JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

ARTICLE

pubs.acs.org/JACS

Frontier Residues Lining Globin Internal Cavities Present Specific Mechanical Properties

Anthony Bocahut,[†] Sophie Bernad,[‡] Pierre Sebban,[‡,§] and Sophie Sacquin-Mora*[,†]

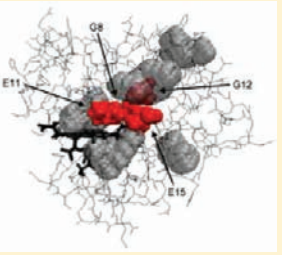
[†]Laboratoire de Biochimie Théorique, UMR 9080 CNRS, Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, 75005 Paris, France

[‡]Laboratoire de Chimie Physique, CNRS UMR8000, Bât. 350, Université Paris-sud, 91405 Orsay, France

[§]Université des Sciences et des Technologies de Hanoi, 18 Hoang Quoc Viet, Cau Giay, Hanoi, Vietnam

Supporting Information

ABSTRACT: The internal cavity matrix of globins plays a key role in their biological function. Previous studies have already highlighted the plasticity of this inner network, which can fluctuate with the proteins breathing motion, and the importance of a few key residues for the regulation of ligand diffusion within the protein. In this Article, we combine all-atom molecular dynamics and coarse-grain Brownian dynamics to establish a complete mechanical landscape for six different globins chain (myoglobin, neuroglobin, cytoglobin, truncated hemoglobin, and chains α and β of hemoglobin). We show that the rigidity profiles of these proteins can fluctuate along time, and how a limited set of residues present specific mechanical properties that are related to their position at the frontier between internal cavities. Eventually, we postulate the existence of conserved positions within the globin fold, which form a mechanical nucleus located at the center of the cavity network, and whose constituent residues are essential for controlling ligand migration in globins.



INTRODUCTION

The globin superfamily is found in all kingdoms of life, and its members can perform a large variety of functions such as NO scavenging, enzymatic activities, oxygen sensing, and, of course, O_2 transport and storage.[1−5] Interestingly, their sequence can be extremely variable, with globins presenting less than 10% homology,[6] and they are best characterized by their common structural feature. This typical 3D fold of a small number of α-helices, named the globin fold, protects a noncovalently bound heme group and alows reversible ligand binding. Despite over 50 years of intensive research,[7] globins still represent a fascinating subject, their structural and functional properties being far from fully understood,[8,9] and with recently discovered members, such as neuroglobin or cytoglobin, whose physiological function has remained elusive until today.[5,10,11]

The internal cavity network located in the matrix of globular proteins usually plays a key role in ligand migration and for the control of protein function.[12−16] In the case of globins, the diffusion pathways of various small ligands have been extensively studied for over 30 years,[17−28] showing great variability among the different members of the family.[29] In their work,[29] Cohen and Schulten also noted that, despite this multiplicity of ligand migrations pathways that could be observed among globins, some specific positions within the globin fold could actually present a propensity to be located near a ligand passageway. In a previous study on human neuroglobin (Ngb),[30] we showed that the mechanical properties of the residues lying at the border between two internal cavities could be related to the ligand migration pathways that were observed via metadynamics simulations. In a recently published paper on myoglobin (Mgb),[31] Scorciapino et al. identified a set of key residues likely to work as switches regulating ligand migration from one cavity to the other. After noting that these "frontier" residues did occupy similar positions along both the Mgb and Ngb sequences, we sat about investigating their mechanical properties in an extended set of globin chains comprising also cytoglobin (Cgb), truncated hemoglobin (Tr. Hb), and the α and β chains of human hemoglobin (Hb). Although the common general features of globins dynamics have already been studied,[32,33] we chose here to focus on frontier residues, to understand how their mechanical properties can affect ligand migration within the protein cavity network. In this perspective, we use an approach combining all-atom classical molecular dynamics (MD) and coarse-grain Brownian dynamics simulations to draw a complete picture of globins mechanics and show how a limited set of residues might be playing a key role for ligand diffusion in the protein matrix.

MATERIALS AND METHODS

The starting coordinates employed for the simulations were taken from the experimental X-ray structure of each globin. In the case of human Ngb, we took the B chain of the 1OJ6[34] PDB file (with 1.95 Å resolution), and we performed three mutations in silico (G46C, S55C, and S120C) to retrieve the wild-type cysteines, which are not present in the crystal. For the other globins, we chose the following PDB entries:

Received: March 22, 2011
Published: May 09, 2011

dx.doi.org/10.1021/ja202587a | J. Am. Chem. Soc. 2011, 133, 8753–8761

Table 1. Clustering of the 17 501 Conformers Obtained for Each of the 35 ns MD Simulations of Six Representative Globins

Human Neuroglobin (10J6, B Chain), 151 Residues				
NGB0	NGB1	NGB2	NGB3	NGB4
20%	15%	4%	16%	44%
Myoglobin (1YMB), 153 Residues				
MGB0	MGB1	MGB2	MGB3	MGB4
53%	18%	17%	11%	2%
Truncated Hemoglobin (1IDR, B Chain), 126 Residues				
THB0	THB1	THB2	THB3	THB4
45%	17%	16%	15%	7%
Cytoglobin (2DC3, A Chain), 155 Residues				
CGB0	CGB1	CGB2	CGB3	CGB4
16%	one conformer	44%	13%	27%
αHemoglobin (2HHB, A Chain), 141 Residues				
AHB0	AHB1	AHB2	AHB3	AHB4
31%	14%	16%	30%	9%
βHemoglobin (2HHB, B Chain), 146 Residues				
BHB0	BHB1	BHB2	BHB3	BHB4
51%	12%	37%	one conformer	one conformer

Horse heart Mgb from 1YMB[35] at 2.8 Å resolution; Human Cgb from 2DC3[36] at 1.68 Å resolution (A chain); Tr. Hb of *Mycobacterium tuberculosis* from 1IDR[18] at 1.9 Å resolution (B chain); and human Hb from 2HHB[37] at 1.74 Å resolution (chains A for αHb and B for βHb).

Classical Molecular Dynamics. MD simulations were performed with the Gromacs[38−40] software package using the OPLS all atoms force field.[41] Quantum chemical calculations with Gaussian[42] were performed to determine the charges of the hexacoordinated heme group (Ngb, Cytg, TrHb) and pentacoordinated heme group (Mgb, αHb, βHb) using B3LYP[43] and the 6-31G* basis set. The other force field parameters for the prosthetic group were taken from previous studies done on Mgb.[44] The protein was solvated in a cubic box of side length 78 Å, using periodic boundary conditions, with explicit single-point charge water molecules.[45] When necessary, Na^+ ions (from two to six) were added to neutralize the system, which contained between 47 000 and 52 000 atoms depending on the globin under study. All simulations were performed at 1 atm and 300 K, maintained with the Berendsen barostat and thermostat.[46] Long-range electrostatic interactions were treated using the Particle Mesh Ewald (PME) method,[47] with a grid spacing of 0.12 nm and a nonbond pair list cutoff of 9.0 Å with an updating of the pair list every five steps. We could choose a time step of 2 fs by constraining bond lengths involving H atoms with the LINCS algorithm.[48] The solvent was first relaxed by an energy minimization, which was followed by a 100 ps equilibration step under restraint, and then heated slowly until 300 K; 50 ns production runs were eventually performed from which the last 35 ns were kept for analysis. The g_cluster algorithm from the Gromacs suite was then used to obtain five representative structures for each globin over the simulation production period (see Table 1). We used the single linkage method, where a new structure is added to the cluster when the distance between two conformations is less than a chosen cutoff, and employed a different clustering cutoff for each globin, depending on the weight of the system. For Mgb we used a 0.0788 cutoff; 0.078 nm for Ngb; 0.0763 nm for Cgb; 0.0802 nm for TrHb; 0.0777 nm for αHb; and 0.0795 nm for βHb. The 30 resulting structures (listed in Table 1) are identified via a three-letter code indicating the original globin (MGB, NGB, CGB, THB, AHB, or BHB) and a number (from 0 to 4) for the cluster. MGB0, for example, corresponds to the first clusterized structure obtained for horse heart Mgb.

Finally, the online software Pocket-Finder (http://www.modelling.leeds.ac.uk/pocketfinder/)[49] was used for detecting cavities in the various globin structures that were produced and calculating their volumes. These calculations were performed on the clusterized structures with their prosthetic group but in the absence of ligand.

Brownian Dynamics Simulations. BD simulations have been carried out on the globins clusterized structures using the ProPHet (Probing Protein Heterogeneity) program.[50−52] The simulations used a coarse-grained protein model, in which each amino acid is represented by one pseudoatom located at the Cα position, and either one or two (for larger residues) pseudoatoms replacing the side chain (with the exception of Gly).[53] Interactions between the pseudoatoms are treated according to the standard elastic network model;[54] that is, all pseudoatoms lying closer than 9 Å are joined with quadratic springs having the same force constant of 0.6 kcal mol^{-1} $Å^{-2}$. Springs are assumed to be relaxed in the reference conformation of the protein, derived either from the crystallographic data or from the clusterized structures produced by the MD simulations. Following earlier studies, which showed how ligands as large as a heme group actually had little influence on calculated force constants,[50,51] we chose not to include the prosthetic group in the protein representation. The simulations use an implicit solvent representation via the diffusion and random displacement terms in the equation of motion,[55] and hydrodynamic interactions are included through the diffusion tensor.[56]

From the positional fluctuations resulting from BD simulations, carried out for 100 000 steps at a temperature of 300 K, effective force constants for displacing each particle i are calculated as

$$k_i = \frac{3k_B T}{\langle (d_i - \langle d_i \rangle)^2 \rangle} \quad (1)$$

where the brackets indicate an average taken over the whole simulation, k_B is the Boltzmann constant, and d_i is the average distance of particle i from the other particles j in the protein, excluding the pseudoatoms, which belong to the same residue m to which particle i belongs. Also, the distances between the Cα pseudoatom of residue m and the Cα pseudoatoms of the adjacent residues $m + 1$ and $m - 1$ are not included in the average. The force constant associated with each residue m is taken to be the average of the force constants calculated according to eq 1 for each of the pseudoatoms i forming this residue. Within this framework, the mechanical properties of the protein are described at the residue level by its "rigidity profile", that is, by the ordered sequence of the force constants calculated for each residue.

RESULTS

Globins Cavity Network. For the six studied globin chains, the clusterized structures were analyzed with the Pocket-Finder program, and the 10 main cavities detected in each of these structures are listed with their lining residues in Supporting Information Tables 1−6. Similarly to what we observed in our previous work on human neuroglobin,[30] the cavity network of each protein can show considerable reorganization from one cluster to the other, thus inducing large variations of the total volume of the cavities, which can range from 289 to 543 $Å^3$ for βHb and from 558 to 913 $Å^3$ for Mgb, if we take the two globins presenting the most extreme volume variations. A number of

these inner pockets can be related to the xenon cavities (Xe1, Xe2, Xe3, and Xe4) and the distal pocket (DP) that have been observed experimentally in sperm whale Mgb,[12] the phantom 1 cavity that was detected in the same protein by MD simulations,[21,22] the site 1 that was observed in a Xe adduct of human Hb,[57] or the numerous ligand exit pathways that could be identified via MD simulations on Mgb,[24−27] Hbs,[18,23] or Ngb;[30,34,58] see Figure 1 for a typical representation of the cavity network and its fluctuations in Mgb.

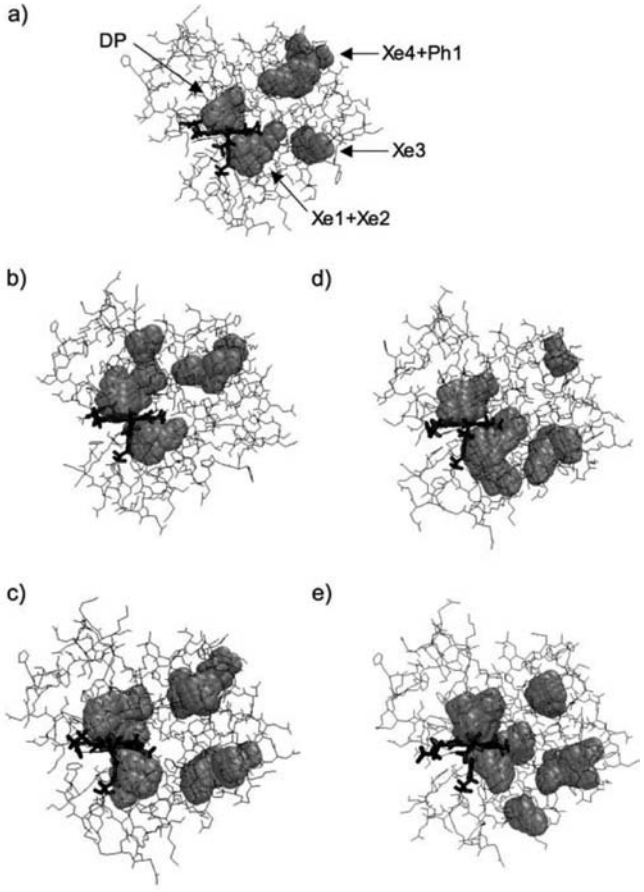


Figure 1. Representation of the five main cavities in the clusterized structures of horse-heart Mgb as detected by Pocket-Finder. (a) MGBO with arrows pointing to the standard Xe and Ph1 cavities, (b) MGB1, (c) MGB2, (d) MGB3, and (e) MGB4. This figure and Figures 5 and 6 were prepared using Visual Molecular Dynamics.[85]

From the list of the cavity lining residues, we could define two subgroups of what we call frontier residues (FR), that is, residues lining two or more internal pockets in the protein. For a given globin, transient frontier residues (TFR) are located at the border between two cavities in only one of the five clusterized structures, while recurrent frontier residues (RFR) can be found in at least two of the clusterized structures; both groups are listed in tTable 2.

Globins Mechanical Properties. The force constant profiles obtained for the main structural cluster of each protein are plotted in Figure 2. Similar to what has been observed in our previous studies on hemoproteins[50] and neuroglobin,[30] the analogous aspect of the profiles reflects the α-helical globin fold, with α-helices appearing as grouped rigidity peaks along the protein sequence (see the shaded areas in Figure 2a) and flexible regions between, denoting in particular the CD and EF loops. In their work of 1999 made on 728 sequences of different globin subfamilies, Ptitsyn and Ting[59] identified 13 conserved heme-binding residues. It turns out 12 out of these 13 residues (which are indicated by empty squares in Figure 2a) actually correspond to local peaks in the proteins rigidity profiles. This suggests how important the tight binding of the prosthetic group is for the biological activity of the protein.[51] Likewise, the five residues forming the folding nucleus of globins (indicated by "●" in Figure 2a) correspond to rigidity peaks, thus underlying the strong correspondence between a protein mechanics and its functional and structural properties.

Even though the five clusterized structures obtained for each globin do not present important variations, with Cα rmsd's between two conformations that are always inferior to 2 Å and with an average value of 1.2 Å, these small structural changes are nonetheless sufficient to induce noticeable variations in the mechanical properties of a limited number of residues in the six globin chains. For every studied protein, we made a pairwise comparison of all five rigidity profiles, and for each residue we kept the maximum value that could be observed for its force constant variation. The resulting max(Δk) profiles are plotted in Figure 3. We then defined as "mechanically sensitive" (MS) those residues presenting a max(Δk) value over a given threshold of 10 kcal mol^{-1} Å^{-2} for Ngb, Mgb, Cgb, and αHb, 7 kcal mol^{-1} Å^{-2} for Tr. Hb, and 20 kcal mol^{-1} Å^{-2} for βHb. This procedure led to the selection of 8−14 residues for each globin that are listed in

Table 2. List of the Frontier Residues (Lining Two or More Internal Cavities), Which Were Obtained via Pocket-Finder for Each Globin[a]

Mgb	transient	10/14/28/29/39/42/46/61/66/78/81/82/86/89/93/97/100/101/11/115/118/123/142
	recurrent	17/21/25/43/64/65/69/72/75/76/77/86/99/104/107/138/146
Ngb	transient	41/42/71/82/99/102/103/105/111/147
	recurrent	27/28/38/68/72/75/85/89/92/95/96/101/106/109/110/113/133/136/137/140/144
Cgb	transient	30/41/42/45/49/84/88/93/135/143/156
	recurrent	31/34/56/60/81/85/86/92/102/106/109/124/127/128/131/134/151/154/157/158/161
Tr. Hb	transient	16/22/29/36/53/54/65/72/84/95/126
	recurrent	19/25/32/33/46/58/61/63/66/77/80/86/94/98/102/115/116/119/122
αHb	transient	21/25/30/58/95/102/117/121/130/132
	recurrent	14/17/24/29/33/43/48/55/62/63/66/101/105/106/109/117/125/129/133
βHb	transient	11/24/25/30/31/33/35/45/48/54/63/72/75/76/81/84/98/103/106/107/110/114/134/137/139/140
	recurrent	15/23/26/28/32/42/60/67/68/71/78/85/130

[a] These can be either transient (appearing in only one of the clusterized structures) or recurrent (present in two or more of the clusterized structures).

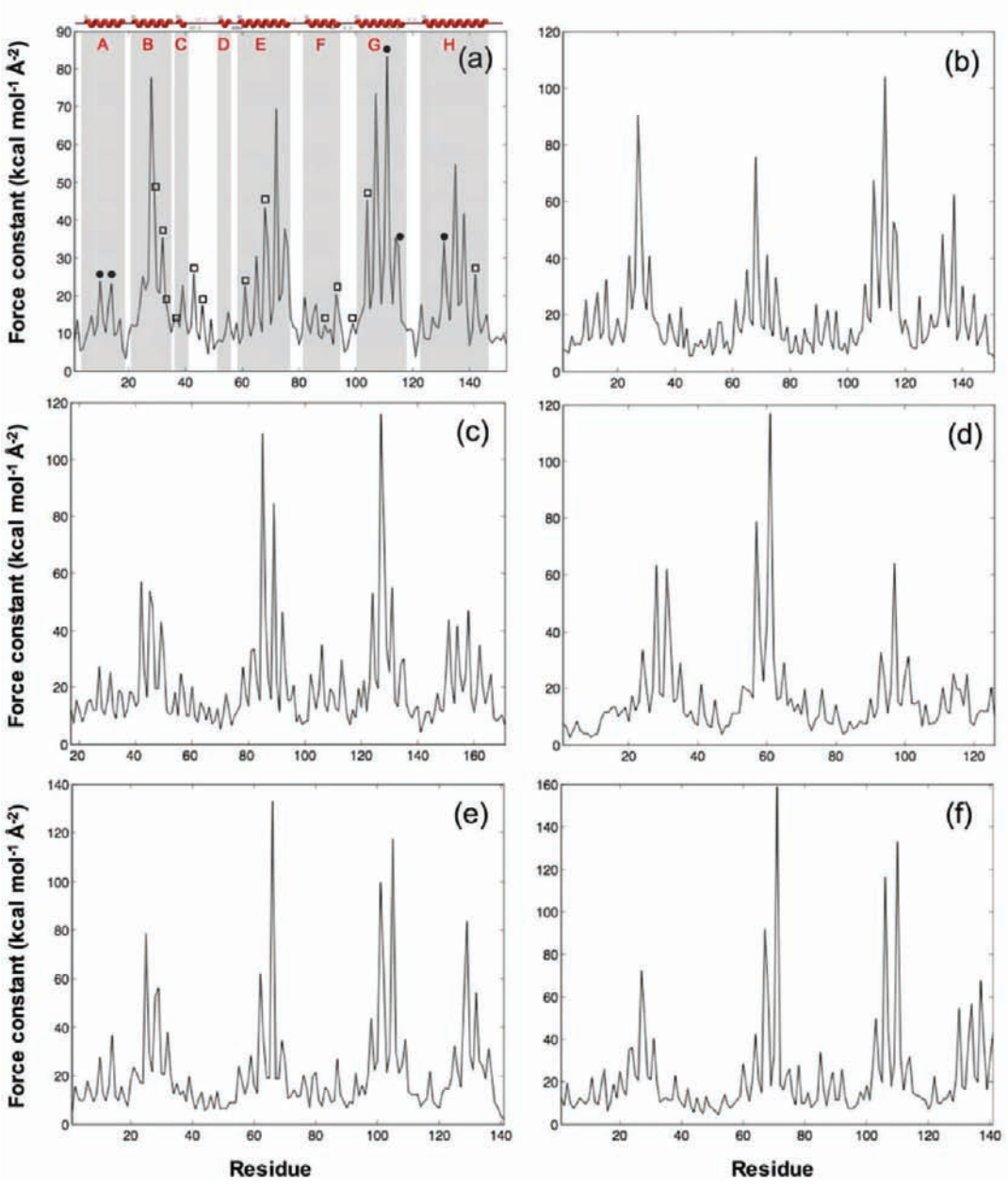


Figure 2. Rigidity profiles (in kcal mol^{-1} Å^{-2}) of the main cluster for the six globin chains under study. (a) Mgb, (b) Ngb, (c) Cgb, (d) Tr. Hb, (e) αHb, (f) βHb. In (a), the areas shaded in gray correspond to α-helices, as indicated by the red secondary structure plot at the top of the structure, the "□" indicate heme-binding conserved residues of globins (from left to right: Leu29-B10, Leu32-B13, Phe33-B14, Pro37-C2, Phe43-CD1, Phe46-CD4, Leu61-E4, Val68-E11, Leu89-F4, His93-F8, Ile99-FG5, Leu104-G5, and Ile142-H19), and the "●" indicate the conserved folding nucleus (from left to right: Val10-A8, Trp14-A12, Ile11-G12, Leu115-G16, and Met131-H8).

Table 3. Interestingly, these MS residues, which represent a subset of the rigid residues from the original rigidity profiles, systematically correspond to frontier residues in Mgb and Cgb. In the case of Ngb, Tr. Hb, αHb, and βHb, the few MS residues that are not frontier residues are nonetheless cavity lining residues, with the only exception of Ser112-G11 in Ngb.

Conservation of the Mechanical Properties along the Sequence. We used the clustalw[60] web server to align the sequences of the six globin chains under study. Despite the high mechanical similarity that could be observed in the rigidity profiles of Figure 2, theses sequences present relatively low identities, ranging from 17% to 50% (see Supporting Information Table 7). The multiple sequence alignment is presented in Figure 4 with the positions that are occupied by MS residues highlighted in red. We can see that most of these positions are indeed common to one or more globins, with the particular case of positions G8 and G12, which correspond to MS residues in all six chains. Two other positions that are extremely well conserved in terms of mechanical properties are E11, which presents a MS residue in all chains but Mgb, and E15, where Ngb is the only chain not showing a MS residue. However, the max(Δk) value of Val68-E11 in Mgb is actually right under the chosen cutoff with 9.14 kcal mol^{-1} Å^{-2}, which means that this residue does actually present mechanical sensitivity. In the case of Ile72-E15 of Ngb,

dx.doi.org/10.1021/ja202587a |J. Am. Chem. Soc. 2011, 133, 8753–8761

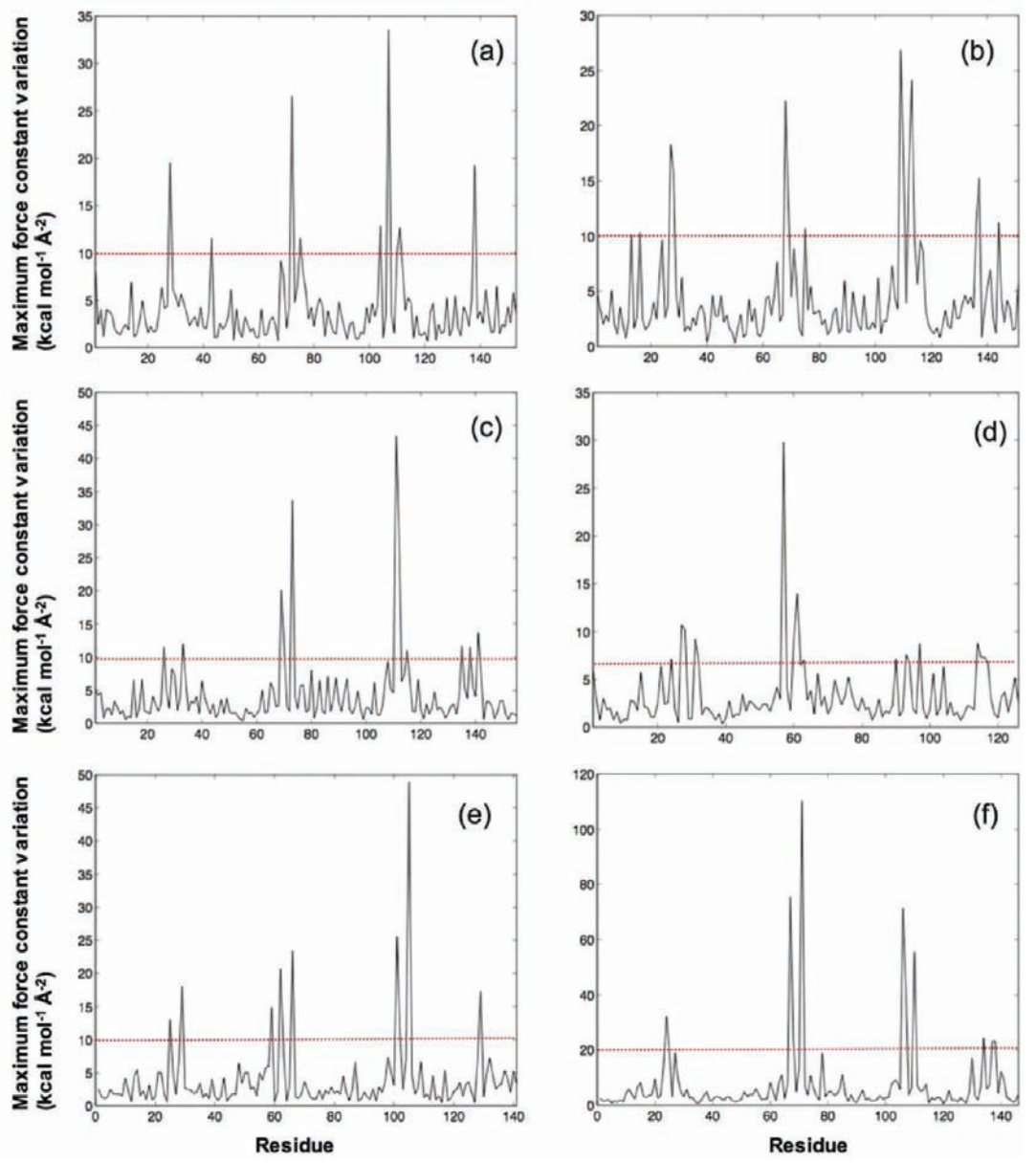


Figure 3. Maximum variation (in kcal mol^{-1} Å^{-2}) of the force constant upon changing the globin structure. (a) Mgb, (b) Ngb, (c) Cgb, (d) Tr. Hb, (e) αHb, (f) βHb. The red horizontal dotted line indicates the threshold value chosen for the selection of mechanically sensitive residues that are listed in Table 2.

Table 3. List of the Mechanically Sensitive Residues in Each of the Six Globins with Their Position along the Protein Fold[a]

Mgb	Val28-B9, Phe43-CD1, Leu72-E15, Ile75-E18, Leu104-G5, **Ile107-G8**, **Ile111-G12**, Phe138-H15
Ngb	Trp13-A12, Val16-A15, Leu27-B9, Phe28-B10, Val68-E11, Met69-E12, Ala75-E18, **Val109-G8**, Gly110-G9, Ser112-G11, **Leu113-G12**, Leu136-H11, Tyr137-H12, Met144-H19
Cgb	Gly42-B6, Phe49-B13, Val85-E11, Met86-E12, Leu89-E15, **Leu127-G8**, Ser128-G9, **Ile131-G12**, Trp151-H8, Leu154-H11
Tr. Hb	Ile25-B6, Val28-B9, Val29-B10, Phe32-B13, Gln58-E11, Phe61-E14, Phe62-E15, Ala64-E17, Phe91-G5, **Val94-G8**, **Leu98-G12**, Ile115-H8, Leu116-H9, Gly117-H10
αHb	Gly25-B6, Leu29-B10, Gly59-E8, Val62-E11, Leu66-E15, **Leu101-G8**, Ser102-G9, **Leu105-G12**, Leu129-H12
βHb	Gly24-B6, Val67-E11, Leu68-E12, Phe75-E15, **Leu106-G8**, Gly107-G9, **Leu110-G12**, Val134-H12, Val137-H15, Ala138-H16

[a] Positions common to more than one protein are underlined; positions common to all proteins are in bold.

this residue did show some specific mechanical properties in our previous work on human Ngb,[30] where we investigated its mechanical variations upon formation of an internal disulfide bond in the pentacoordinated state of the protein (which is the

dx.doi.org/10.1021/ja202587a |J. Am. Chem. Soc. 2011, 133, 8753–8761

```
                  1234567890123456 12345678901234567123456          1234561
                  AAAAAAAAAAAAAAAA BBBBBBBBBBBBBBBBBCCCCCC          DDDDDDE
1YMB_A      -GLSDGEWQQVLNVWGKVEADIAGHGQEVLIRLFTGHPETLEKFDKFK-HLKTEAEMKAS 58
1OJ6_B      --MERPEPELIRQSWRAVSRSPLEHGTVLFARLFALEPDLLPLFQYNCRQFSSPEDCLSS 58
2DC3_A      EELSEAERKAVQAMWARLYANCEDVGVAILVRFFVNFPSAKQYFSQFK-HMEDPLEMERS 59
1IDR_B      GLLSRLRKREPISIYDKIG--GHEAIEVVVEDFYVRVLADDQLSAFFS----------G 47
2HHB_A      -VLSPADKTNVKAAWGKVGAHAGEYGAEALERMFLSFPTTKTYFPHF--DLSH-----GS 52
2HHB_B      VHLTPEEKSAVTALWGKV--NVDEVGGEALGRLLVVYPWTQRFFESFG-DLSTPDAVMGN 57

                  2345678901234567890       1234567890123456  1234567890123456
                  EEEEEEEEEEEEEEEEEEE       FFFFFFFFFFFFFFFF  GGGGGGGGGGGGGGGG
1YMB_A      EDLKKHGTVVLTALGGILKKKGH---HEAELKPLAQSHATKHKIPIKYLEFISDAIIHVL 115
1OJ6_B      PEFLDHIRKVMLVIDAAVTNVEDLSSLEEYLASLGRKHR-AVGVKLSSFSTVGESLLYML 117
2DC3_A      PQLRKHACRVMGALNTVVENLHDPDKVSSVLALVGKAHALKHKVEPVYFKILSGVILEVV 119
1IDR_B      TNMSRLKGKQVEFFAAALGGPEP-----YTGAPMKQVHQ-GRGITMHHFSLVAGHLADAL 101
2HHB_A      AQVKGHGKKVADALTNAVAHVDD---MPNALSALSDLHAHKLRVDPVNFKLLSHCLLVTL 109
2HHB_B      PKVKAHGKKVLGAFSDGLAHLDN---LKGTFATLSELHCDKLHVDPENFRLLGNVLVCVL 114

                  7890     12345678901234567890123456
                  GGGG     HHHHHHHHHHHHHHHHHHHHHHHHHH
1YMB_A      HSKHPGDFGADAQGAMTKALELFRNDIAAKYKELGFQG 153
1OJ6_B      EKCLGPAFTPATRAAWSQLYGAVVQAMSRGWDGE---- 151
2DC3_A      AEEFASDFPPETQRAWAKLRGLIYSHVTAAYKEVGW-- 155
1IDR_B      T---AAGVPSETITEILGVIAPLAVDVT---------- 126
2HHB_A      AAHLPAEFTPAVHASLDKFLASVSTVLTSKYR------ 141
2HHB_B      AHHFGKEFTPPVQAAYQKVVAGVANALAHKYH------ 146
```

Figure 4. Alignment of the six globin sequences. The first column in each block displays the PDB code and chain of the protein, and the last column shows the number of residues up to that line. Green annotations indicate the positions of the α-helices along the Mgb sequence, while mechanically sensitive residues are highlighted in red.

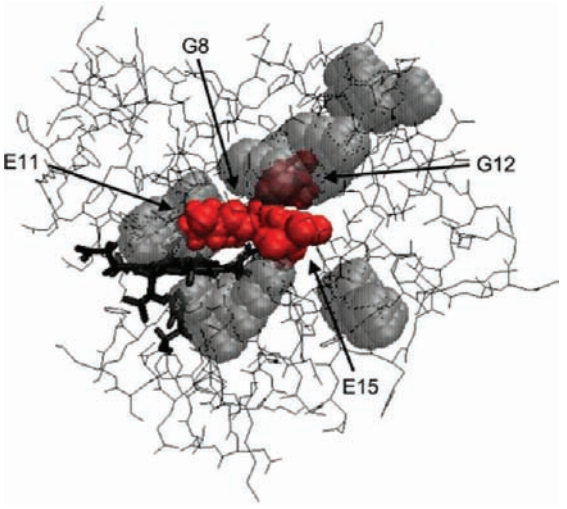


Figure 5. Conserved mechanical nucleus formed by positions E11, E15, G8, and G12 (in red) at the heart of the MGB0 structure.

ligand binding state), while the simulations in the current work were carried out on the hexacoordinated state of human Ngb. So eventually, we can select positions E11, E15, G8, and G12 along the globin sequence to define a conserved mechanical nucleus that appears to form a central gate system right at the heart of the globin fold and at the frontier of the DP, Xe2, and Xe4 cavities; see Figure 5.

■ DISCUSSION

The six globin chains in this study presented very similar rigidity profiles, thus reflecting the conservation of proteins dynamics within a structural family.[33,61,62] More interestingly, the variations of these profiles due to the proteins structural fluctuations are also comparable and allow us to select a restricted set of residues occupying that we called "mechanically sensitive" positions. A search in the literature shows that most of the positions bearing that label had already been highlighted as corresponding to cavity lining of frontier residues in numerous experimental or theoretical works on Mgb,[21,22,26,27,63] Ngb,[30,64−69] Cgb,[70−72] Tr. Hb,[18,73,74] and human Hb.[23] Interestingly, our results concur with data obtained from molecular dynamics performed using not only the OPLS force field like us,[27] but also Amber 95[31] and 99[73] or Charmm 22[23,29] and 27,[72] thus showing the robustness of molecular simulation studies for the investigation of protein properties.

The positions of the mechanical nucleus residues in particular have been shown to play an important role for ligand diffusion in various globins. For example, in the case of Mgb, several mutational studies focused on the importance of positions E11 and G8,[75−77] showing how the replacement of the isoleucine in G8 does not modify the protein's structure, but substantially affects ligand binding. In Scorciapino et al.'s work on Mgb breathing motions,[31] all four positions E11, E15, G8, and G12 appear in the central gate area between cavities DP, Xe4, and Xe2. In Tr. Hb, ligand migration along the protein's internal tunnel is thought to be regulated by residues Gln58-E11 and Phe62-E15, with both side-chains acting as gate-opening molecular switches.[78−81] For human Hb, gating movements of the leucine residue in G12 govern the hopping of gaseous ligand from and to different binding sites.[57]

As we have already seen, most MS residues are also frontier residues adjacent to two or more of the internal cavities that were detected in the various structures produced during our MD simulations. If we look more precisely into the structural rearrangement of our globins inner pockets, it appears clearly that the mechanical variations of the proteins are closely related to the cavity network fluctuations. As an example, we superimposed in Figure 6a and b the five main cavities of Tr. Hb in its THB2 (in blue) and THB4 (in yellow) conformations. The five residues undergoing the most important variations of their force

dx.doi.org/10.1021/ja202587a |J. Am. Chem. Soc. 2011, 133, 8753–8761

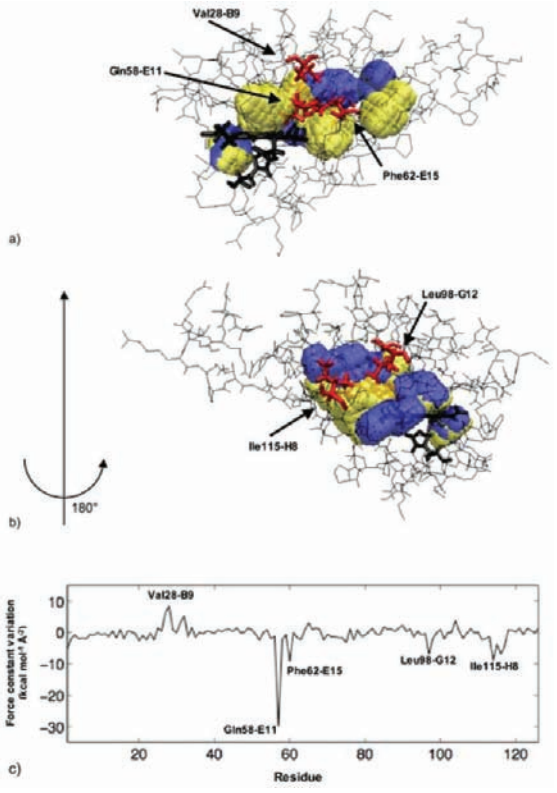


Figure 6. Superposition of the cavity networks from the THB2 (in blue) and THB4 (in yellow) structures from Tr. Hb. The residues undergoing the most important mechanical perturbations upon the structural transition are plotted in red. (a) Front view, (b) back view. (c) Variation of the force constants upon transition from the THB2 to THB4 structures from Tr. Hb.

constant upon the THB2 → THB4 transition are signaled in Figure 6c. Among these are positions E11, E15 (whose importance we underlined in the previous paragraph), and G8 from the mechanical nucleus. Gln58-E11 in particular shows a remarkable decrease of its rigidity (~ -30 kcal mol^{-1} Å^{-2}), which can be related to its central position in the globin's structure. As we can see in Figure 6a and b, the five MS residues from Figure 6c, which are drawn here in red, tightly surround the internal cavities of Tr. Hb. Gln58-E11 and Phe62-E15 lie right at the frontiers between three successive pockets leading to the prosthetic heme group that are found in the THB4 structure (in yellow). Hence, for the ligand to access the heme binding site by diffusing along the protein's cavity network, it is essential for the side-chains of these residues to show some flexibility.

CONCLUDING REMARKS

In our previous works on protein mechanics, we did compare rigidity profiles for various protein oxidation or coordination states and were able to relate residues mechanical properties to their role in the protein's functional activity.[30,50,51] Here, we used classical MD simulations to produce several representative clusterized structures for a single protein state, and the mechanical properties of each structure were then studied via coarse-grain Brownian dynamics. By comparing the rigidity profiles of the clusterized structures, we show that these mechanical properties do have a dynamic quality. While the rigidity profile of a protein remains qualitatively the same along time, with its main peaks associated with a given set of amino acids, it can nonetheless present noticeable variation from one structure to the other for a limited number of residues. In the case of globins, the resulting "mechanically sensitive" residues are connected with the breathing motions of the protein and the fluctuations of its internal cavity network. We also note that these residues positions are well conserved along the protein's sequence. In particular, we could identify what we called a mechanical nucleus, formed by positions E11, E15, G8, and G12. Residues occupying these positions have already been shown separately to play a key role for ligand diffusion in Mgb, Tr. Hb, and human Hb using various experimental and theoretical approaches. Here, we suggest that this quartet might actually be essential for the regulation of ligand migration within the cavity network all throughout the whole protein family. More generally, our findings are of interest for the study of the numerous globular proteins that possess internal cavities and channels, such as redox enzymes.[82−84] From a protein engineering perspective, the study of their mechanical properties should bring us valuable information regarding the key residues that could represent potential mutation targets to modulate or improve their enzymatic activity.

ASSOCIATED CONTENT

Supporting Information. Supplementary Tables S1−6 summarizing the 10 main cavities with their volume and lining residues for each of the 30 globin structures (five for each globin chain) that were produced during this study. Supplementary Table 7 giving the percentage of sequence identity for the six globin chains. Complete refs 42 and 84. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
sacquin@ibpc.fr

REFERENCES

(1) Hardison, R. C. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 5675–5679.
(2) Vinogradov, S. N.; Hoogewijs, D.; Bailly, X.; Arredondo-Peter, R.; Gough, J.; Dewilde, S.; Moens, L.; Vanfleteren, J. R. *BMC Evol. Biol.* **2006**, *6*, 31–67.
(3) Vinogradov, S. N.; Hoogewijs, D.; Bailly, X.; Mizuguchi, K.; Dewilde, S.; Moens, L.; Vanfleteren, J. R. *Gene* **2007**, *398*, 132–142.
(4) Vinogradov, S. N.; Moens, L. *J. Biol. Chem.* **2008**, *283*, 8773–8777.
(5) Kakar, S.; Hoffman, F. G.; Storz, J. F.; Fabian, M.; Hargrove, M. S. *Biophys. Chem.* **2010**, *152*, 1–14.
(6) Wajcman, H.; Kiger, L.; Marden, M. C. *C. R. Biol.* **2009**, *332*, 273–282.
(7) Kendrew, J. C.; Dickerson, R. E.; Strandberg, B. E.; Hart, R. G.; Davies, D. R.; Phillips, D. C.; Shore, V. C. *Nature* **1960**, *185*, 422–427.
(8) Bettati, S.; Viappiani, C.; Mozzarelli, A. *Biochim. Biophys. Acta* **2009**, *1794*, 1317–1324.
(9) Frauenfelder, H. *Chem. Phys.* **2010**, *375*, 612–615.
(10) Nienhaus, K.; Nienhaus, G. U. *IUMB Life* **2007**, *59*, 490–497.
(11) Burmester, T.; Hankeln, T. *J. Exp. Biol.* **2009**, *212*, 1423–1428.
(12) Tilton, R. F.; Kuntz, I. D.; Petsko, G. A. *Biochemistry* **1984**, *23*, 2849–2857.
(13) Brunori, M.; Gibson, Q. H. *EMBO Rep.* **2001**, *2*, 674–679.
(14) Hubbard, S. J.; Gross, K. H.; Argos, P. *Protein Eng.* **1994**, *7*, 613–626.

(15) Carugo, O.; Argos, P. *Proteins* **1998**, *31*, 201–213.
(16) Tomita, A.; Kreutzer, U.; Adachi, S.; Koshihara, S.; Jue, T. *J. Exp. Biol.* **2010**, *213*, 2748–2754.
(17) Case, D. A.; Karplus, M. *J. Mol. Biol.* **1979**, *132*, 343–368.
(18) Milani, M.; Pesce, A.; Ouellet, Y.; Ascenzi, P.; Guertin, M.; Bolognesi, M. *EMBO J.* **2001**, *20*, 3902–3909.
(19) Bourgeois, D.; Vallone, B.; Schotte, F.; Arcovito, A.; Miele, A. E.; Sciara, G.; Wulff, M.; Anfinrud, P.; Brunori, M. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 8704–8709.
(20) Schotte, F.; Lim, M. H.; Jackson, T. A.; Smirnov, A. V.; Soman, J.; Olson, J. S.; Phillips, G. N.; Wulff, M.; Anfinrud, P. A. *Science* **2003**, *300*, 1944–1947.
(21) Bossa, C.; Anselmi, M.; Roccatano, D.; Amadei, A.; Vallone, B.; Brunori, M.; Di Nola, A. *Biophys. J.* **2004**, *86*, 3855–3862.
(22) Bossa, C.; Amadei, A.; Daidone, I.; Anselmi, M.; Vallone, B.; Brunori, M.; Di Nola, A. *Biophys. J.* **2005**, *89*, 465–474.
(23) Mouawad, L.; Marechal, J. D.; Perahia, D. *Biochim. Biophys. Acta* **2005**, *1724*, 385–393.
(24) Cohen, J.; Arkhipov, A.; Braun, R.; Schulten, K. *Biophys. J.* **2006**, *91*, 1844–1857.
(25) Ceccarelli, M.; Anedda, R.; Casu, M.; Ruggerone, P. *Proteins* **2008**, *71*, 1231–1236.
(26) Nishihara, Y.; Hayashi, S.; Kato, S. *Chem. Phys. Lett.* **2008**, *464*, 220–225.
(27) Elber, R.; Gibson, Q. H. *J. Phys. Chem. B* **2008**, *112*, 6147–6154.
(28) Elber, R. *Curr. Opin. Struct. Biol.* **2010**, *20*, 162–167.
(29) Cohen, J.; Schulten, K. *Biophys. J.* **2007**, *93*, 3591–3600.
(30) Bocahut, A.; Bernad, S.; Sebban, P.; Sacquin-Mora, S. *J. Phys. Chem. B* **2009**, *113*, 16257–16267.
(31) Scorciapino, M. A.; Robertazzi, A.; Casu, M.; Ruggerone, P.; Ceccarelli, M. *J. Am. Chem. Soc.* **2009**, *131*, 11825–11832.
(32) Maguid, S.; Fernandez-Alberti, S.; Ferrelli, L.; Echave, J. *Biophys. J.* **2005**, *89*, 3–13.
(33) Laberge, M.; Yonetani, T. *IUBMB Life* **2007**, *59*, 528–534.
(34) Pesce, A.; Dewilde, S.; Nardini, M.; Moens, L.; Ascenzi, P.; Hankeln, T.; Burmester, T.; Bolognesi, M. *Structure* **2003**, *11*, 1087–1095.
(35) Evans, S. V.; Brayer, G. D. *J. Mol. Biol.* **1990**, *213*, 885–897.
(36) Makino, M.; Sugimoto, H.; Sawai, H.; Kawada, N.; Yoshizato, K.; Shiro, Y. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2006**, *62*, 671–677.
(37) Fermi, G.; Perutz, M. F.; Shaanan, B.; Fourme, R. *J. Mol. Biol.* **1984**, *175*, 159–174.
(38) Berendsen, H. J. C.; Vanderspoel, D.; Vandrunen, R. *Comput. Phys. Commun.* **1995**, *91*, 43–56.
(39) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306–317.
(40) Van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comput. Chem.* **2005**, *26*, 1701–1718.
(41) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105*, 6474–6487.
(42) Frisch, M. J.; et al. *Gaussian 03*; Gaussian, Inc.: Wallingford, CT, 2003.
(43) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
(44) Li, H. Y.; Elber, R.; Straub, J. E. *J. Biol. Chem.* **1993**, *268*, 17908–17916.
(45) Miyamoto, S.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 952–962.
(46) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
(47) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
(48) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
(49) Hendlich, M.; Rippmann, F.; Barnickel, G. *J. Mol. Graphics Modell.* **1997**, *15*, 359.
(50) Sacquin-Mora, S.; Lavery, R. *Biophys. J.* **2006**, *90*, 2706–2717.
(51) Sacquin-Mora, S.; Laforet, E.; Lavery, R. *Proteins* **2007**, *67*, 350–359.
(52) Lavery, R.; Sacquin-Mora, S. *J. Biosci.* **2007**, *32*, 891–898.
(53) Zacharias, M. *Protein Sci.* **2003**, *12*, 1271–82.
(54) Tozzini, V. *Curr. Opin. Struct. Biol.* **2005**, *15*, 144–50.
(55) Ermak, D. L.; McCammon, J. A. *J. Chem. Phys.* **1978**, *69*, 1352–1360.
(56) Pastor, R. W.; Venable, R.; Karplus, M. *J. Chem. Phys.* **1988**, *89*, 1112–1127.
(57) Savino, C.; Miele, A. E.; Draghi, F.; Johnson, K. A.; Sciara, G.; Brunori, M.; Vallone, B. *Biopolymers* **2009**, *91*, 1097–1107.
(58) Vallone, B.; Nienhaus, K.; Brunori, M.; Nienhaus, G. U. *Proteins* **2004**, *56*, 85–92.
(59) Ptitsyn, O. B.; Ting, K. L. *J. Mol. Biol.* **1999**, *291*, 671–82.
(60) Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G. *Bioinformatics* **2007**, *23*, 2947–2948.
(61) Maguid, S.; Fernandez-Alberti, S.; Parisi, G.; Echave, J. *J. Mol. Evol.* **2006**, *63*, 448–457.
(62) Hollup, S. M.; Fuglebakk, E.; Taylor, W. R.; Reuter, N. *Protein Sci.* **2010**, *20*, 197–209.
(63) Olson, J. S.; Soman, J.; Phillips, G. N. *IUBMB Life* **2007**, *59*, 552–562.
(64) Vallone, B.; Nienhaus, K.; Matthes, A.; Brunori, M.; Nienhaus, G. U. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17351–17356.
(65) Anselmi, M.; Brunori, M.; Vallone, B.; Di Nola, A. *Biophys. J.* **2007**, *93*, 434–441.
(66) Lutz, S.; Nienhaus, K.; Nienhaus, G. U.; Meuwly, M. *J. Phys. Chem. B* **2009**, *113*, 15334–15343.
(67) Abbruzzetti, S.; Faggiano, S.; Bruno, S.; Spyrakis, F.; Mozzarelli, A.; Dewilde, S.; Moens, L.; Viappiani, C. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 18984–18989.
(68) Nienhaus, K.; Lutz, S.; Meuwly, M.; Nienhaus, G. U. *ChemPhysChem* **2010**, *11*, 119–129.
(69) Anselmi, M.; Di Nola, A.; Amadei, A. *J. Phys. Chem. B* **2011**, *115*, 2436–2446.
(70) de Sanctis, D.; Dewilde, S.; Pesce, A.; Moens, L.; Ascenzi, P.; Hankeln, T.; Burmester, T.; Bolognesi, M. *J. Mol. Biol.* **2004**, *336*, 917–927.
(71) de Sanctis, D.; Dewilde, S.; Pesce, A.; Moens, L.; Ascenzi, P.; Hankeln, T.; Burmester, T.; Bolognesi, M. *Biochem. Biophys. Res. Commun.* **2004**, *316*, 1217–1221.
(72) Orlowski, S.; Nowak, W. *BioSystems* **2008**, *94*, 263–266.
(73) Crespo, A.; Marti, M. A.; Kalko, S. G.; Morreale, A.; Orozco, M.; Gelpi, J. L.; Luque, F. J.; Estrin, D. A. *J. Am. Chem. Soc.* **2005**, *127*, 4433–4444.
(74) Golden, S. D.; Olsen, K. W. *Globins and Other Nitric Oxide-Reactive Proteins, Part B*; Elsevier Academic Press, San Diego, CA, 2008; Vol. 437, pp 417−437.
(75) Quillin, M. L.; Li, T. S.; Olson, J. S.; Phillips, G. N.; Dou, Y.; Ikedasaito, M.; Regan, R.; Carlson, M.; Gibson, Q. H.; Li, H. Y.; Elber, R. *J. Mol. Biol.* **1995**, *245*, 416–436.
(76) Ishikawa, H.; Uchida, T.; Takahashi, S.; Ishimori, K.; Morishima, I. *Biophys. J.* **2001**, *80*, 1507–1517.
(77) Dantsker, D.; Roche, C.; Samuni, U.; Blouin, G.; Olson, J. S.; Friedman, J. M. *J. Biol. Chem.* **2005**, *280*, 38740–38755.
(78) Bidon-Chanal, A.; Marti, M. A.; Crespo, A.; Milani, M.; Orozco, M.; Bolognesi, M.; Luque, F. J.; Estrin, D. A. *Proteins* **2006**, *64*, 457–464.
(79) Bidon-Chanal, A.; Marti, M. A.; Estrin, D. A.; Luque, F. J. *J. Am. Chem. Soc.* **2007**, *129*, 6782–6788.
(80) Mishra, S.; Meuwly, M. *Biophys. J.* **2009**, *96*, 2105–2118.
(81) Lama, A.; Pawaria, S.; Bidon-Chanal, A.; Anand, A.; Gelpi, J. L.; Arya, S.; Marti, M.; Estrin, D. A.; Luque, F. J.; Dikshit, K. L. *J. Biol. Chem.* **2009**, *284*, 14457–14468.
(82) Fontecilla-Camps, J. C.; Amara, P.; Cavazza, C.; Nicolet, Y.; Volbeda, A. *Nature* **2009**, *460*, 814–822.

(83) Baron, R.; Riley, C.; Chenprakhon, P.; Thotsaporn, K.; Winter, R. T.; Alfieri, A.; Forneris, F.; van Berkel, W. J. H.; Chaiyen, P.; Fraaije, M. W.; Mattevi, A.; McCammon, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 10603–10608.
(84) Liebgott, P. P.; et al. *Nat. Chem. Biol.* **2010**, *6*, 63–70.
(85) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 27–8.