. 77, no. 9, pp. 1905–1908, 1996.
- [15] I. Bahar, A. R. Atilgan, and B. Erman, "Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential," *Folding and Design*, vol. 2, no. 3, pp. 173–181, 1997.
- [16] A. R. Atilgan, S. R. Durell, R. L. Jernigan, M. C. Demirel, O. Keskin, and I. Bahar, "Anisotropy of fluctuation dynamics of proteins with an elastic network model," *Biophysical Journal*, vol. 80, no. 1, pp. 505–515, 2001.
- [17] G. Hu, S. Michielssens, S. L. C. Moors, and A. Ceulemans, "The harmonic analysis of cylindrically symmetric proteins: a comparison of Dronpa and a DNA sliding clamp," *Journal of Molecular Graphics and Modelling*, vol. 34, pp. 28–37, 2012.
- [18] S. P. Tiwari and N. Reuter, "Similarity in shape dictates signature intrinsic dynamics despite no functional conservation in TIM barrel enzymes," *PLoS Computational Biology*, vol. 12, no. 3, Article ID e1004834, 2016.
- [19] E. Fuglebakk, S. P. Tiwari, and N. Reuter, "Comparing the intrinsic dynamics of multiple protein structures using elastic network models," *Biochimica et Biophysica Acta (BBA)—General Subjects*, vol. 1850, no. 5, pp. 911–922, 2015.
- [20] X.-Y. Li, J.-C. Zhang, Y.-Y. Zhu, and J.-G. Su, "Domain motions and functionally-key residues of l-alanine dehydrogenase revealed by an elastic network model," *International Journal of Molecular Sciences*, vol. 16, no. 12, pp. 29383–29397, 2015.
- [21] S. Mahajan and Y.-H. Sanejouand, "On the relationship between low-frequency normal modes and the large-scale conformational changes of proteins," *Archives of Biochemistry and Biophysics*, vol. 567, pp. 59–65, 2015.
- [22] A. Uyar, O. Kurkcuoglu, L. Nilsson, and P. Doruker, "The elastic network model reveals a consistent picture on intrinsic functional dynamics of type II restriction endonucleases," *Physical Biology*, vol. 8, no. 5, Article ID 056001, 2011.
- [23] N. Kannan and S. Vishveshwara, "Identification of side-chain clusters in protein structures by a graph spectral method," *Journal of Molecular Biology*, vol. 292, no. 2, pp. 441–464, 1999.
- [24] L. Di Paola, M. De Ruvo, P. Paci, D. Santoni, and A. Giuliani, "Protein contact networks: an emerging paradigm in chemistry," *Chemical Reviews*, vol. 113, no. 3, pp. 1598–1613, 2013.

- [25] G. Hu, J. Zhou, W. Yan, J. Chen, and B. Shen, "The topology and dynamics of protein complexes: insights from intra-molecular network theory," *Current Protein and Peptide Science*, vol. 14, no. 2, pp. 121–132, 2013.
- [26] W. Yan, J. Zhou, M. Sun, J. Chen, G. Hu, and B. Shen, "The construction of an amino acid network for understanding protein structure and function," *Amino Acids*, vol. 46, no. 6, pp. 1419–1439, 2014.
- [27] L. Di Paola and A. Giuliani, "Protein contact network topology: a natural language for allostery," *Current Opinion in Structural Biology*, vol. 31, pp. 43–48, 2015.
- [28] L. H. Greene, "Protein structure networks," *Briefings in Functional Genomics*, vol. 11, no. 6, pp. 469–478, 2012.
- [29] S. Cheng, H.-L. Fu, and D.-X. Cui, "Characteristics analyses and comparisons of the protein structure networks constructed by different methods," *Interdisciplinary Sciences: Computational Life Sciences*, vol. 8, no. 1, pp. 65–74, 2016.
- [30] N. T. Doncheva, K. Klein, F. S. Domingues, and M. Albrecht, "Analyzing and visualizing residue networks of protein structures," *Trends in Biochemical Sciences*, vol. 36, no. 4, pp. 179–182, 2011.
- [31] R. K. Grewal and S. Roy, "Modeling proteins as residue interaction networks," *Protein and Peptide Letters*, vol. 22, no. 10, pp. 923–933, 2015.
- [32] L. Vuillon and C. Lesieur, "From local to global changes in proteins: a network view," *Current Opinion in Structural Biology*, vol. 31, pp. 1–8, 2015.
- [33] I. Bahar, T. R. Lezon, L.-W. Yang, and E. Eyal, "Global dynamics of proteins: bridging between structure and function," *Annual Review of Biophysics*, vol. 39, no. 1, pp. 23–42, 2010.
- [34] X. Zhang, T. Perica, and S. A. Teichmann, "Evolution of protein structures and interactions from the perspective of residue contact networks," *Current Opinion in Structural Biology*, vol. 23, no. 6, pp. 954–963, 2013.
- [35] M. De Ruvo, A. Giuliani, P. Paci, D. Santoni, and L. Di Paola, "Shedding light on protein-ligand binding by graph theory: the topological nature of allostery," *Biophysical Chemistry*, vol. 165–166, pp. 21–29, 2012.
- [36] G. Hu, S. Michielssens, S. L. C. Moors, and A. Ceulemans, "Normal mode analysis of Trp RNA binding attenuation protein: structure and collective motions," *Journal of Chemical Information and Modeling*, vol. 51, no. 9, pp. 2361–2371, 2011.
- [37] L. Di Paola, C. B. M. Platania, G. Oliva, R. Setola, F. Pascucci, and A. Giuliani, "Characterization of protein-protein interfaces through a protein contact network approach," *Frontiers in Bioengineering and Biotechnology*, vol. 3, article 170, 2015.
- [38] W. I. Karain and N. I. Qaraeen, "Weighted protein residue networks based on joint recurrences between residues," *BMC Bioinformatics*, vol. 16, article 173, 2015.
- [39] E. Marcos, R. Crehuet, and I. Bahar, "Changes in dynamics upon oligomerization regulate substrate binding and allostery in amino acid kinase family members," *PLoS Computational Biology*, vol. 7, no. 9, Article ID e1002201, 2011.
- [40] O. Schueler-Furman and S. J. Wodak, "Computational approaches to investigating allostery," *Current Opinion in Structural Biology*, vol. 41, pp. 159–171, 2016.
- [41] C. Xu, D. Tobi, and I. Bahar, "Allosteric changes in protein structure computed by a simple mechanical model: hemoglobin T \leftrightarrow R2 transition," *Journal of Molecular Biology*, vol. 333, no. 1, pp. 153–168, 2003.
- [42] M. Tekpinar and W. Zheng, "Coarse-grained and all-atom modeling of structural states and transitions in hemoglobin," *Proteins: Structure, Function and Bioinformatics*, vol. 81, no. 2, pp. 240–252, 2013.
- [43] M. Davis and D. Tobi, "Multiple Gaussian network modes alignment reveals dynamically variable regions: the hemoglobin case," *Proteins: Structure, Function and Bioinformatics*, vol. 82, no. 9, pp. 2097–2105, 2014.
- [44] A. Giuliani, L. Di Paola, and R. Setola, "Proteins as networks: a mesoscopic approach using haemoglobin molecule as case study," *Current Proteomics*, vol. 6, no. 4, pp. 235–245, 2009.
- [45] L. B. Caruso, A. Giuliani, and A. Colosimo, "Allosteric transitions of proteins studied by topological networks: a preliminary investigation on human haemoglobin," *Biophysics and Bioengineering Letters*, vol. 5, no. 1, pp. 1–10, 2012.
- [46] L. Meireles, M. Gur, A. Bakan, and I. Bahar, "Pre-existing soft modes of motion uniquely defined by native contact topology facilitate ligand binding to proteins," *Protein Science*, vol. 20, no. 10, pp. 1645–1658, 2011.
- [47] C. Chennubhotla and I. Bahar, "Markov propagation of allosteric effects in biomolecular systems: application to GroEL-GroES," *Molecular Systems Biology*, vol. 2, article no. 36, 2006.
- [48] C. Chennubhotla, Z. Yang, and I. Bahar, "Coupling between global dynamics and signal transduction pathways: a mechanism of allostery for chaperonin GroEL," *Molecular BioSystems*, vol. 4, no. 4, pp. 287–292, 2008.
- [49] K. V. Brinda and S. Vishveshwara, "A network representation of protein structures: implications for protein stability," *Biophysical Journal*, vol. 89, no. 6, pp. 4159–4170, 2005.
- [50] A. Giuliani and L. Di Paola, "Protein as networks: will contact maps hold the promise to represent the 'structural-formula' of protein molecules?" *Current Protein and Peptide Science*, vol. 17, no. 1, article 3, 2016.
- [51] P. Paci, L. Di Paola, D. Santoni, M. de Ruvo, and A. Giuliani, "Structural and functional analysis of hemoglobin and serum albumin through protein long-range interaction networks," *Current Proteomics*, vol. 9, no. 3, pp. 160–166, 2012.
- [52] D. Santoni, P. Paci, L. Di Paola, and A. Giuliani, "Are proteins just coiled cords? Local and global analysis of contact maps reveals the backbone-dependent nature of proteins," *Current Protein and Peptide Science*, vol. 17, no. 1, pp. 26–29, 2016.
- [53] J. D. Ullman and M. Yannakakis, "High-probability parallel transitive-closure algorithms," *SIAM Journal on Computing*, vol. 20, no. 1, pp. 100–125, 1991.
- [54] G. Amitai, A. Shemesh, E. Sitbon et al., "Network analysis of protein structures identifies functional residues," *Journal of Molecular Biology*, vol. 344, no. 4, pp. 1135–1146, 2004.
- [55] R. Guimerà, M. Sales-Pardo, and L. A. N. Amaral, "Classes of complex networks defined by role-to-role connectivity profiles," *Nature Physics*, vol. 3, no. 1, pp. 63–69, 2007.
- [56] S. Tasdighian, L. Di Paola, M. De Ruvo et al., "Modules identification in protein structures: the topological and geometrical solutions," *Journal of Chemical Information and Modeling*, vol. 54, no. 1, pp. 159–168, 2014.
- [57] F. Cumbo, P. Paci, D. Santoni, L. Di Paola, and A. Giuliani, "GIANT: a cytoscape plugin for modular networks," *PLoS ONE*, vol. 9, no. 10, Article ID e105001, 2014.
- [58] A. Krishnan, J. P. Zbilut, M. Tomita, and A. Giuliani, "Proteins as networks: usefulness of graph theory in protein science," *Current Protein and Peptide Science*, vol. 9, no. 1, pp. 28–38, 2008.

- [59] H. Li, S. Sakuraba, A. Chandrasekaran, and L.-W. Yang, "Molecular binding sites are located near the interface of intrinsic dynamics domains (IDDs)," *Journal of Chemical Information and Modeling*, vol. 54, no. 8, pp. 2275–2285, 2014.
- [60] U. Emekli, D. Schneidman-Duhovny, H. J. Wolfson, R. Nussinov, and T. Haliloglu, "HingeProt: automated prediction of hinges in protein structures," *Proteins: Structure, Function and Genetics*, vol. 70, no. 4, pp. 1219–1227, 2008.
- [61] L.-W. Yang and I. Bahar, "Coupling between catalytic site and collective dynamics: a requirement for mechanochemical activity of enzymes," *Structure*, vol. 13, no. 6, pp. 893–904, 2005.
- [62] T. L. Rodgers, P. D. Townsend, D. Burnell et al., "Modulation of global low-frequency motions underlies allosteric regulation: demonstration in CRP/FNR family transcription factors," *PLoS Biology*, vol. 11, no. 9, Article ID e1001651, 2013.
- [63] C. Chennubhotla and I. Bahar, "Signal propagation in proteins and relation to equilibrium fluctuations," *PLoS Computational Biology*, vol. 3, no. 9, pp. 1716–1726, 2007.
- [64] A. Dutta and I. Bahar, "Metal-binding sites are designed to achieve optimal mechanical and signaling properties," *Structure*, vol. 18, no. 9, pp. 1140–1148, 2010.
- [65] B. M. Butler, Z. N. Gerek, S. Kumar, and S. B. Ozkan, "Conformational dynamics of nonsynonymous variants at protein interfaces reveals disease association," *Proteins: Structure, Function and Bioinformatics*, vol. 83, no. 3, pp. 428–435, 2015.
- [66] N. Kantarci, P. Doruker, and T. Haliloglu, "Cooperative fluctuations point to the dimerization interface of p53 core domain," *Biophysical Journal*, vol. 91, no. 2, pp. 421–432, 2006.
- [67] A. Zen, C. Micheletti, O. Keskin, and R. Nussinov, "Comparing interfacial dynamics in protein-protein complexes: an elastic network approach," *BMC Structural Biology*, vol. 10, article 26, 2010.
- [68] S. Soner, P. Ozbek, J. I. Garzon, N. Ben-Tal, and T. Haliloglu, "DynaFace: discrimination between obligatory and non-obligatory protein-protein interactions based on the complex's dynamics," *PLoS Computational Biology*, vol. 11, no. 10, Article ID e1004461, 2015.
- [69] O. K. Yusuff, J. O. Babalola, G. Bussi, and S. Raugei, "Role of the subunit interactions in the conformational transitions in adult human hemoglobin: an explicit solvent molecular dynamics study," *Journal of Physical Chemistry B*, vol. 116, no. 36, pp. 11004–11009, 2012.
- [70] O. N. Yagurtcu, S. B. Erdemli, R. Nussinov, M. Turkyay, and O. Keskin, "Restricted mobility of conserved residues in protein-protein interfaces in molecular simulations," *Biophysical Journal*, vol. 94, no. 9, pp. 3475–3485, 2008.
- [71] E. Cukuroglu, H. B. Engin, A. Gursoy, and O. Keskin, "Hot spots in protein-protein interfaces: towards drug discovery," *Progress in Biophysics and Molecular Biology*, vol. 116, no. 2-3, pp. 165–173, 2014.
- [72] C. Chennubhotla, A. J. Rader, L.-W. Yang, and I. Bahar, "Elastic network models for understanding biomolecular machinery: from enzymes to supramolecular assemblies," *Physical Biology*, vol. 2, no. 4, pp. S173–S180, 2005.
- [73] I. Bahar, A. R. Atilgan, M. C. Demirel, and B. Erman, "Vibrational dynamics of folded proteins: significance of slow and fast motions in relation to function and stability," *Physical Review Letters*, vol. 80, no. 12, pp. 2733–2736, 1998.
- [74] P. Ozbek, S. Soner, and T. Haliloglu, "Hot spots in a network of functional sites," *PLoS ONE*, vol. 8, no. 9, Article ID e74320, 2013.
- [75] C. B. M. Platania, L. Di Paola, G. M. Leggio et al., "Molecular features of interaction between VEGFA and anti-angiogenic drugs used in retinal diseases: a computational approach," *Frontiers in Pharmacology*, vol. 6, p. 248, 2015.
- [76] A. R. Atilgan, P. Akan, and C. Baysal, "Small-world communication of residues and significance for protein dynamics," *Biophysical Journal*, vol. 86, no. 1, pp. 85–91, 2004.
- [77] A. Del Sol, H. Fujihashi, D. Amoros, and R. Nussinov, "Residues crucial for maintaining short paths in network communication mediate signaling in proteins," *Molecular Systems Biology*, vol. 2, no. 1, 2006.
- [78] G. Bagler and S. Sinha, "Assortative mixing in protein contact networks and protein folding kinetics," *Bioinformatics*, vol. 23, no. 14, pp. 1760–1767, 2007.
- [79] G. Hu, W. Yan, J. Zhou, and B. Shen, "Residue interaction network analysis of Dronpa and a DNA clamp," *Journal of Theoretical Biology*, vol. 348, pp. 55–64, 2014.
- [80] F. Raimondi, A. Felling, M. Seeber, S. Mariani, and F. Fanelli, "A mixed protein structure network and elastic network model approach to predict the structural communication in biomolecular systems: The PDZ2 domain from tyrosine phosphatase 1E as a case study," *Journal of Chemical Theory and Computation*, vol. 9, no. 5, pp. 2504–2518, 2013.