

# Coarse-grained and all-atom modeling of structural states and transitions in hemoglobin

Mustafa Tekpinar and Wenjun Zheng\*

Physics Department, University at Buffalo, Buffalo, New York 14260

## ABSTRACT

Hemoglobin (Hb), an oxygen-binding protein composed of four subunits ( $\alpha_1$ ,  $\alpha_2$ ,  $\beta_1$ , and  $\beta_2$ ), is a well-known example of allosteric proteins that are capable of cooperative ligand binding. Despite decades of studies, the structural basis of its cooperativity remains controversial. In this study, we have integrated coarse-grained (CG) modeling, all-atom simulation, and structural data from X-ray crystallography and wide-angle X-ray scattering (WAXS), aiming to probe dynamic properties of the two structural states of Hb (T and R state) and the transitions between them. First, by analyzing the WAXS data of unliganded and liganded Hb, we have found that the structural ensemble of T or R state is dominated by one crystal structure of Hb with small contributions from other crystal structures of Hb. Second, we have used normal mode analysis to identify two distinct quaternary rotations between the  $\alpha_1\beta_1$  and  $\alpha_2\beta_2$  dimer, which drive the transitions between T and R state. We have also identified the hot-spot residues whose mutations are predicted to greatly change these quaternary motions. Third, we have generated a CG transition pathway between T and R state, which predicts a clear order of quaternary and tertiary changes involving  $\alpha$  and  $\beta$  subunits in Hb. Fourth, we have used the accelerated molecular dynamics to perform an all-atom simulation starting from the T state of Hb, and we have observed a transition toward the R state of Hb. Further analysis of crystal structural data and the all-atom simulation trajectory has corroborated the order of quaternary and tertiary changes predicted by CG modeling.

Proteins 2013; 81:240–252.  
© 2012 Wiley Periodicals, Inc.

**Key words:** hemoglobin; coarse-grained modeling; elastic network model; normal mode analysis; transition pathway; reaction coordinate; accelerated molecular dynamics; wide-angle X-ray scattering.

## INTRODUCTION

Hemoglobin (Hb) is a tetrameric protein composed of four subunits ( $\alpha_1$ ,  $\alpha_2$ ,  $\beta_1$ , and  $\beta_2$ ) symmetrically arranged as a dimer of dimers ( $\alpha_1\beta_1$  and  $\alpha_2\beta_2$ ). Each Hb subunit contains a heme group that can reversibly bind exogenous ligand such as oxygen ( $O_2$ ) or carbon monoxide (CO). Physiologically, Hb acts as an  $O_2$  carrier to transport  $O_2$  from lungs to tissues. This function requires Hb to bind  $O_2$  in a cooperative way<sup>1</sup>—the binding of  $O_2$  to the first heme increases the  $O_2$ -binding affinity of the second heme and successive hemes during the four-step  $O_2$ -binding process. The cooperative binding of  $O_2$  in Hb is modulated by endogenous heterotropic effectors (such as carbon dioxide and hydrogen ions), which are of great interest in medicine.<sup>2</sup>

The cooperativity mechanism of Hb and other allosteric proteins has been extensively studied for many decades. Two competing models were advanced in the 1960s. Koshland, Nemethy, and Filmer (KNF) proposed the

induced-fit sequential model,<sup>3</sup> which postulates that binding of a ligand to a multisubunit protein induces a conformational change in the ligand-binding subunit, which in turn propagates to neighboring subunits to alter their conformation and reactivity. Monod, Wyman, and Changeux (MWC) proposed the two-state concerted model,<sup>4</sup> which postulates that an allosteric protein has only two states (the T state with low affinity for ligand and the R state with high affinity for ligand), and ligand binding induces a transition from T to R state without involving any intermediate state. The MWC model was supported by early structural studies of Hb,<sup>5,6</sup> which

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: American Heart Association; Grant number: 0835292N; Grant sponsor: National Science Foundation; Grant number: 0952736

\*Correspondence to: Wenjun Zheng, Physics Department, University at Buffalo, Buffalo, NY 14260. E-mail: wjzheng@buffalo.edu

Received 9 July 2012; Revised 27 August 2012; Accepted 10 September 2012

Published online 14 September 2012 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.24180

yielded only two distinct quaternary structures—one for unliganded Hb (named T structure) and the other for fully liganded Hb (named R structure). These two structures differ by quaternary changes featuring a torsional rotation ( $\sim 15^\circ$ ) of the  $\alpha_1\beta_1$  dimer relative to the  $\alpha_2\beta_2$  dimer. To further explain how ligand binding triggers the T  $\rightarrow$  R quaternary transition, Perutz proposed a detailed stereochemical model with the following signal transmission pathway—O<sub>2</sub> binding causes a shift in the heme iron position, and a displacement of the proximal F helix, followed by a series of helical rearrangements, leading to successive breakage of the intersubunit salt bridges, eventually resulting in the T  $\rightarrow$  R quaternary structural changes.<sup>7</sup> The Perutz model was later substantiated and extended by Szabo and Karplus.<sup>8</sup>

Recent structural studies have challenged the premise of two structurally unique states for Hb. In addition to the classical R structure, several different quaternary structures of fully liganded Hb were solved by X-ray crystallography (see a review<sup>2</sup>), which were termed R2,<sup>9</sup> RR2,<sup>10</sup> RR3,<sup>2</sup> R3,<sup>10</sup> etc. Other quaternary structures of Hb were reported which are intermediate between the R and R2 structure.<sup>11</sup> Different unliganded Hb structures were also solved, although they are much more similar to one another than the liganded Hb structures.<sup>12</sup> Together, these findings suggest that the T or R state of Hb corresponds to an ensemble of energetically accessible structures instead of a unique structure. Further supporting the ensemble nature of Hb states, previous NMR studies found that the solution structure of CO-bound Hb<sup>13,14</sup> or unliganded Hb<sup>15</sup> does not correspond to any of the Hb crystal structures. In solution, the fully liganded Hb was found to adopt a time-averaged quaternary structure that is intermediate between the R and R2 structure.<sup>13</sup>

Another challenge to the MWC model concerns the functional role of tertiary structural changes, which were found to be uncoupled from quaternary structural changes in Hb.<sup>16</sup> Therefore, kinetic models of Hb need to be revised to incorporate both quaternary and tertiary structural transitions.<sup>16</sup> Along this line, Henry *et al.* proposed a tertiary two-state (TTS) allosteric model,<sup>17</sup> which postulates that high- and low-affinity conformations of individual subunits of Hb (named r and t) exist in equilibrium within each quaternary structure (R and T). It remains unclear what is the dynamic order of quaternary and tertiary structural changes in Hb during the T  $\leftrightarrow$  R transition.

To probe the dynamic events during the R  $\rightarrow$  T transition of Hb, time-resolved spectroscopic studies were carried out following photolysis of ligand.<sup>18,19</sup> The consensus picture emerging from these studies is that protein tertiary relaxation takes place first (<1  $\mu$ s after photolysis), whereas the quaternary transition occurs later ( $\sim 20$   $\mu$ s after photolysis).<sup>20–22</sup> To directly probe structural changes with nanosecond time resolution, the time-resolved wide-angle X-ray scattering (TR-WAXS) was

utilized to probe the R  $\rightarrow$  T transition triggered by laser-induced photolysis.<sup>23,24</sup> As revealed by TR-WAXS, after the breaking of Hb-CO bonds, Hb undergoes tertiary structural changes in <150 ns, followed by the quaternary transition in 2–3  $\mu$ s.<sup>23,24</sup> The structural details for the above dynamic events are still under investigation.<sup>24</sup>

In complement with experimental studies, computer simulations promise to offer molecular details of the cooperativity mechanism in Hb. To simulate the dynamics of Hb with full details, the all-atom molecular dynamics (MD) simulation is the method of choice given its unparalleled ability to probe the atomic details of protein dynamics in solution. Previous MD simulations were attempted for Hb, with simulation time ranging from 200 ps to 300 ns.<sup>25–28</sup> A recent long-time MD study<sup>28</sup> successfully observed the T  $\rightarrow$  R quaternary transition of the Hb tetramer and tertiary transitions of the  $\alpha$  and  $\beta$  subunits, but the R  $\rightarrow$  T transition remains out of reach to MD simulations with today's computing power. As alternatives to MD simulations, Mouawad and Perahia used normal mode analysis (NMA) to explore the intermediate conformations between T and R state,<sup>29</sup> and Fischer *et al.* computed a minimum-energy pathway from T to R state.<sup>30</sup>

To overcome the time-scale barrier to all-atom MD simulations, coarse-grained (CG) modeling has been developed using simplified structural representations and energy functions.<sup>31</sup> At residue level of details, the elastic network model (ENM) represents a protein structure as a network of C <sub>$\alpha$</sub>  atoms connected by elastic springs<sup>32–34</sup> with a uniform force constant.<sup>35</sup> The NMA of ENM yields low-frequency modes, which were found to compare well with many crystallographically observed structural changes.<sup>34,36</sup> Many studies have established ENM as an effective means to simulate protein conformational dynamics with no limit in timescale or system size (see reviews<sup>37,38</sup>). ENM has also provided a useful framework for developing CG computational techniques for identification of dynamically important residues,<sup>39</sup> transition pathway modeling,<sup>40–47</sup> flexible fitting of distance constraints,<sup>48,49</sup> and low-resolution structural data such as electron-microscopy maps<sup>50,51</sup> and solution X-ray scattering data.<sup>52</sup>

In this study, we have applied the ENM-based CG modeling and all-atom simulation to the crystal structures and WAXS data of Hb, aiming to probe dynamic properties of the structural states (T and R state) and the transitions between them:

First, we have used the WAXS data of unliganded and CO-bound Hb<sup>53</sup> to analyze the structural ensembles of T and R state of Hb. We use two different WAXS-based modeling methods—one based on all-atom representation of a fixed protein and explicit solvents<sup>54</sup> and the other based on CG representation of a flexible protein and an implicit hydration shell.<sup>52</sup> We have found that the T-state ensemble is dominated by the T structure of Hb, whereas the R-state ensemble is dominated by the R2 structure of Hb.

Second, we have used NMA to identify two key normal modes that drive the transitions between T and R state—the first mode captures a well-known torsional rotation of the  $\alpha 1\beta 1$  dimer relative to the  $\alpha 2\beta 2$  dimer; the second mode captures a new hinge-bending rotation between the  $\alpha 1\beta 1$  dimer and the  $\alpha 2\beta 2$  dimer, which affects the packing in the interdimer interface. We have also identified the hot-spot residues dynamically involved in these quaternary motions.

Third, we have used iENM<sup>41</sup> to generate a CG transition pathway between T and R state, which predicts a clear order of quaternary and tertiary changes in Hb: quaternary changes of  $\beta 1\beta 2$  subunits  $\rightarrow$  quaternary changes of  $\alpha 1\alpha 2$  subunits  $\rightarrow$  tertiary changes in  $\beta$  subunit  $\rightarrow$  tertiary changes in  $\alpha$  subunit. This predicted order is in good agreement with crystal structural data and past studies.

Fourth, we have used the accelerated MD protocol<sup>55</sup> to perform an all-atom simulation starting from the T state of Hb, and we have observed a transition toward the R2 structure of Hb. Further analysis of the simulation trajectory has corroborated the order of quaternary and tertiary changes predicted by CG modeling.

## METHODS

### ENM and NMA

In an ENM, a protein structure is represented as a network of beads each corresponding to the  $C_\alpha$  atom of an amino acid residue. A harmonic potential accounts for the elastic interaction between a pair of  $C_\alpha$  atoms that are within a cutoff distance  $R_c$  chosen to be 10 Å (we have also tried  $R_c = 13$  Å following another ENM study of Hb,<sup>56</sup> which gives very similar results). The ENM potential energy is

$$E = \frac{1}{2} \sum_{i=2}^N \sum_{j=1}^{i-1} k_{ij} \theta(R_c - d_{ij}^0) (d_{ij} - d_{ij}^0)^2, \quad (1)$$

where  $N$  is the number of  $C_\alpha$  atoms,  $\theta(x)$  is the Heaviside function,  $d_{ij}$  is the distance between the  $C_\alpha$  atom  $i$  and  $j$ , and  $d_{ij}^0$  is the value of  $d_{ij}$  as given by a protein crystal structure.  $k_{ij}$  is the force constant, which is 1 for nonbonded interactions and 10 for bonded interactions between residues.

We can expand the ENM potential energy to second order:

$$E \approx \frac{1}{2} X^T H X, \quad (2)$$

where  $X$  is a  $3N$ -dimension vector representing the 3D displacement of  $N C_\alpha$  atoms away from their equilibrium positions,  $H$  is the Hessian matrix, which is obtained by calculating the second derivatives of ENM potential energy with respect to the 3D coordinates of  $C_\alpha$  atoms.

NMA solves  $HW_m = \lambda_m W_m$ , where  $\lambda_m$  and  $W_m$  represent the eigenvalue and eigenvector of mode  $m$ . After excluding six zero modes corresponding to three rotations and three translations, we keep  $3N - 6$  nonzero modes, which are numbered from 1 to  $3N - 6$  in order of ascending eigenvalue.

To validate NMA, we compare each mode (mode  $m$ ) with the observed structural changes between two crystal structures (represented by a  $3N$ -dimension vector  $X_{obs}$ ) by calculating the following overlap:

$$I_m = \frac{|X_{obs}^T W_m|}{|X_{obs}| \cdot |W_m|}, \quad (3)$$

where  $X_{obs}^T W_m$  is the dot product between vectors  $X_{obs}$  and  $W_m$ ,  $|X_{obs}|$  and  $|W_m|$  represent their magnitudes.

The ENM-based NMA is available via a webserver at <http://enm.lobos.nih.gov/start.html>.

### CG transition pathway modeling by iENM

Recently, we have developed an interpolated ENM (iENM) protocol for modeling a transition pathway between two given protein structures.<sup>41</sup> iENM solves the saddle points of a double-well potential  $F(E_{ENM1} + E_{col}, E_{ENM2} + E_{col})$ , where  $E_{ENM1}$  and  $E_{ENM2}$  are two ENM potentials [see Eq. (1)] based on the beginning and end structure of the transition, and  $E_{col}$  is a steric collision energy.<sup>41</sup> All the saddle points trace a pathway that connects the beginning and end structure, which is independent of the mathematic form of function  $F(x, y)$ . This method is available via a web server at [http://enm.lobos.nih.gov/start\\_ienm.html](http://enm.lobos.nih.gov/start_ienm.html).

### Quantification of motional order of key protein parts during a transition

We use the predicted pathway to determine the motional order of two protein parts (such as  $\alpha 1$  subunit vs.  $\beta 1$  subunit of Hb) during a transition. To this end, the following reaction coordinate (RC) is defined for a given part S<sup>41</sup>:

$$RC_S = 0.5 \left( 1 + \frac{RMSD_{S,1}^2 - RMSD_{S,2}^2}{RMSD_{S,obs}^2} \right), \quad (4)$$

where  $RMSD_{S,1}$  ( $RMSD_{S,2}$ ) is the RMSD of  $C_\alpha$  atoms of part S between a given intermediate conformation and the beginning (end) conformation, and  $RMSD_{S,obs}$  is the corresponding RMSD between the beginning and end conformation.  $RC_S$  varies from 0 to 1 as the transition proceeds from the beginning to the end conformation. For two protein parts S and S', if  $RC_S < RC_{S'}$  along the pathway, we can infer that the motion of S' precedes that of S.

## Fitting of WAXS data with known crystal structures of Hb

We minimize the following score function  $\chi$  by linearly combining the WAXS profiles calculated from three crystal structure of Hb:

$$\chi =$$

$$\sqrt{\min \left\{ \frac{1}{N_s} \sum_{n=1}^{N_s} [c_1 I_R(s_n) + c_2 I_{R2}(s_n) + c_3 I_T(s_n) - I_{exp}(s_n)]^2 / \sigma_n^2 \right\}}. \quad (5)$$

where  $I_R, I_{R2}$ , and  $I_T$  are the WAXS profiles calculated from the R, R2, and T structure of Hb,  $I_{exp}$  is the experimental WAXS data,  $\sigma_n$  is the experimental error at  $s = s_n$ ,  $N_s$  is the number of  $s$  values between 0 and  $0.15 \text{ \AA}^{-1}$ ,  $c_1$ ,  $c_2$ , and  $c_3$  are the weights to be optimized.

## Flexible fitting of WAXS data

We have recently developed a new flexible fitting method based on a CG (one-bead-per-residue) protein representation, an implicit hydration shell model and a modified form of ENM that allows large-scale conformational changes while maintaining pseudobonds and secondary structures.<sup>52</sup> By optimizing a pseudoenergy that combines the modified ENM energy with a WAXS-fitting score and a collision energy that penalizes steric collisions, we can generate a new CG structural model that fits best with a given WAXS profile.

## Accelerated MD simulation

To speed up MD simulations of Hb, we have used the accelerated molecular dynamics<sup>55</sup> (aMD) to simulate the  $T \rightarrow R$  transition. The NAMD implementation of aMD is used.<sup>57</sup> aMD accelerates slow conformational transitions of proteins by adding a boost potential  $\Delta V = \frac{(E-V)^2}{\alpha + (E-V)}$  to the dihedral potential  $V$  of the system. During an aMD simulation, the boost potential is turned on if the dihedral potential energy  $V$  falls below a threshold energy  $E$ . Another parameter for the boost potential is the acceleration factor  $\alpha$ , which determines the depth of dihedral potential basin below  $E$ . These parameters are chosen as follows:  $E = 3381.9 \text{ kcal/(mol A)}$  and  $\alpha = 114.8 \text{ kcal/(mol A)}$ .

The T structure of Hb (PDB id: 2hhb) has been selected as the starting conformation of aMD simulation. None of the histidines is protonated. The protein is submerged within a solvent box. The system is neutralized by adding sodium ions. We applied 2000 steps of energy minimization, followed by a 3-ns equilibration run (with a 2-fs time step), then a 20-ns production run under constant temperature ( $T = 293 \text{ K}$ ) and constant volume conditions. Two hundred snapshots of the production run are saved every 100 ps.

## RESULTS AND DISCUSSION

### Analysis of T/R-state structural ensembles of Hb using WAXS data

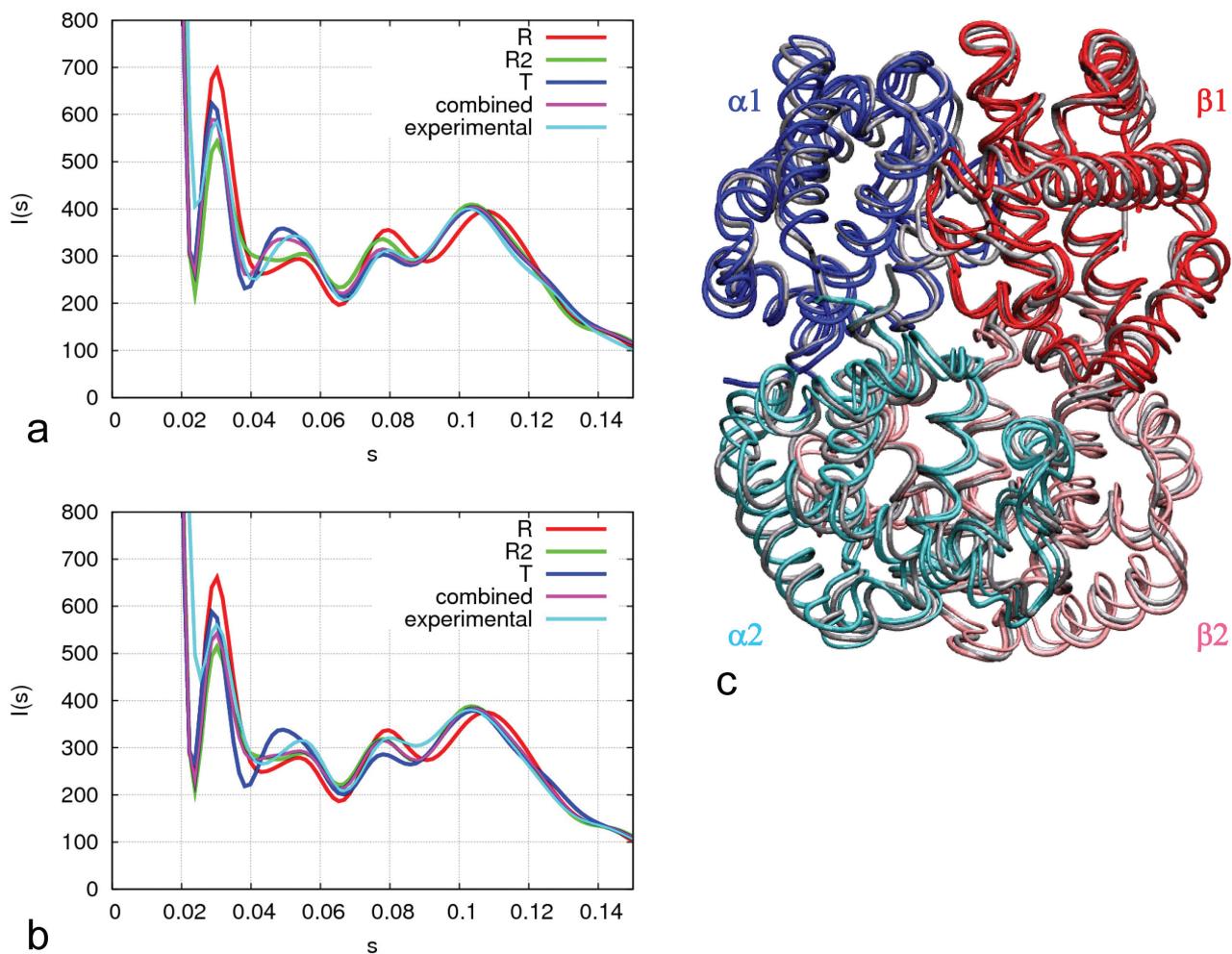
WAXS, as an extension of small-angle X-ray scattering, provides information about the global structure of a protein, its local secondary structures and their arrangements. WAXS represents a powerful approach for exploring structural relationships between different forms of a protein in crystals versus those in solution. The WAXS data for unliganded and CO-bound Hb, corresponding to T and R state, were collected at 50 mg/mL protein concentration by Makowski *et al.*<sup>53</sup> The WAXS patterns from Hb include four distinct peaks<sup>53</sup>: the first peak at  $\sim 0.03 \text{ \AA}^{-1}$  indicates degrees of rotation of  $\alpha\beta$  dimers relative to each other; the first and second peaks at 0.03 and  $0.05 \text{ \AA}^{-1}$  are indicative of the shape of Hb; the third and fourth peaks at 0.08 and  $0.1 \text{ \AA}^{-1}$  are indicative of local structures (such as packing of  $\alpha$ -helices).

Recent NMR<sup>13–15</sup> and WAXS<sup>53</sup> studies of Hb in solution indicate that Hb adopts an ensemble of conformations in both T and R state. To explore the structural ensembles of T/R state using a combination of different crystal structures of Hb, we have used the XS program developed by Park *et al.*<sup>54</sup> to calculate the WAXS profile for each crystal structure of Hb and then linearly combine those calculated profiles to best fit the experimental WAXS data (the combination weight of each structure gives its population in the ensemble). The XS program uses short MD simulation to sample configurations of explicit solvents around a fixed protein, and it was shown to accurately reproduce high-angle features of WAXS data.<sup>54</sup> To fully equilibrate the solvent degrees of freedom, we have run 1-ns MD simulation instead of 100 ps used by Makowski *et al.*<sup>53</sup> Indeed, we have found that the use of longer MD simulations has substantially reduced the sharp peak at  $\sim 0.03 \text{ \AA}^{-1}$  found by Makowski *et al.*<sup>53</sup>

### T-state ensemble

We have compared calculated WAXS profiles from individual Hb structures with the experimental WAXS data for unliganded Hb, which yields  $\chi = 2.86, 4.03$ , and  $3.19$  for the T, R, and R2 structure (PDB ids: 2dn2, 2dn3, and 1bbb, respectively). As expected, the T structure fits best with the WAXS data. Indeed, the WAXS profile calculated from the T structure nicely reproduces the four peaks observed experimentally, although it slightly overestimates the height of the first peak and the second peak is slightly shifted [see Fig. 1(a)].

We have optimized the fitting of experimental WAXS data for unliganded Hb using a linear combination of the three calculated WAXS profiles, which has led to slightly improved fitting with  $\chi = 2.73$  [see Fig. 1(a)]. The optimized weight is 66, 0, and 34% for the T, R, and

**Figure 1**

Results for analysis of the R/T-state ensembles using WAXS data. (a) Fitting of experimental WAXS data of unliganded Hb (cyan) using WAXS profiles calculated from three crystal structures of Hb [PDB ids: 2dn2 (blue), 2dn3 (red), and 1bbb (green)] and their optimized linear combination (purple). (b) Fitting of experimental WAXS data of CO-bound Hb (cyan) using WAXS profiles calculated from three crystal structures of Hb [PDB ids: 2dn2 (blue), 2dn3 (red), and 1bbb (green)] and their optimized linear combination (purple). (c) Comparison of the flexibly fitted Hb model of R state (shown as gray tube) with the R structure (PDB id: 2dn3, shown as thin colored tube) and the R2 structure (PDB id: 1bbb, shown as thick colored tube), where the four subunits of R/R2 structure are colored as follows:  $\alpha_1$  (blue),  $\alpha_2$  (cyan),  $\beta_1$  (red), and  $\beta_2$  (pink).

R2 structure, respectively. To test the robustness of these weights, we have recalculated WAXS profiles using different MD trajectories. We have found that the optimal weight of the T structure ( $\sim 70\%$ ) is reproducible, whereas the smaller weights of the R and R2 structure are not. Therefore, we conclude that the T-state ensemble of Hb is dominated by the T structure with small contributions from the R or R2 or both structures.

#### R-state ensemble

We have compared calculated WAXS profiles from individual Hb structures with the experimental WAXS data for CO-bound Hb, which yields  $\chi = 20.22$ ,  $19.78$ , and  $14.83$  for the T, R, and R2 structure, respectively. We

have also tried other liganded structures of Hb (RR2, R3, and RR3), which do not fit better than the R2 structure. Therefore, the R2 structure fits best with the WAXS data, whereas the T structure fits worst. Indeed, the WAXS profile calculated from the R2 structure nicely reproduces the four peaks observed experimentally, although it slightly underestimates the height of the first and second peaks [see Fig. 1(b)].

We have optimized the fitting of experimental WAXS data for CO-bound Hb using a linear combination of the three calculated WAXS profiles, which has led to slightly improved fitting with  $\chi = 14.38$  [see Fig. 1(b)]. The optimized weight is 16, 11, and 72% for the T, R, and R2 structure, respectively. To test the robustness of these weights, we have recalculated WAXS profiles using

different MD trajectories. We have found that the optimal weight of the R2 structure ( $\sim 70\%$ ) is reproducible, whereas the small weights of the T and R structure are not (varying by up to 10%). Therefore, we conclude that the R-state ensemble of Hb is dominated by the R2 structure with small contributions from the R or T or both structures.

In support of our finding that the R2 structure is the best representative of R-state ensemble, an NMR study showed that the R2 structure fits better to the residual dipolar coupling data than the R structure<sup>13</sup>; a TR-WAXS study found that the use of the R2 structure instead of the R structure for calculating WAXS profiles led to better agreement with the 100- $\mu$ s difference of WAXS scattering data.<sup>23</sup> Therefore, we will use the R2 structure of Hb as a representative of R state for the CG modeling in this study.

#### Flexible WAXS fitting at R state

An alternative way to explore structural ensembles of Hb using WAXS data is to build a new Hb conformation with optimal fitting to the given WAXS data, which may correspond to the average conformation of a structural ensemble. To this end, we have used a recently developed flexible fitting method<sup>52</sup> (see Methods section). We start from the R structure of Hb and iteratively deform the Hb conformation to improve the fitting to the WAXS data for CO-bound Hb. During the iterations, the overall RMSD to the R2 structure decreases slightly from 1.7 to 1.4 Å. Interestingly, the RMSD for  $\beta_1\beta_2$  subunits decreases significantly from 1.6 to 1.0 Å, whereas the RMSD for  $\alpha_1\alpha_2$  subunits barely changes (from 1.8 to 1.6 Å). Therefore, we infer that the average conformation of R-state ensemble of Hb lies intermediate between the R and R2 structure with the  $\beta_1\beta_2$  subunits more similar to the R2 structure than the  $\alpha_1\alpha_2$  subunits [see Fig. 1(c)]. Such a marked asymmetry between the  $\alpha$  and  $\beta$  subunits was also observed in a recent MD simulation study.<sup>28</sup>

#### NMA predicts two quaternary motions in Hb driving T $\leftrightarrow$ R transition

After establishing the T and R2 structure as representatives for the T and R state of Hb, we will use NMA to analyze those quaternary motions intrinsic to these two structures, which may drive the T  $\leftrightarrow$  R structural transition. This transition involves quaternary changes of the  $\alpha_1\beta_1$  dimer relative to the  $\alpha_2\beta_2$  dimer, which affect two key regions in the  $\alpha_1-\beta_2$  (or  $\alpha_2-\beta_1$ ) interface: (i) the contacts between the FG corner (residues 91–97) of  $\alpha_1/\alpha_2$  subunit and the C helix (residues 35–43) of  $\beta_2/\beta_1$  subunit known as the “joint” region; (ii) the contacts between the C helix (residues 38–44) of  $\alpha_1/\alpha_2$  subunit and the FG corner (residues 97–102) and C-terminus of  $\beta_2/\beta_1$  subunit known as the “switch” region. The transition from the T to R2 structure features a sliding motion

between the C helix of  $\alpha_1/\alpha_2$  subunit and the FG corner of  $\beta_2/\beta_1$  subunit, and loosened packing in the switch region (see Fig. 2).

#### NMA of T state

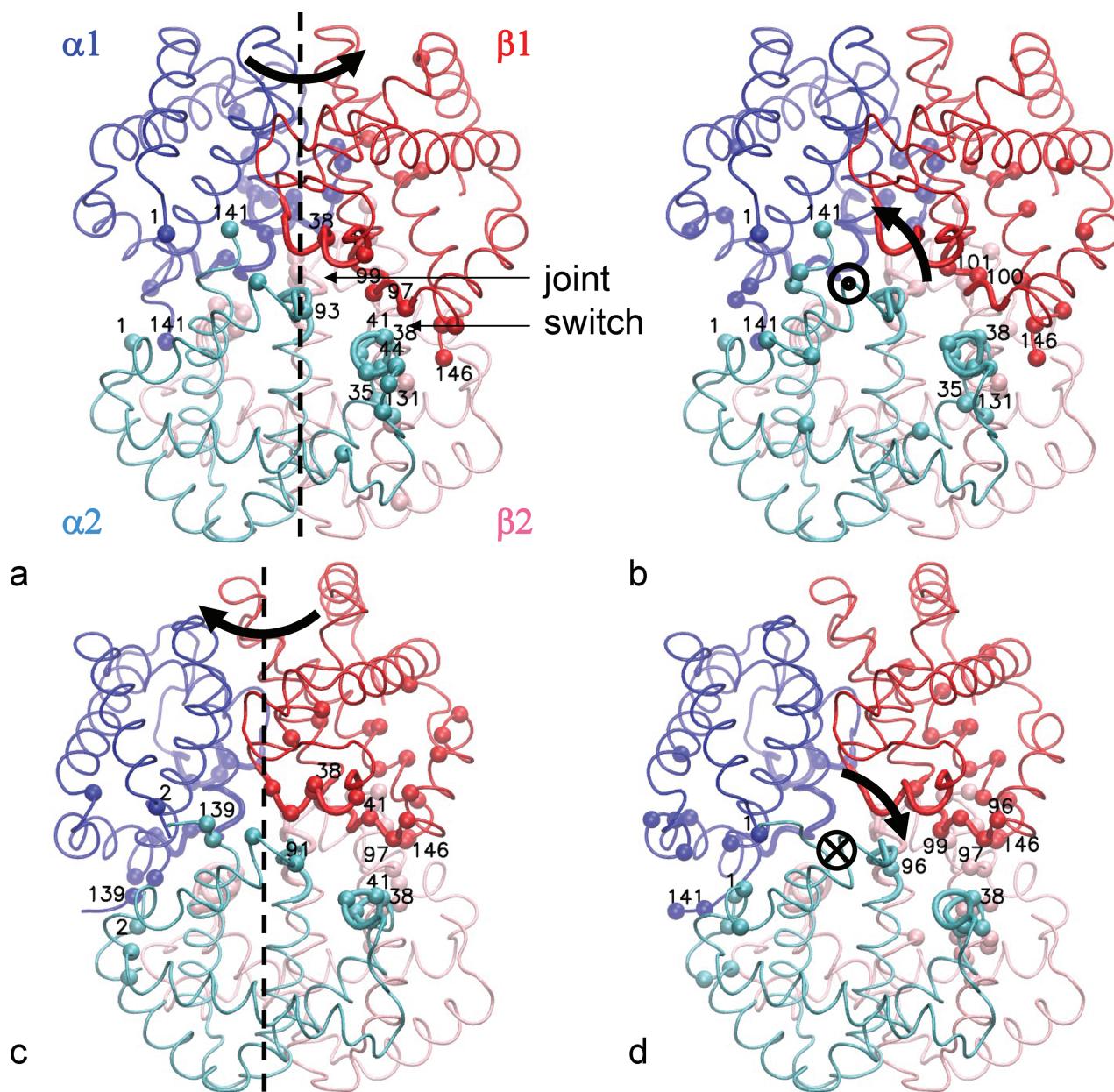
We first perform NMA for an ENM constructed from the T structure (PDB id: 2dn2) and then compare each normal mode with the observed conformational changes from the T to R2 structure (PDB id: 1bbb). Only two modes (#1 and #2) are found to contribute significantly to the T  $\rightarrow$  R2 structural transition, with mode #2 more involved than mode #1 (see Table I).

Mode #1 describes a torsional rotation of the  $\alpha_1\beta_1$  dimer relative to the  $\alpha_2\beta_2$  dimer [see Fig. 2(a) and Supporting Information movie S1], which is well known as the hallmark quaternary motion for the T  $\rightarrow$  R transition. This rotation results in a sliding between the FG corner of  $\beta_2/\beta_1$  subunit and the C helix of  $\alpha_1/\alpha_2$  subunit (in the switch region), with the distance between His97 of  $\beta_2/\beta_1$  subunit and Thr38 of  $\alpha_1/\alpha_2$  subunit decreasing and the distance between His97 of  $\beta_2/\beta_1$  subunit and Pro44 of  $\alpha_1/\alpha_2$  subunit increasing. This sliding causes His97 of  $\beta_2/\beta_1$  subunit to move from the groove between Pro44 and Thr41 of  $\alpha_1/\alpha_2$  subunit by a turn to the groove between Thr41 and Thr38 of  $\alpha_1/\alpha_2$  subunit, as observed between the T and R structure.<sup>2</sup> This rotation also causes the narrowing of  $\beta_1-\beta_2$  cleft, which was observed between the T and R structure.<sup>2</sup>

Mode #2 describes a hinge-bending rotation of the  $\alpha_1\beta_1$  dimer relative to the  $\alpha_2\beta_2$  dimer [see Fig. 2(b) and Supporting Information movie S2] with the rotational axis passing near the two joint regions in the interdimer interface. This previously unnoticed rotation results in slight opening of the interdimer interface (in the switch region), with the distance between Pro100 of  $\beta_2/\beta_1$  subunit and Thr38 of  $\alpha_1/\alpha_2$  subunit increasing. Such opening may loosen the packing in the switch region, which will enable the sliding motion between the FG corner of  $\beta_2/\beta_1$  subunit and the C helix of  $\alpha_1/\alpha_2$  subunit without encountering a steric barrier,<sup>9</sup> and allow the FG corner of  $\beta_2/\beta_1$  subunit to disengage from the C helix of  $\alpha_1/\alpha_2$  subunit in the R2 structure.<sup>2</sup>

The finding of two relevant quaternary motions adds more complexity to the T  $\rightarrow$  R quaternary transition in Hb. The combination of these two quaternary motions may give Hb more flexibility to visit an ensemble of R-state conformations (such as R, R2, RR2, RR3, and R3). Indeed, the conformational changes from the T structure to those R-state conformations all involve different combinations of mode #1 and #2 (see Table I).

To further identify the hot-spot residues dynamically involved in the above two quaternary motions, we have used an ENM-based perturbation method,<sup>58</sup> which evaluates the perturbational effect of turning off the elastic interactions involving each residue position on a given

**Figure 2**

Results of NMA for the T and R2 structures of Hb: (a) torsional rotation of the  $\alpha_1\beta_1$  dimer relative to the  $\alpha_2\beta_2$  dimer described by mode #1 of the T structure; (b) hinge-bending rotation of the  $\alpha_1\beta_2$  dimer relative to the  $\alpha_2\beta_2$  dimer described by mode #2 of the T structure; (c) torsional rotation of the  $\alpha_1\beta_1$  dimer relative to the  $\alpha_1\beta_1$  dimer described by mode #1 of the R2 structure; (d) hinge-bending rotation of the  $\alpha_1\beta_1$  dimer relative to the  $\alpha_2\beta_2$  dimer described by mode #3 of the R2 structure. The T structure in (a) and (b) is aligned with the R2 structure in (c) and (d) along the  $\alpha_2\beta_2$  dimer. Same color scheme as Figure 1(c) is used for  $\alpha$  and  $\beta$  subunits. The rotational directions are shown by curved arrows. The rotational axes in (a) and (c) are shown as dashed lines. The rotational axes in (b) and (d) are roughly perpendicular to the paper. The switch and joint regions are labeled and shown as thick tube. Hot-spot residues are shown as spheres with residue numbers labeled for some of them. For dynamic visualization of the above quaternary motions, see Supporting Information movies S1–S4.

mode, and then selects the top 10% residues with the largest effect.<sup>58</sup>

The hot-spot residues for mode #1 are distributed in the following intersubunit interfaces [see Fig. 2(a)]: some are in the  $\alpha_1-\beta_2$  and  $\alpha_2-\beta_1$  interfaces, including the

switch region (residues 38, 39, 40, 41, 43, and 44 of  $\alpha_1\alpha_2$ , and 97, 98, 99, 145, and 146 of  $\beta_1\beta_2$ ) and the joint region (residues 93 and 97 of  $\alpha_1/\alpha_2$  and 38, 40, and 41 of  $\beta_1/\beta_2$ ); some are in the  $\alpha_1-\alpha_2$  interface (residues 1 and 141 of  $\alpha_1/\alpha_2$ ); some are in the  $\alpha_1-\beta_1$  and  $\alpha_2-\beta_2$

**Table I**

Results of NMA for the T and R2 Structures of Hb

PDB ids	RMSD (Å)	Mode#	Eigenvalue	Overlap
2dn2 (T) → 2dn3 (R)	2.4	1	0.025312	0.643
		2	0.032012	0.533
2dn2 (T) → 1bbb (R2)	3.5	1		0.391
		2		0.714
2dn2 (T) → 1mko (RR2)	3.0	1		0.543
		2		0.597
2dn2 (T) → 1yzi (R3)	3.8	1		0.827
		2		0.322
2dn2 (T) → 3d17 (RR3)	2.8	1		0.825
		2		0.268
1bbb (R2) → 2dn2 (T)	3.5	1	0.039239	0.399
		3	0.053021	0.652
		4		<0.2
1bbb (R2) → 2dn3 (R)	1.7	1		0.688
		3		0.212
1bbb (R2) → 1mko (RR2)	1.0	1		0.657
		3		0.583
1bbb (R2) → 1yzi (R3)	2.6	1		0.522
		3		0.327
1bbb (R2) → 3d17 (RR3)	2.5	1		0.700
		3		

interfaces (residues 34, 35, 36, and 37 of  $\alpha_1/\alpha_2$ , and 128, 131, 132, and 133 of  $\beta_1/\beta_2$ ).

The hot-spot residues for mode #2 are distributed in the following intersubunit interfaces [see Fig. 2(b)]: some are in the switch region of the  $\alpha_1-\beta_2$  and  $\alpha_2-\beta_1$  interfaces (residues 38, 39, 40, and 42 of  $\alpha_1/\alpha_2$ , and 100, 101, 144, 145, and 146 of  $\beta_1/\beta_2$ ); some are in the  $\alpha_1-\alpha_2$  interface (residues 1 and 141 of  $\alpha_1/\alpha_2$ ); some are in the  $\alpha_1-\beta_1$  and  $\alpha_2-\beta_2$  interfaces (residues 33, 34, 35, 36, and 37 of  $\alpha_1/\alpha_2$ , and 131, 132, and 133 of  $\beta_1/\beta_2$ ).

Some of the above hot-spot residues were studied previously. Arg141 of  $\alpha$  subunits and His146 of  $\beta$  subunits contribute to the interdimer salt bridges, which lock Hb in the low-affinity T structure. An NMR study found that the mobility of His146 increases significantly upon CO binding.<sup>59</sup> Arg141 was also found to be highly mobile in the R2 structure.<sup>9</sup> The loss of Arg141 or His146 was found to increase the overall ligand-binding affinity of Hb,<sup>60–63</sup> presumably by tuning the  $T \leftrightarrow R$  transition. Mutation Tyr145Gly in  $\beta$  subunits was found to have a substantial kinetic effect on CO binding in the T state.<sup>64</sup> Mutations of Gln131 of  $\beta$  subunits resulted in lower O<sub>2</sub> affinity and perturbed hydrogen bonds in the  $\alpha_1-\beta_1$  interface.<sup>65</sup> In agreement with our finding, a previous NMA study identified both the joint and switch region as hinge centers for the lowest normal mode.<sup>56</sup> Future mutational studies will be needed to further assess the functional roles of those unexplored hot-spot residues.

#### NMA of R state

We then perform NMA for an ENM constructed from the R2 structure (PDB id: 1bbb) and compare each mode with the observed conformational changes from the R2 to T structure (PDB id: 2dn2). Only two modes (#1 and

#3) are found to contribute significantly to the  $R2 \rightarrow T$  transition, with mode #3 contributing more than mode #1 (see Table I).

Mode #1 describes a torsional rotation of the  $\alpha_1\beta_1$  dimer relative to the  $\alpha_2\beta_2$  dimer [see Fig. 2(c) and Supporting Information movie S3], which is very similar to mode #1 of the T structure (but in opposite direction). This rotation results in a sliding between the FG corner of  $\beta_2/\beta_1$  subunit and the C helix of  $\alpha_1/\alpha_2$  subunit (in the switch region), causing His97 of  $\beta_2/\beta_1$  subunit to move from the groove between Thr41 and Thr38 of  $\alpha_1/\alpha_2$  subunit toward the groove between Pro44 and Thr41 of  $\alpha_1/\alpha_2$  subunit.<sup>2</sup> This rotation also leads to the opening of  $\beta_1-\beta_2$  cleft.<sup>2</sup>

Mode #3 describes a hinge-bending rotation of the  $\alpha_1\beta_1$  dimer relative to the  $\alpha_2\beta_2$  dimer [see Fig. 2(d) and Supporting Information movie S4], which resembles mode #2 of the T structure (but in opposite direction). This rotation results in slight closing of the interdimer interface and tighter packing in the switch region, which will enable the formation of interdimer salt bridges characteristic of the T state.

The combination of these two quaternary motions may give Hb more flexibility to transition from the R2 to T structure and other R-state conformations (such as R, RR2, RR3, and R3, see Table I).

The hot-spot residues for mode #1 are distributed in the following intersubunit interfaces [see Fig. 2(c)]: some are in the  $\alpha_1-\beta_2$  and  $\alpha_2-\beta_1$  interfaces, including the switch region (residues 38, 39, 41, and 42 of  $\alpha_1/\alpha_2$ , and 97, 98, 145, and 146 of  $\beta_1/\beta_2$ ) and the joint region (residues 91 and 93 of  $\alpha_1/\alpha_2$ , and 36, 37, 38, 39, and 41 of  $\beta_1/\beta_2$ ); some are in the  $\alpha_1-\alpha_2$  interface (residues 2, 138, and 139 of  $\alpha_1/\alpha_2$ ).

The hot-spot residues for mode #3 are distributed in the following intersubunit interfaces [see Fig. 2(d)]: some are in the switch region of the  $\alpha_1-\beta_2$  and  $\alpha_2-\beta_1$  interfaces (residues 38, 39, 40, and 42 of  $\alpha_1/\alpha_2$ , and 97, 98, 99, 102, 145, and 146 of  $\beta_1/\beta_2$ ); some are in the  $\alpha_1-\alpha_2$  interface (residues 1, 140, and 141 of  $\alpha_1/\alpha_2$ ).

The above hot-spot residues mostly overlap with those identified for the T structure of Hb. Some of them were studied previously. For example, a resonance Raman spectroscopy study<sup>66</sup> showed that the first step of  $R \rightarrow T$  transition involves the joint-region contacts (as monitored by the hydrogen bond between Trp37 of  $\beta_2/\beta_1$  subunit and Asp94 of  $\alpha_1/\alpha_2$  subunit), whereas the second involves the switch-region contacts (as monitored by the hydrogen bond between Tyr42 of  $\alpha_1/\alpha_2$  subunit and Asp99 of  $\beta_2/\beta_1$  subunit). Future mutational studies will be needed to further assess the functional roles of those unexplored hot-spot residues.

In summary, the NMA of the T and R2 structures has revealed two key quaternary motions (torsional and hinge-bending rotation of the  $\alpha_1\beta_1$  dimer relative to the  $\alpha_2\beta_2$  dimer), which may drive the  $T \leftrightarrow R$  transition in

both directions. These two quaternary motions enable Hb to visit various R-state conformations via sliding between the  $\alpha_1\beta_1$  and  $\alpha_2\beta_2$  dimer,<sup>11</sup> which may be facilitated by slight opening of the interdimer interface.

The involvement of two modes suggests that it is not sufficient to only consider the lowest mode for exploring the T  $\leftrightarrow$  R transition.<sup>29,56</sup> Indeed, starting from the R2 structure and following the lowest mode<sup>56</sup> will only lower the RMSD (relative to the T structure) slightly from 3.5 to 3.2 Å. In contrast, if mode #3 is followed, the RMSD can be lowered significantly to 2.6 Å. Therefore, by following both modes, Hb may transition more readily toward the T state without being trapped in the R state.<sup>29,56</sup>

Despite high similarity between the quaternary motions captured by modes #1 and #2 of the T structure and modes #1 and #3 of the R2 structure, these modes are energetically different. Indeed, the eigenvalues of the two modes of the T structure are significantly lower than the corresponding modes of the R2 structure (see Table I). Because lower eigenvalue implies that lower energy is needed to excite the mode, we infer that the T state has higher quaternary structural fluctuations than the R state, which is consistent with previous studies by MD simulation<sup>67</sup> and WAXS.<sup>53</sup>

#### **Transition pathway modeling predicts the order of tertiary and quaternary changes in Hb during T $\leftrightarrow$ R transition**

NMA is limited to the prediction of possible collective motions accessible to a protein near an equilibrium state corresponding to an energy minimum. To probe intermediate conformational changes taking place during a transition between two states, one must go beyond NMA and explore transition pathways between two distinct energy minima. To this end, we have recently developed a ENM-based transition pathway modeling method (iENM),<sup>41</sup> which generates a transition pathway (consisting of a series of intermediate conformations) that connects two given conformations (named beginning and end conformation). This new method is able to predict a clear sequence of structural events involving different protein parts during a protein conformational transition. It has been used to illuminate the allosteric coupling mechanisms in molecular motors like myosin<sup>44,46</sup> and dynein.<sup>47</sup> In this study, we will use iENM to analyze the dynamic order of quaternary and tertiary structural changes involving  $\alpha$  and  $\beta$  subunits of Hb.

We have simulated the conformational transition between T and R state of Hb, which are represented by the T and R2 structure of Hb (PDB ids: 2dn2 and 1bbb). To this end, we have generated a transition pathway from the T to R2 structure using iENM (see Supporting Information movie S5). We then validate the predicted pathway by comparing the predicted order of quaternary and

tertiary changes with the observed order from a set of 257 Hb structures (see Supporting Information Table S1). The comparison is quantified using the following RCs (see Methods section):  $RC_{all}$ ,  $RC_{\alpha_1\alpha_2}$ , and  $RC_{\beta_1\beta_2}$  quantify the progress of quaternary changes involving the entire tetramer, the  $\alpha_1\alpha_2$  subunits, and the  $\beta_1\beta_2$  subunits, respectively;  $RC_{\alpha_1}$  and  $RC_{\beta_1}$  quantify the progress of tertiary changes in the  $\alpha_1$  and  $\beta_1$  subunit, respectively.  $RC = 0$  corresponds to the T state, and  $RC = 1$  corresponds to the R state [see Fig. 3(a) and Supporting Information Fig. S1].

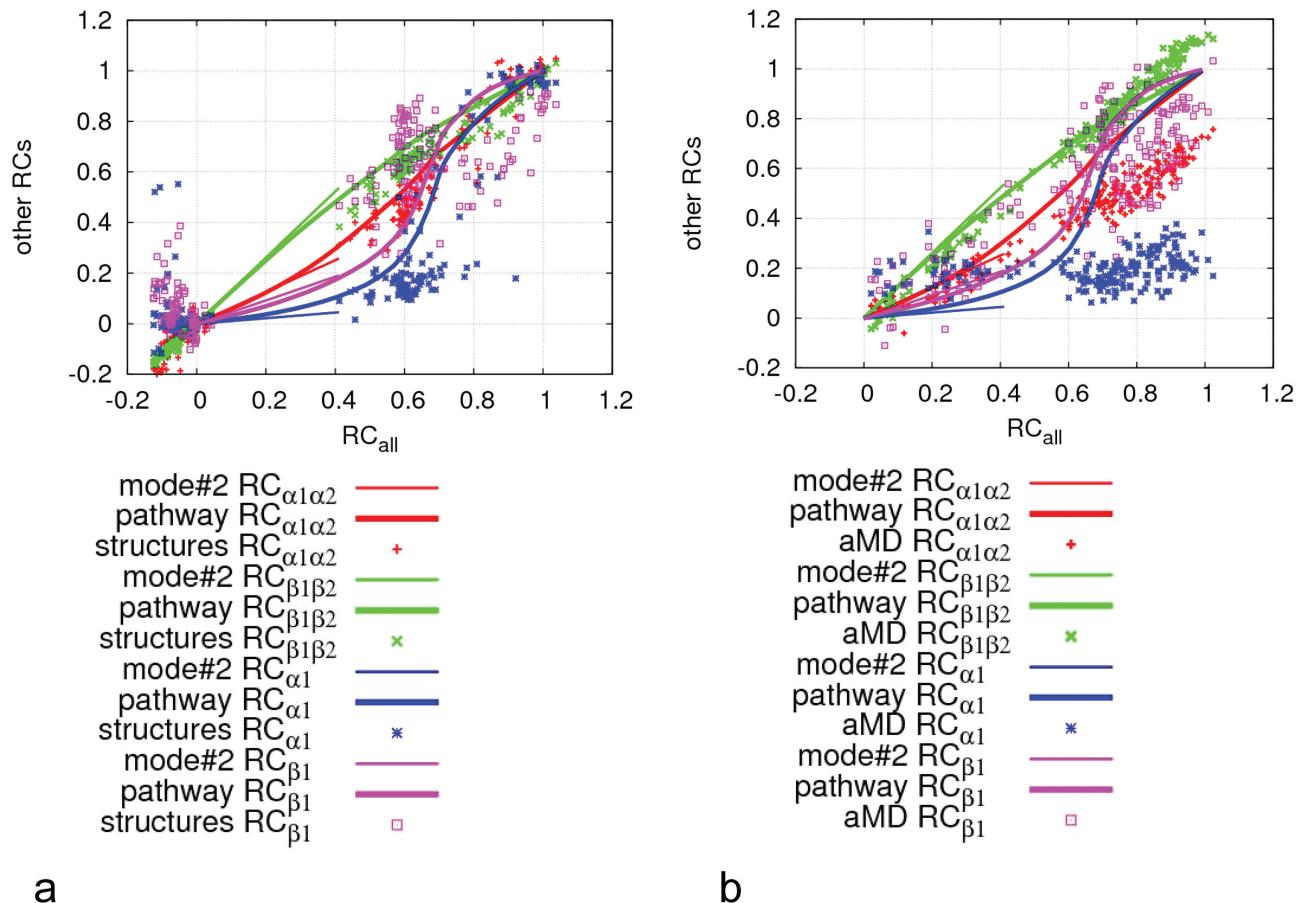
The analysis of iENM pathway indicates  $RC_{\beta_1\beta_2} > RC_{\alpha_1\alpha_2} > RC_{\beta_1} > RC_{\alpha_1}$  during most of the transition [for  $0 < RC_{all} < 0.7$ , see Fig. 3(a)], which implies the following order: quaternary changes of  $\beta_1\beta_2$  subunits  $\rightarrow$  quaternary changes of  $\alpha_1\alpha_2$  subunits  $\rightarrow$  tertiary changes of  $\beta$  subunit  $\rightarrow$  tertiary changes of  $\alpha$  subunit. For  $RC_{all} > 0.7$ ,  $RC_{\beta_1}$  and  $RC_{\alpha_1}$  increase sharply to catch up with  $RC_{\beta_1\beta_2}$  and  $RC_{\alpha_1\alpha_2}$ , so that they are more intimately coupled to one another [see Fig. 3(a)]. The above order is also encoded in the normal mode driving the T  $\rightarrow$  R transition. Indeed, the RC analysis of a trajectory generated by following mode #2 of the T structure also gives  $RC_{\beta_1\beta_2} > RC_{\alpha_1\alpha_2} > RC_{\beta_1} > RC_{\alpha_1}$  [see Fig. 3(a)].

To validate the predicted pathway, we have calculated RCs for 257 Hb structures [see Fig. 3(a)]. The T, R, and R2 structures of Hb are distributed around  $RC_{all} = 0$ , 0.6, and 1, respectively. In addition, there are some intermediate structures between the T and R structures (PDB ids: 1sdk, 1sdl, etc., see Supporting Information Table S1), and some between the R and R2 structures (PDB ids: 1g08, 1g09, 1g0a, etc., see Supporting Information Table S1). The finding that the R structures lie between the T and R2 structures in the RC plot [see Fig. 3(a)] supports the proposal that the R structure is the pathway from the T to R2 structure.<sup>68–70</sup>

The RC analysis of those intermediate Hb structures with  $0.4 < RC_{all} < 0.7$  clearly indicates  $RC_{\beta_1\beta_2} > RC_{\alpha_1\alpha_2}$  and  $RC_{\beta_1} > RC_{\alpha_1}$  [see Fig. 3(a)], supporting the predicted order of quaternary changes of  $\beta_1\beta_2$  subunits  $\rightarrow$  quaternary changes of  $\alpha_1\alpha_2$  subunits and tertiary changes of  $\beta$  subunit  $\rightarrow$  tertiary changes of  $\alpha$  subunit. However, as  $RC_{\beta_1}$  is widely distributed [see Fig. 3(a)], it is hard to assign the relative order between the tertiary changes of  $\beta$  subunit and the two quaternary changes.

Despite qualitative agreement, the asymmetry between the tertiary changes of  $\alpha$  and  $\beta$  subunit predicted by our CG modeling is less pronounced than indicated by the structural data [see Fig. 3(a)]. This may be due to the inaccuracy in CG modeling of tertiary changes in Hb. Another source of errors is that our CG modeling does not consider ligand binding explicitly, so it cannot model ligation-induced structural changes, which may amplify the asymmetry between  $\alpha$  and  $\beta$  subunit.

To probe the origin of dynamic asymmetry between  $\alpha$  and  $\beta$  subunits, we have repeated the CG modeling using

**Figure 3**

Analysis of predicted transition pathway from the T to R2 structure using reaction coordinates (RCs): (a) comparison with crystal structures of Hb; (b) comparison with aMD simulation. The reaction coordinates  $RC_{\beta_1\beta_2}$ ,  $RC_{\alpha_1\alpha_2}$ ,  $RC_{\beta_1}$ , and  $RC_{\alpha_1}$  are plotted versus  $RC_{all}$  and colored green, red, purple, and blue, respectively. The curves correspond to the predicted pathway. The data points correspond to crystal structures of Hb in (a) and snapshots of aMD simulation in (b). The straight lines correspond to mode #2 of the T structure. For dynamic visualization of the predicted pathway, see Supporting Information movie S5.

a smaller force constant for the intersubunit contacts than the intrasubunit contacts. We have found that the  $\alpha$ - $\beta$  asymmetry is reduced as a result. Therefore, the intersubunit interactions are responsible for the observed  $\alpha$ - $\beta$  asymmetry in Hb.

Next, we will relate our CG modeling results to past studies of the  $T \leftrightarrow R$  transition in Hb.

#### ***T → R transition***

Our results suggest that the  $T \rightarrow R$  transition proceeds in the order of quaternary changes followed by tertiary changes. This is consistent with the notion that the T-state quaternary assembly constrains a T-like tertiary conformation of each subunit leading to low O<sub>2</sub> affinity.<sup>16</sup> Therefore, the removal of such quaternary constraints by early quaternary changes is needed to allow the  $T \rightarrow R$  tertiary changes to proceed. Indeed, under conditions in which the  $T \rightarrow R$  quaternary changes are

impaired by crystal lattice, a noncooperative binding isotherm was observed and characterized by very low affinity,<sup>71</sup> indicating the tertiary changes that enable high affinity are also impaired. In agreement with our results, an early NMA study found that the intermediate conformations between the T and R structure are quaternary R-like and tertiary T-like.<sup>29</sup> Moreover, the same authors computed the transition pathway from the T to R structure with the Path Exploration with Distance Constraints method,<sup>25</sup> which showed that the  $T \rightarrow R$  transition proceeds in two steps—the quaternary changes take place before the tertiary ones.

Another prediction for the  $T \rightarrow R$  transition is that  $\beta$  subunits are more active than  $\alpha$  subunits at both tertiary and quaternary levels. In support of this prediction, an early crystallographic study by Paoli *et al.* found that the structural effect of partial ligation of the deoxy Hb crystals was more pronounced in  $\beta$  subunits than  $\alpha$  subunits.<sup>72</sup> Another study also found that the structural

changes associated with ligation in the T state are larger in  $\beta$  subunits than  $\alpha$  subunits.<sup>73</sup> A Raman spectroscopy study observed larger structural flexibility around the  $\beta$ -heme group than the  $\alpha$ -heme group.<sup>74</sup> The MD simulations by Hub *et al.* also revealed a marked asymmetry between the  $\alpha$  and  $\beta$  subunits with the tertiary transition of  $\beta$  subunits more strongly correlated to the quaternary transition than  $\alpha$  subunits.<sup>28</sup>

#### **R $\rightarrow$ T transition**

Our results suggest that the R  $\rightarrow$  T transition proceeds in the order of tertiary changes followed by quaternary changes. This is consistent with the consensus picture from time-resolved spectroscopy studies—protein tertiary relaxation takes place before the quaternary transition after photolysis.<sup>20–22</sup>

Another prediction for the R  $\rightarrow$  T transition is that  $\alpha$  subunits are more active than  $\beta$  subunits at both tertiary and quaternary levels. In support of our prediction, Adachi *et al.* studied structural changes induced by photodissociation of fully liganded Hb in the R state and they found that the tertiary changes involving the heme and the F-helix are more marked in  $\alpha$  subunits than  $\beta$  subunits.<sup>75</sup> Wilson *et al.* solved a crystal structure of deoxy Hb trapped in the R state, and they found that the tertiary changes associated with ligand binding/unbinding in the R state differ significantly between  $\alpha$  and  $\beta$  subunits, with the magnitude smaller in  $\beta$  subunits than  $\alpha$  subunits.<sup>76</sup> Samaja *et al.* studied the dissociation of CO from several intermediate ligation states of Hb and found that  $\alpha$  subunits dissociate CO faster than  $\beta$  subunits.<sup>77</sup>

In summary, our finding suggests that the quaternary and tertiary changes in Hb are not tightly coupled during the R  $\leftrightarrow$  T transition. This may allow allosteric effectors to induce tertiary changes to regulate O<sub>2</sub> affinity without quaternary changes,<sup>78</sup> which cannot be explained by the MWC model. Our finding is also consistent with the TTS model,<sup>17</sup> which postulates that the unliganded T state contains nearly 100% of subunits in the t conformation (the t  $\rightarrow$  r tertiary changes cannot proceed before the T  $\rightarrow$  R quaternary changes), whereas the unliganded R state contains significant populations of both r and t conformations (the r  $\rightarrow$  t tertiary changes can occur before the R  $\rightarrow$  T quaternary changes).

#### **Accelerated MD simulation of T $\rightarrow$ R transition in Hb**

To evaluate the predictions based on the CG modeling of T  $\leftrightarrow$  R transition in Hb, we have performed all-atom simulations of Hb in solution. To overcome the time-scale barrier of regular MD simulations, we have used the aMD protocol<sup>55</sup> to simulate the T  $\rightarrow$  R transition for 20 ns, starting from the T structure of Hb (see Methods section). We have generated seven aMD trajectories. Encouragingly, in one aMD trajectory, we have observed that Hb moves

toward the R2 structure (see Supporting Information movie S6), reaching a minimal RMSD of  $\sim$ 2.4 Å (relative to the R2 structure) in  $\sim$ 3 ns. Similar conformational transition was also observed in another aMD trajectory.

To assess the order of tertiary and quaternary changes in Hb as predicted by the CG modeling, we have calculated RCs for 200 snapshots of the aMD trajectory [see Fig. 3(b)]. The distributions of RCs suggest  $RC_{\beta_1\beta_2} > RC_{\alpha_1\alpha_2}$  and  $RC_{\beta_1} > RC_{\alpha_1}$  for  $RC_{\text{all}} > 0.6$  [see Fig. 3(b)], which supports the predicted order of quaternary changes of  $\beta_1\beta_2$  subunits  $\rightarrow$  quaternary changes of  $\alpha_1\alpha_2$  subunits and tertiary changes of  $\beta$  subunit  $\rightarrow$  tertiary changes of  $\alpha$  subunit. Similar to the RCs of Hb crystal structures, the distribution of  $RC_{\beta_1}$  is wider than  $RC_{\alpha_1}$  [see Fig. 3(a,b)]. Unlike the RCs of iENM pathway,  $RC_{\alpha_1\alpha_2}$  and  $RC_{\alpha_1}$  remain well below 1 throughout the aMD simulation [see Fig. 3(b)], which suggests that the full structural changes in  $\alpha$  subunits are beyond the reach of 20-ns aMD simulation.

We have also attempted an aMD simulation starting from the R2 structure of Hb, which failed to produce a trajectory toward the T structure of Hb. This is consistent with past unsuccessful efforts of MD simulations to observe the R  $\rightarrow$  T transition.<sup>28</sup>

## **CONCLUSION**

In conclusion, we have performed a computational study that combines CG modeling, all-atom simulation, and structural data from X-ray crystallography and WAXS, which aims to probe dynamic properties of the two structural states of Hb (T and R state) and the transitions between them. First, by analyzing the WAXS data of unliganded and liganded Hb, we have found that the structural ensemble of T/R state is dominated by the T/R2 structure of Hb with small contributions from other crystal structures of Hb. Second, we have used NMA to identify two quaternary rotations between the  $\alpha_1\beta_1$  and  $\alpha_2\beta_2$  dimer that drive the transitions between T and R state, and we have identified the hot-spot residues dynamically involved in these quaternary rotations. Third, we have generated a CG transition pathway between T and R state, which predicts a clear order of quaternary and tertiary changes in Hb: quaternary changes of  $\beta_1\beta_2$  subunits  $\rightarrow$  quaternary changes of  $\alpha_1\alpha_2$  subunits  $\rightarrow$  tertiary changes in  $\beta$  subunit  $\rightarrow$  tertiary changes in  $\alpha$  subunit. Fourth, we have used the accelerated MD to perform an all-atom simulation starting from the T state of Hb, and we have observed a transition toward the R state of Hb. Further analysis of the simulation trajectory has corroborated the order of quaternary and tertiary changes predicted by CG modeling. This work has offered structural insights for past studies of Hb and new predictions for future experiments. The integrated methods used in this study will be useful to future study of other allosteric proteins.

## ACKNOWLEDGMENTS

The WAXS data of Hb from Dr. Makowski are gratefully acknowledged.

## REFERENCES

- Eaton WA, Henry ER, Hofrichter J, Bettati S, Viappiani C, Mozzarelli A. Evolution of allosteric models for hemoglobin. *IUBMB Life* 2007;59:586–599.
- Safo MK, Ahmed MH, Ghatge MS, Boyiri T. Hemoglobin-ligand binding: understanding Hb function and allostery on atomic level. *Biochim Biophys Acta* 2011;1814:797–809.
- Koshland DE, Jr, Nemethy G, Filmer D. Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochemistry* 1966;5:365–385.
- Monod J, Wyman J, Changeux JP. On the nature of allosteric transitions: a plausible model. *J Mol Biol* 1965;12:88–118.
- Muirhead H, Perutz MF. Structure of haemoglobin. A three-dimensional Fourier synthesis of reduced human haemoglobin at 5-5 Å resolution. *Nature* 1963;199:633–638.
- Fermi G, Perutz MF, Shaanan B, Fourme R. The crystal structure of human deoxyhaemoglobin at 1.74 Å resolution. *J Mol Biol* 1984;175:159–174.
- Perutz MF. Stereochemistry of cooperative effects in haemoglobin. *Nature* 1970;228:726–739.
- Szabo A, Karplus M. A mathematical model for structure-function relations in hemoglobin. *J Mol Biol* 1972;72:163–197.
- Silva MM, Rogers PH, Arnone A. A third quaternary structure of human hemoglobin A at 1.7-A resolution. *J Biol Chem* 1992;267:17248–17256.
- Safo MK, Abraham DJ. The enigma of the liganded hemoglobin end state: a novel quaternary structure of human carbonmonoxy hemoglobin. *Biochemistry* 2005;44:8347–8359.
- Mueser TC, Rogers PH, Arnone A. Interface sliding as illustrated by the multiple quaternary structures of liganded hemoglobin. *Biochemistry* 2000;39:15353–15364.
- Kavanaugh JS, Rogers PH, Arnone A. Crystallographic evidence for a new ensemble of ligand-induced allosteric transitions in hemoglobin: the T-to-T(high) quaternary transitions. *Biochemistry* 2005;44:6101–6121.
- Lukin JA, Kontaxis G, Simplaceanu V, Yuan Y, Bax A, Ho C. Quaternary structure of hemoglobin in solution. *Proc Natl Acad Sci USA* 2003;100:517–520.
- Gong Q, Simplaceanu V, Lukin JA, Giovannelli JL, Ho NT, Ho C. Quaternary structure of carbonmonoxyhemoglobins in solution: structural changes induced by the allosteric effector inositol hexaphosphate. *Biochemistry* 2006;45:5140–5148.
- Sahu SC, Simplaceanu V, Gong Q, Ho NT, Tian F, Prestegard JH, Ho C. Insights into the solution structure of human deoxyhemoglobin in the absence and presence of an allosteric effector. *Biochemistry* 2007;46:9973–9980.
- Bellelli A, Brunori M. Hemoglobin allostery: variations on the theme. *Biochim Biophys Acta* 2011;1807:1262–1272.
- Henry ER, Bettati S, Hofrichter J, Eaton WA. A tertiary two-state allosteric model for hemoglobin. *Biophys Chem* 2002;98:149–164.
- Hofrichter J, Sommer JH, Henry ER, Eaton WA. Nanosecond absorption spectroscopy of hemoglobin: elementary processes in kinetic cooperativity. *Proc Natl Acad Sci USA* 1983;80:2235–2239.
- Jayaraman V, Rodgers KR, Mukerji I, Spiro TG. Hemoglobin allostery: resonance Raman spectroscopy of kinetic intermediates. *Science* 1995;269:1843–1848.
- Bjorling SC, Goldbeck RA, Paquette SJ, Milder SJ, Kligler DS. Allosteric intermediates in hemoglobin. I. Nanosecond time-resolved circular dichroism spectroscopy. *Biochemistry* 1996;35:8619–8627.
- Goldbeck RA, Paquette SJ, Bjorling SC, Kligler DS. Allosteric intermediates in hemoglobin. II. Kinetic modeling of HbCO photolysis. *Biochemistry* 1996;35:8628–8639.
- Goldbeck RA, Esquerra RM, Kligler DS. Hydrogen bonding to Trp beta37 is the first step in a compound pathway for hemoglobin allostery. *J Am Chem Soc* 2002;124:7646–7647.
- Camarata M, Levantino M, Schotte F, Anfinrud PA, Ewald F, Choi J, Cupane A, Wulff M, Ihee H. Tracking the structural dynamics of proteins in solution using time-resolved wide-angle X-ray scattering. *Nat Methods* 2008;5:881–886.
- Camarata M, Levantino M, Wulff M, Cupane A. Unveiling the timescale of the R-T transition in human hemoglobin. *J Mol Biol* 2010;400:951–962.
- Mouawad L, Perahia D, Robert CH, Guillet C. New insights into the allosteric mechanism of human hemoglobin from molecular dynamics simulations. *Biophys J* 2002;82:3224–3245.
- Saito M, Okazaki I. A 45-ns molecular dynamics simulation of hemoglobin in water by vectorizing and parallelizing COSMOS90 on the earth simulator: dynamics of tertiary and quaternary structures. *J Comput Chem* 2007;28:1129–1136.
- Laberge M, Yonetani T. Molecular dynamics simulations of hemoglobin A in different states and bound to DPG: effector-linked perturbation of tertiary conformations and HbA concerted dynamics. *Biophys J* 2008;94:2737–2751.
- Hub JS, Kubitzki MB, de Groot BL. Spontaneous quaternary and tertiary T-R transitions of human hemoglobin in molecular dynamics simulation. *PLoS Comput Biol* 2010;6:e1000774.
- Mouawad L, Perahia D. Motions in hemoglobin studied by normal mode analysis and energy minimization: evidence for the existence of tertiary T-like, quaternary R-like intermediate structures. *J Mol Biol* 1996;258:393–410.
- Fischer S, Olsen KW, Nam K, Karplus M. Unsuspected pathway of the allosteric transition in hemoglobin. *Proc Natl Acad Sci USA* 2011;108:5608–5613.
- Tozzini V. Coarse-grained models for proteins. *Curr Opin Struct Biol* 2005;15:144–150.
- Atilgan AR, Durell SR, Jernigan RL, Demirel MC, Keskin O, Bahar I. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys J* 2001;80:505–515.
- Hinsen K. Analysis of domain motions by approximate normal mode calculations. *Proteins* 1998;33:417–429.
- Tama F, Sanejouand YH. Conformational change of proteins arising from normal mode calculations. *Protein Eng* 2001;14:1–6.
- Tirion MM. Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Phys Rev Lett* 1996;77:1905–1908.
- Krebs WG, Alexandrov V, Wilson CA, Echols N, Yu H, Gerstein M. Normal mode analysis of macromolecular motions in a database framework: developing mode concentration as a useful classifying statistic. *Proteins* 2002;48:682–695.
- Bahar I, Rader AJ. Coarse-grained normal mode analysis in structural biology. *Curr Opin Struct Biol* 2005;15:586–592.
- Zheng W, Brooks BR, Thirumalai D. Allosteric transitions in biological nanomachines are described by robust normal modes of elastic networks. *Curr Protein Pept Sci* 2009;10:128–132.
- Zheng W, Tekpinar M. Large-scale evaluation of dynamically important residues in proteins predicted by the perturbation analysis of a coarse-grained elastic model. *BMC Struct Biol* 2009;9:45.
- Zheng W, Brooks BR, Hummer G. Protein conformational transitions explored by mixed elastic network models. *Proteins* 2007;69:43–57.
- Tekpinar M, Zheng W. Predicting order of conformational changes during protein conformational transitions using an interpolated elastic network model. *Proteins* 2010;78:2469–2481.
- Franklin J, Koehl P, Doniach S, Delarue M. MinActionPath: maximum likelihood trajectory for large-scale structural transitions in a coarse-grained locally harmonic energy landscape. *Nucleic Acids Res* 2007;35:W477–W482.

43. Maragakis P, Karplus M. Large amplitude conformational change in proteins explored with a plastic network model: adenylate kinase. *J Mol Biol* 2005;352:807–822.
44. Zheng W. Multiscale modeling of structural dynamics underlying force generation and product release in actomyosin complex. *Proteins* 2010;78:638–660.
45. Zheng W, Auerbach A. Decrypting the sequence of structural events during the gating transition of pentameric ligand-gated ion channels based on an interpolated elastic network model. *PLoS Comput Biol* 2011;7:e1001046.
46. Zheng W. Coarse-grained modeling of conformational transitions underlying the processive stepping of myosin V dimer along filamentous actin. *Proteins* 2011;79:2291–2305.
47. Zheng W. Coarse-grained modeling of the structural states and transition underlying the powerstroke of dynein motor domain. *J Chem Phys* 2012;136:155103.
48. Zheng W, Brooks BR. Normal-modes-based prediction of protein conformational changes guided by distance constraints. *Biophys J* 2005;88:3109–3117.
49. Zheng W, Brooks BR. Modeling protein conformational changes by iterative fitting of distance constraints using reoriented normal modes. *Biophys J* 2006;90:4327–4336.
50. Tama F, Miyashita O, Brooks CL, III. Normal mode based flexible fitting of high-resolution structure into low-resolution experimental data from cryo-EM. *J Struct Biol* 2004;147:315–326.
51. Zheng W. Accurate flexible fitting of high-resolution protein structures into cryo-electron microscopy maps using coarse-grained pseudo-energy minimization. *Biophys J* 2011;100:478–488.
52. Zheng W, Tekpinar M. Accurate flexible fitting of high-resolution protein structures to small-angle x-ray scattering data using a coarse-grained model with implicit hydration shell. *Biophys J* 2011;101:2981–2991.
53. Makowski L, Bardhan J, Gore D, Lal J, Mandava S, Park S, Rodi DJ, Ho NT, Ho C, Fischetti RF. WAXS studies of the structural diversity of hemoglobin in solution. *J Mol Biol* 2011;408:909–921.
54. Park S, Bardhan JP, Roux B, Makowski L. Simulated x-ray scattering of protein solutions using explicit-solvent models. *J Chem Phys* 2009;130:134114.
55. Hamelberg D, Mongan J, McCammon JA. Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. *J Chem Phys* 2004;120:11919–11929.
56. Xu C, Tobi D, Bahar I. Allosteric changes in protein structure computed by a simple mechanical model: hemoglobin T-R2 transition. *J Mol Biol* 2003;333:153–168.
57. Wang Y, Harrison CB, Schulten K, McCammon JA. Implementation of accelerated molecular dynamics in NAMD. *Comput Sci Discov* 4:015002, 2011.
58. Zheng W, Brooks BR, Doniach S, Thirumalai D. Network of dynamically important residues in the open/closed transition in polymerases is strongly conserved. *Structure* 2005;13:565–577.
59. Song XJ, Yuan Y, Simplaceanu V, Sahu SC, Ho NT, Ho C. A comparative NMR study of the polypeptide backbone dynamics of hemoglobin in the deoxy and carbonmonoxy forms. *Biochemistry* 2007;46:6795–6803.
60. Kilmartin JV, Wootton JF. Inhibition of Bohr effect after removal of C-terminal histidines from haemoglobin beta-chains. *Nature* 1970;228:766–767.
61. Antonini E, Wyman J, Zito R, Rossi-Fanelli A, Caputo A. Studies on carboxypeptidase digests of human hemoglobin. *J Biol Chem* 1961;236:PC60–PC63.
62. Kavanagh JS, Chafin DR, Arnone A, Mozzarelli A, Rivetti C, Rossi GL, Kwiatkowski LD, Noble RW. Structure and oxygen affinity of crystalline desArg141 alpha human hemoglobin A in the T state. *J Mol Biol* 1995;248:136–150.
63. Bettati S, Kwiatkowski LD, Kavanagh JS, Mozzarelli A, Arnone A, Rossi GL, Noble RW. Structure and oxygen affinity of crystalline des-his-146beta human hemoglobin in the T state. *J Biol Chem* 1997;272:33077–33084.
64. Noble RW, Hui HL, Kwiatkowski LD, Paily P, DeYoung A, Wierzba A, Colby JE, Bruno S, Mozzarelli A. Mutational effects at the subunit interfaces of human hemoglobin: evidence for a unique sensitivity of the T quaternary state to changes in the hinge region of the alpha 1 beta 2 interface. *Biochemistry* 2001;40:12357–12368.
65. Chang CK, Simplaceanu V, Ho C. Effects of amino acid substitutions at beta 131 on the structure and properties of hemoglobin: evidence for communication between alpha 1 beta 1- and alpha 1 beta 2-subunit interfaces. *Biochemistry* 2002;41:5644–5655.
66. Balakrishnan G, Case MA, Pevsner A, Zhao X, Tengroth C, McLendon GL, Spiro TG. Time-resolved absorption and UV resonance Raman spectra reveal stepwise formation of T quaternary contacts in the allosteric pathway of hemoglobin. *J Mol Biol* 2004;340:843–856.
67. Yonetani T, Laberge M. Protein dynamics explain the allosteric behaviors of hemoglobin. *Biochim Biophys Acta* 2008;1784:1146–1158.
68. Janin J, Wodak SJ. The quaternary structure of carbonmonoxy hemoglobin ypsilonlanti. *Proteins* 1993;15:1–4.
69. Srinivasan R, Rose GD. The T-to-R transformation in hemoglobin: a reevaluation. *Proc Natl Acad Sci USA* 1994;91:11113–11117.
70. Schumacher MA, Zheleznova EE, Poundstone KS, Kluger R, Jones RT, Brennan RG. Allosteric intermediates indicate R2 is the liganded hemoglobin end state. *Proc Natl Acad Sci USA* 1997;94:7841–7844.
71. Mozzarelli A, Rivetti C, Rossi GL, Henry ER, Eaton WA. Crystals of haemoglobin with the T quaternary structure bind oxygen noncooperatively with no Bohr effect. *Nature* 1991;351:416–419.
72. Paoli M, Liddington R, Tame J, Wilkinson A, Dodson G. Crystal structure of T state haemoglobin with oxygen bound at all four haems. *J Mol Biol* 1996;256:775–792.
73. Liddington R, Derewenda Z, Dodson E, Hubbard R, Dodson G. High resolution crystal structures and comparisons of T-state deoxyhaemoglobin and two liganded T-state haemoglobins: T(alpha)-oxy)haemoglobin and T(met)haemoglobin. *J Mol Biol* 1992;228:551–579.
74. Gilch H, Schweitzer-Stenner R, Dreybrodt W. Structural heterogeneity of the Fe(2+)-N epsilon (HisF8) bond in various hemoglobin and myoglobin derivatives probed by the Raman-active iron histidine stretching mode. *Biophys J* 1993;65:1470–1485.
75. Adachi S, Park SY, Tame JR, Shiro Y, Shibayama N. Direct observation of photolysis-induced tertiary structural changes in hemoglobin. *Proc Natl Acad Sci USA* 2003;100:7039–7044.
76. Wilson J, Phillips K, Luisi B. The crystal structure of horse deoxyhaemoglobin trapped in the high-affinity (R) state. *J Mol Biol* 1996;264:743–756.
77. Samaja M, Rovida E, Niggeler M, Perrella M, Rossi-Bernardi L. The dissociation of carbon monoxide from hemoglobin intermediate. *J Biol Chem* 1987;262:4528–4533.
78. Yonetani T, Park SI, Tsunehige A, Imai K, Kanaori K. Global allostery model of hemoglobin. Modulation of O(2) affinity, cooperativity, and Bohr effect by heterotropic allosteric effectors. *J Biol Chem* 2002;277:34508–34520.