Investigating the Local Flexibility of Functional Residues in Hemoproteins

Sophie Sacquin-Mora and Richard Lavery
Laboratoire de Biochimie Théorique, UMR 9080 CNRS, Institut de Biologie Physico-Chimique, Paris, France

ABSTRACT It is now widely accepted that protein function depends not only on structure, but also on flexibility. However, the way mechanical properties contribute to catalytic mechanisms remains unclear. Here, we propose a method for investigating local flexibility within protein structures that combines a reduced protein representation with Brownian dynamics simulations. An analysis of residue fluctuations during the dynamics simulation yields a rigidity profile for the protein made up of force constants describing the ease of displacing each residue with respect to the rest of the structure. This approach has been applied to the analysis of a set of hemoproteins, one of the functionally most diverse protein families. Six proteins containing one or two heme groups have been studied, paying particular attention to the mechanical properties of the active-site residues. The calculated rigidity profiles show that active site residues are generally associated with high force constants and thus rigidly held in place. This observation also holds for diheme proteins if their mechanical properties are analyzed domain by domain. We note, however, that residues other than those in the active site can also have high force constants, as in the case of residues belonging to the folding nucleus of *c*-type hemoproteins.

INTRODUCTION

There are now enough protein structures available to obtain a good overall picture of protein structural classes, as analyzed in, for example, the CATH database (1). However, when we pass from static structural descriptions to either mechanical or dynamical analyses, much less data is available and there is no simple way of deducing these properties from the corresponding structure. It has nevertheless been well established that protein fluctuations are essential for biological function (2,3). Proteins clearly need to resist thermal agitation within the cell while allowing for the deformations necessary for recognition, substrate uptake, catalytic activity, etc. At the same time, they have to maintain the structures of the interaction surfaces or active sites sufficiently accurately to remain specific and, in the case of enzymes, to carry out catalysis in a stereospecific or regiospecific manner. Molecular dynamics simulations are one tool used to study the relationship between the structural and the mechanical or dynamical properties of proteins. Although currently limited to the nanosecond timescale, such simulations provide important information on local and intermediate-scale movements (4–6). Data regarding protein flexibility can also be obtained from normal-mode calculations, where low-frequency modes, obtained using either atomic resolution (7–9) or various coarse-grained models (10–13), provide useful information on large-scale collective movements (14–17). Related data can be obtained by comparing different conformations of the same protein (18,19) or by using mean-field protein models (20). Experimentally, crystallographic B-factors provide data on atomic fluctuations (21), whereas other approaches include side-chain order parameters (22–24) or hydrogen exchange data (25) derived from NMR experiments and fluctuations obtained from neutron scattering experiments after selective labeling of parts of the protein (26–28).

In an earlier study, we attempted to complement the techniques cited above by developing an original approach for obtaining mechanical data on a residue-by-residue basis (29). This approach was based on calculating a force constant for displacing the Cα atom of a chosen residue within a protein structure. Calculations were carried out using successive energy minimizations, modifying at each step the mean distance from the probe residue to all other Cα atoms within the protein. We have shown that, in contrast to B-factors, these force constants are not so strongly dominated by the local structure surrounding each residue. By looking at the way changes in Cα-Cα distances were coupled during the probing of a protein, it was possible to define domains on the basis of their mechanical coherence. In addition, since the probing only involved a scalar restraint, the displacement of probed residues provided vectorial information on protein mechanics. In this study, we again use this technique. However, we have been able to replace residue-by-residue energy minimizations with a single Brownian dynamics simulation, thereby considerably speeding up the calculations and allowing larger proteins to be studied.

Our aim in this work is to turn from general to specific properties of the proteins studied. We are particularly interested in investigating the mechanical properties of residues located in the vicinity of active sites. Several recent works have reported that catalytic residues are predominantly associated with low B-factors (21,30,31) and are therefore, on average, less flexible than the other residues. Given that our probing is less dominated by local structure, will this distinction also apply to our calculated force constants?

Submitted September 23, 2005, and accepted for publication December 29, 2005.

Address reprint requests to Richard Lavery, Laboratoire de Biochimie Théorique, UMR 9080 CNRS, Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, 75005 Paris, France. E-mail: rlavery@ibpc.fr.

0006-3495/06/04/2706/12 $2.00

doi: 10.1529/biophysj.105.074997

Rather than studying a large database of proteins and looking at overall trends (30,31), we chose to focus our attention on one of the structurally and functionally most diverse families of proteins, the hemoproteins (32). Using only the simple framework of a heme prosthetic group coupled to a polypeptide backbone, these proteins can perform a wide range of functions, including oxygen transport (hemeglobin and myoglobin), nitric oxide transport (nitrophorin), activation of molecular oxygen for oxygenation of organic substrates (cytochrome P450), catalysis (peroxidases), and electron transport (*c*-type cytochromes). There are four commonly found heme prosthetic groups in biological systems (32), among which heme-*b* and heme-*c* are the most frequent. Heme-*b* (found in globins, cytochrome-*b*, or peroxidases) is held in the protein by axial ligation of the iron to one or two amino acid side chains (see Table 1), whereas heme-*c* (found in *c*-type cytochromes) differs by the covalent attachment to the protein backbone formed between cysteine residues and the porphyrin macrocycle. One might imagine that the residues surrounding, and, in particular, chelating the heme group would be very important within the structure of such proteins, and we will therefore study whether these residues are indeed easy to recognize from a mechanical point of view. The proteins we investigate include members of the heme-*b* and heme-*c* families, containing either one or two heme groups. In most cases, the heme prosthetic groups are integral parts of the protein structure, but we have also studied a heme transport protein that can capture and release a heme group.

MATERIALS AND METHODS

In this section we describe the methodologies developed to study local protein flexibility and their application to a set of six hemoproteins.

Protein representation

Whereas our earlier studies used a reduced representation with a single pseudoatom per residue (29), for this study we chose a more refined model based on Zacharias (33). In this model, each amino acid is represented by one pseudoatom located at the Cα position, and either one or two pseudoatoms representing the side chain (with the exception of Gly). Ala, Ser, Thr, Val, Leu, Ile, Asn, Asp, and Cys have a single pseudoatom located at the geometrical center of the side-chain heavy atoms. For the remaining amino acids, a first pseudoatom is located midway between the Cβ and Cγ atoms, whereas the second is placed at the geometrical center of the remaining side-chain heavy atoms. This description, which allows different amino acids to be distinguished from one another, has already proven useful in protein-protein docking (33–35). In this case, it notably allowed side-chain flexibility to be modeled at a much lower cost than in all-atom representations. Even though it has been shown that models with a single pseudoatom per residue can demonstrate local fluctuations almost as well as all-atom methods (15,36), it appeared that the refinement introduced by Zacharias would be useful when comparing the properties of closely related conformations of the same protein, where many of the modifications concern side-chain conformations.

Interactions between the pseudoatoms of the Zacharias representation are treated using the standard elastic network model (37), that is, they are restricted to quadratic springs established between those pseudoatoms that lie closer than a cutoff distance, taken here to be 9 Å. All springs have the same force constant (38) and are assumed to be relaxed in the reference conformation of the protein. The spring force constant was taken here to be 0.6 kcal mol^{-1} Å^{-2}, a value slightly smaller than in one-point-per-residue coarse-grained models, which are usually set to roughly 1.0 kcal mol^{-1} Å^{-2} (16,31,36). This value was chosen to offset the higher spring density of the Zacharias representation. The pseudoatom representations of the proteins studied were derived from crystallographic atomic coordinates contained in the Protein Data Bank (39). Table 1 lists the six proteins studied. All structures were resolved to at least 2.5 Å. As mentioned below, calculations involved *apo* proteins, voluntarily excluding the heme prosthetic groups.

Brownian dynamics simulations

Brownian dynamics simulations are now widely used for the computation of biomolecular diffusional association rates (40). These approaches generally use atomic-scale molecular models for the diffusing species, but are still mostly restricted to rigid descriptions of the proteins. The introduction of flexibility (for example, to describe loop movements during a gating process) represents a considerable computational cost, and is usually limited to small parts of the protein (41,42). To study global flexibility, we chose to apply Brownian dynamics to Zacharias coarse-grained models including internal movements within an elastic network force field. The consequent reduction in the number of particles, combined with simplified harmonic interaction potentials, meant that simulations could be carried out inexpensively even for large proteins.

TABLE 1 Summary of the proteins investigated in this study

PDB file	Protein	Function	Resolution	Heme types	Heme ligands	Reference
1ETP	Di-heme cytochrome c_4	Electron tranfer	2.2 Å	*c*-heme *c*-heme	His-18, Met 66 His-123, Met 167	(47)
1ATJ	Horseradish peroxidase isozyme C	Catalysis	2.15 Å	*b*-heme	His-170	(48)
1BCF	Bacterrioferritin, cytochrome b_1	Electron transfer	2.9 Å	*b*-heme	Met-52A, Met-52B	(49)
1NML	Di-heme cytochrome *c* peroxidase, IN form	Catalysis	2.2 Å	*c*-heme *c*-heme	His-55, His -71 His-201, Met-275	(50)
1RZ5	Di-heme cytochrome *c* peroxidase, OUT form	Catalysis	2.4 Å	*c*-heme *c*-heme	His-55 His-201, Met-275	(50)
1QHU	Hemopexin	Heme transport	2.3 Å	*b*-heme	His-213, His-265	(51)
1QKS	Cytochrome cd_1 nitrite reductase	Catalysis	1.55 Å	*c*-heme d_1-heme	His-17, His-69 Tyr-25, His-200	(52)

The Brownian dynamics motion of each particle in the system can be simulated using the equation of Ermak and McCammon (43),

$$\mathbf{r}_i = \mathbf{r}_i^0 + \sum_j \frac{\mathbf{D}_{ij}^0 \mathbf{F}_j^0}{kT}\Delta t + \mathbf{R}_i(\Delta t), \quad (1)$$

where $\mathbf{r}_i^0$ and $\mathbf{r}_i$ denote the position vector of particle i before and after the time step Δt. $\mathbf{D}_{ij}$ is the configuration-dependent diffusion tensor between particles i and j and $\mathbf{F}_i$ is the systematic force on particle i. The displacement $\mathbf{R}_i(\Delta t)$ is a random displacement with a Gaussian distribution, whose average value is zero and whose variance-covariance is

$$\langle \mathbf{R}_i(\Delta t)\mathbf{R}_j(\Delta t)\rangle = 2\mathbf{D}_{ij}^0 \Delta t. \quad (2)$$

Unlike systems where a flexible loop is anchored to a fixed rigid protein (41), rotational and translational motions of the protein are included in these simulations. As a consequence, hydrodynamic interactions should be taken into account to achieve an adequate description of the dynamics of the system (44). These are incorporated into the problem through the diffusion tensors $\mathbf{D}_{ij}$. Hydrodynamic drag can be approximated using the Rotne-Prager tensor and its correction for overlapping spheres (45). To represent the drag on the protein associated with a layer of bound water molecules (46), we added 3 Å to the original pseudoatom radii given by Zacharias (33). In practice, hydrodynamic drag is found to have little effect on the force constants we calculate, but it has been conserved for completeness. The reduced representation of the protein and the implicit solvent model naturally allow a considerably larger time step than all-atom molecular dynamics. After monitoring the stability of the system for a number of simulations with different time steps, we chose $\Delta t = 10$ fs, a value in agreement with the time steps typically chosen in the literature for Brownian dynamics simulations on flexible systems (41). We used a bulk solvent viscosity $\eta = 1.0$ cP (1 P being 0.1 Pa s), corresponding to water at room temperature.

Measuring mechanical properties at the residue level

We began by carrying out Brownian dynamics simulations of each protein for 500 ps at a temperature of 300 K. these simulations typically led to deformations of roughly 1 Å root-mean-square deviation with respect to the starting conformations. Using the fluctuations of the position of each particle around its starting point, we could define the force constant of each particle in the protein as

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

where $d_i = \langle d_{ij} \rangle_{j^*}$ is the average distance from particle i to the other particles j in the protein. (The sum over j^* implies the exclusion of pseudoatoms belonging to residue i. Interactions between the Cα pseudoatom of residue i and and the Cα pseudoatoms of the adjacent residues $i + 1$ and $i - 1$ are also excluded. Note that all the excluded interactions correspond to virtually constant distances.) The brackets $\langle\rangle$ denote an average taken over the whole simulation.

The force constant for each residue k is then taken to be the average of the force constants attributed to all the pseudoatoms i forming this residue:

$$K_k = \langle k_i \rangle_{i \in k}. \quad (4)$$

We use the term ''rigidity profile'' to describe the ordered set of force constants for all the residues of a given protein. Rigidity profiles were obtained for all the proteins studied here. Comparisons with the energy-minimization-based local probing method developed by Navizet et al. (29) gave excellent agreement, but, as mentioned before, the Brownian dynamics approach represents an important time gain and allows the investigation of much larger proteins. As an example of the agreement, Fig. 1 *a* compares the rigidity profile of horseradish peroxidase isozyme C calculated with the Brownian dynamics simulation and the multipoint representation (curve i) with the same data obtained using a single point per residue (curve ii) and with the corresponding results from the local probing method (curve iii). Note that curves ii and iii have been adjusted for a best fit with curve i using a single multiplicative parameter. Curves ii and iii show an overall correlation coefficient of 0.967 for 306 data points. In the case of this protein, multipoint

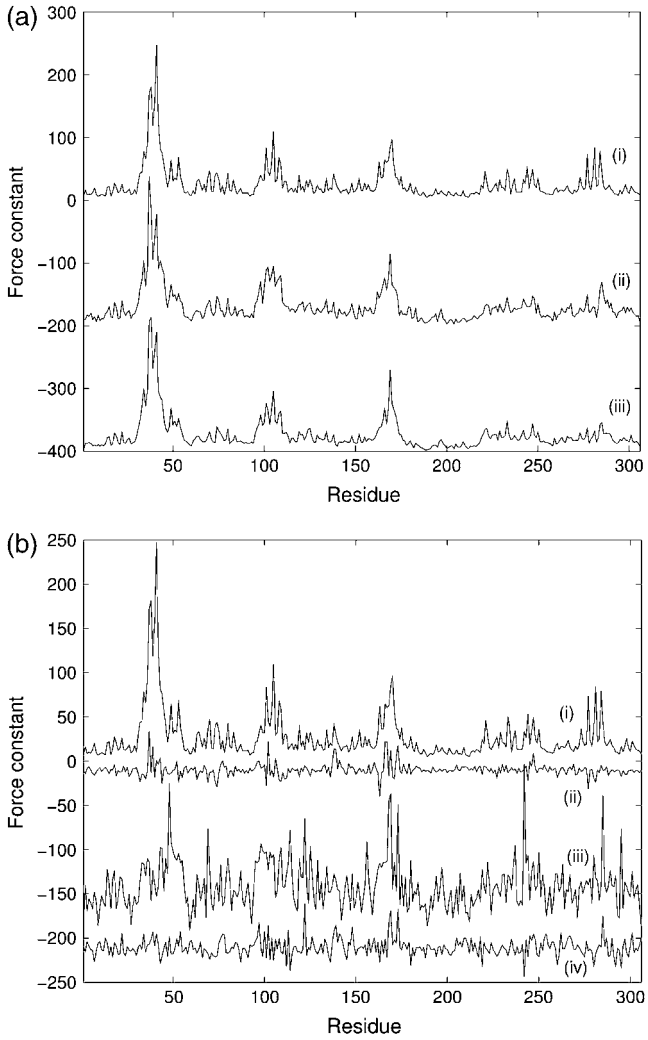

FIGURE 1 (*a*) Rigidity profiles of horseradish peroxidase. Curve i represents force constants calculated with Brownian dynamics, using the Zacharias (33) reduced model for proteins. Curve ii shows force constants calculated with Brownian dynamics, using a coarse-grained, one-point-per-residue model. Curve iii shows force constants calculated with the method of Navizet et al. (29) using a coarse-grained, one-point-per-residue model. (*b*) Effect of the heme group on residue flexibility. Curve i, force constants without the heme group; curve ii, change in force constants upon adding the heme group; curve iii, inverse of B-factors fitted using a proportionality constant and calculated in the absence of the heme group; curve iv, change in inverse B-factors upon adding the heme group. In panels *a* and *b*, curve i is correctly placed with respect to the vertical axis, whereas the remaining curves have been vertically shifted for visibility. The force constants in this figure and in Figs. 2–4, 6–9, and 11 are in kcal mol^{-1} Å^{-2} (note: 1 kcal mol^{-1} Å^{-2} = 0.07 nN Å^{-1}).

and single-point Zacharias representations (curves i and ii, respectively) also lead to very close rigidity profiles with an overall correlation factor of 0.934.

We observe that by superposing snapshots taken from the Brownian dynamics simulations, it is possible to calculate average fluctuations for each residue and thus obtain theoretical estimates of the B-factors (otherwise termed temperature factors). These values, like our force constants, can be obtained for both the complete and the artificially created *apo* forms of the hemoproteins we studied. These B-factors correlate qualitatively with the experimental results for the proteins we studied, as has been noted for other proteins in many earlier publications. However, it should also be remarked that there is a relatively poor correlation between our force constants and the calculated B-factors. This is illustrated in Fig. 1 *b*, again for horseradish peroxidase isozyme C. Curve i in this figure again shows our force constants and curve iii shows the inverse of the calculated B-factors (again adjusted for a best fit to curve i with a single multiplicative factor). The correlation coefficient is 0.368. Although the inverse of the B-factor curve indicates similar regions of rigidity, it does not discriminate individual residues in the same way as our force constants. We will return to this point later.

RESULTS

The method described above was applied to a set of six hemoproteins (see Table 1) with sizes ranging from 190 to >550 residues, which have a variety of biological functions, heme types, and heme iron coordinations. In order of increasing size, these proteins are 1), di-heme cytochrome c_4 (1ETP, first chain), which is involved in electron transfer reactions (47); 2), horseradish peroxidase isozyme C (1ATJ, first chain) (48); 3), the heme-binding site of bacterioferritin (1BCF, chains A and B), also known as cytochrome b_1, an electron transport protein (49); 4), di-heme cytochrome *c* peroxidase in its inactive (IN) form (1NML), and in its active (OUT) form (1RZ5) (50); 5), hemopexin (1QHU), a heme-binding and transport protein (51); and 6), cytochrome cd_1 nitrite reductase (1QKS, first chain) (52).

Although we have developed a pseudoatom representation of the heme group which is compatible in terms of resolution with the Zacharias peptide representation (having one pseudoatom for each of the four rings within the porphyrin and one for the central iron atom), we chose not to include the heme groups in the calculations we carried out. This allowed us to study the mechanical properties of the heme-associated residues produced by the conformation and the constitution of the peptide backbone alone, in the absence of any rigidification linked to interactions with the prosthetic groups. However, it turns out that for all the proteins we studied, the incorporation of the heme group has only a small effect on the calculated force constants. This can be seen in Fig. 1 *b* for the case of horseradish peroxidase isozyme C, where curve ii shows the differences obtained by subtracting the force constants calculated in the presence of the heme group from those calculated for the *apo* protein. Although the heme produces slightly higher values for surrounding residues (see the following section), it does not change the overall form of the curve or modify the outstandingly rigid residues discussed below. In fact, the calculated B-factors also only show small changes associated with introducing the heme group, as can be seen from the corresponding difference curve (curve iv in Fig. 1 *b*). It should be remarked that *apo* forms of the proteins we studied are artificial constructs for the purpose of the computations and involve no changes in the polypeptide conformation. No true experimental *apo* forms of the proteins studied are available.

Hemoproteins with a single heme-binding site

We first consider the relatively simple cases of horseradish peroxidase and bacterioferritin, which each possess a single heme group. For both these proteins, the residues involved in the heme pocket are found to belong to rigid structural areas. In horseradish peroxidase (whose overall conformation is represented in Fig. 2 *b*), the active site comprises a *b*-type heme with a pentacoordinated iron that is bound to the protein via His-170 on the proximal side of the heme pocket (see Fig. 2 *c* for a close-up view of the active site). Three key catalytic residues, Arg-38, Phe-41, and His-42, are located on the distal side. All these residues, along with the proximal His-170, correspond to clear peaks in the force constant

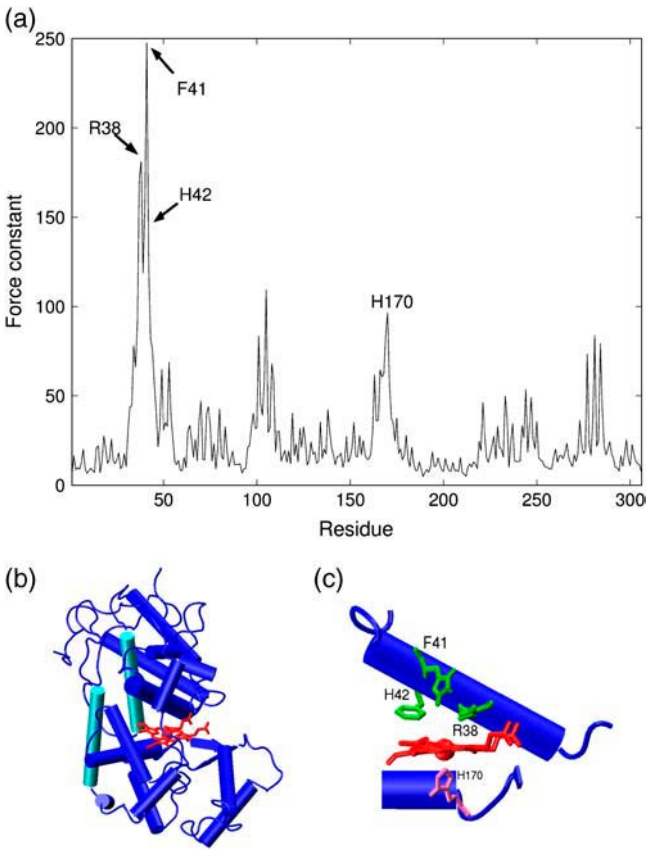

FIGURE 2 (*a*) Force-constant profile for horseradish peroxidase. (*b*) A cartoon representation of the protein with the heme group in red. (*c*) A close-up view of the heme-binding site with the catalytic residues in green and the iron binding histidine side chain in purple. The images in Figs. 2, *b* and *c*; 3, *b* and *c*; 4, *b* and *c*; 5; 6, *b* and *c*; 7, *b* and *c*; 9 *b*; and 10 were prepared using visual molecular dynamics (67).

profile (Fig. 2 *a*). Supplementary peaks in the residue ranges 100–110 and 281–290 correspond to amino acids located in helices D and J (48), which are represented in light blue in Fig. 2 *b*, and whose side chains are also pointing in the direction of the heme group.

In bacterioferritin, the heme pocket is formed by the interface between two symmetry-related subunits of 158 residues each (hence the periodicity in the rigidity profile shown in Fig. 3 *a*). The *b*-heme group is held with its quasi-twofold axis closely aligned with the twofold axis of the dimer (see Fig. 3 *b*). Fig. 3 *b* also shows that each subunit of bacterioferritin contains a binuclear metal-binding site. The iron of the *b*-heme group is hexacoordinated, with its two axial ligands being the sulfur atoms of two equivalent methionyl residues (Met-52) from the symmetry-related monomers. The heme is held in its site by numerous van der Waals contacts with side chains from residues of both subunits belonging to the segments Leu-19–Phe-26, Tyr-45–Asp-56, and Leu-71 (see Fig. 3 *c* for a closer view of the heme-binding site; note, however, that only one set of the symmetrically equivalent heme-binding residues has been labeled in this figure). Once again, all these residues have significant force constants within the profile shown in Fig. 3 *a*. A series of small peaks between residues 120 and 130 can be attributed to involvement of these residues with the binuclear metal-binding sites.

Hemoproteins with multiple domains

The smallest protein in our study, diheme cytochrome c_4 consists of two cytochrome *c*-like domains (from Ala-1 to Ser-91 and from Val-92 to His-190) that are related by a pseudo-twofold axis (see Fig. 4 *b*). The two *c*-heme groups are covalently bound to the protein via Cys-14/Cys-119 and Cys-17/Cys-122, respectively, with each hexacoordinated iron being axially bound by a histidine (His-18/His-123) and a methionine (Met-66/Met-167) residue (see Fig. 4 *c*). As with bacterioferritin, this approximate twofold symmetry in the structure of the proteins is reflected in the periodicity of its rigidity profile, although the peaks in the second domain of the protein are generally seen to be higher (Fig. 4 *a*). The most rigid parts of the protein (in the residue ranges 30–45 and 135–150) again correspond to amino acids with their side chains pointing in the direction of the heme groups

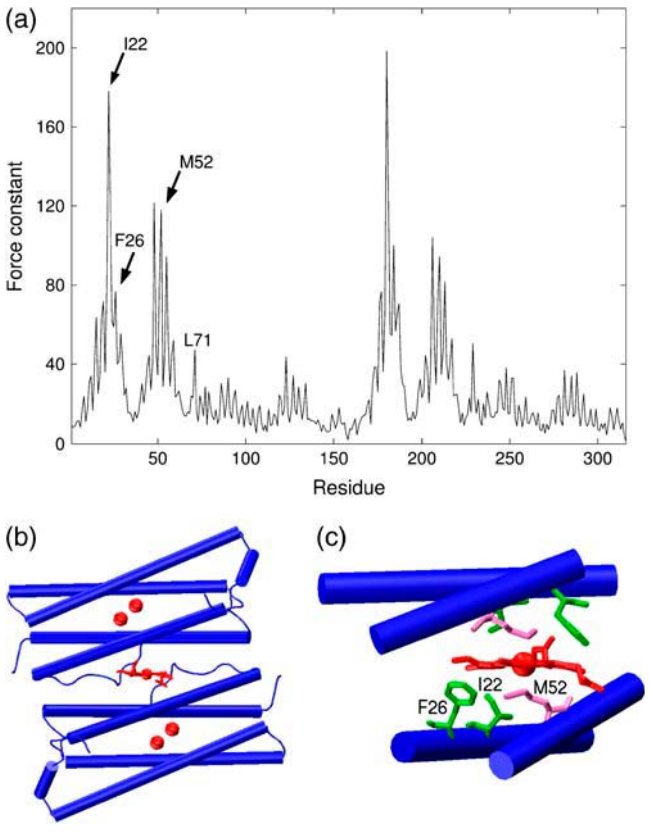

FIGURE 3 (*a*) Force-constant profile for bacterioferritin. (*b*) A cartoon representation of the protein with the heme group and the binuclear binding sites in red. (*c*) A close-up view of the active site, with heme-binding residues in green and the iron ligands in purple.

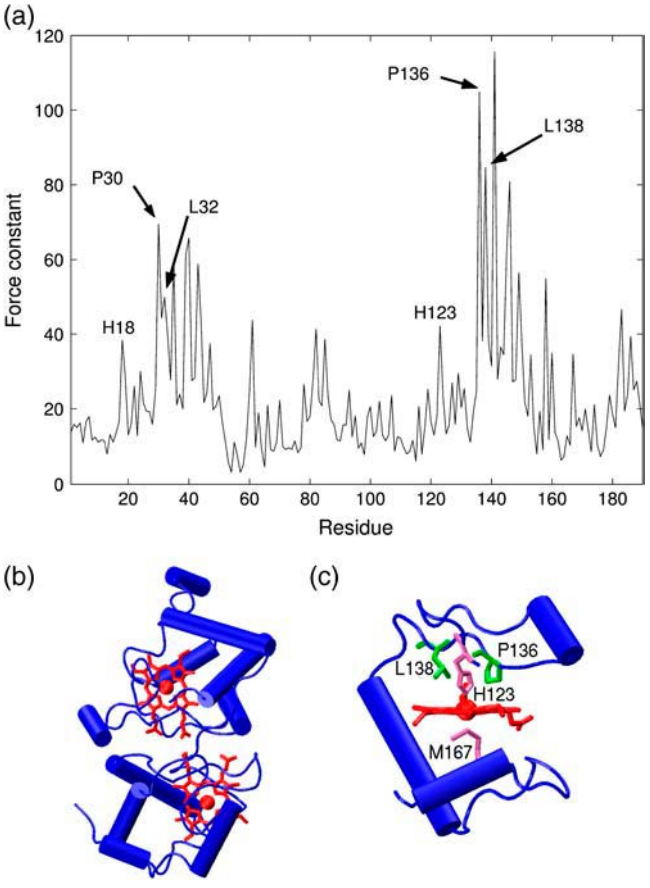

FIGURE 4 (*a*) Force-constant profile for cytochrome c_4. (*b*) A cartoon representation of the protein with the heme groups in red. (*c*) A close-up view of an active site with heme-binding residues in green and iron ligands in purple.

(such as Pro-30/Pro-136 or Leu-32/Leu-138 (see Fig. 4 *c*)). It is worth noting that the highest force constants calculated for cytochrome c_4 ($\sim$100 kcal mol^{-1} Å^{-2}) are significantly smaller than those found for the other hemoproteins studied here ($\sim$200–250 kcal mol^{-1} Å^{-2}). This may simply be related to the size of this small protein, which comprises two virtually independent domains of $\sim$95 residues, compared to $\sim$300 residues per domain for the other proteins. As remarked above, the force constant peaks shown in Fig. 4 *a*, generally associated with residues in the environment of the heme groups, are higher within the second cytochrome *c* domain (residues 92–190; compare, for example, Pro-30 and Pro-136). This observation appears to correlate with the suggestion of Kadziola and Larsen (47) that, due to sequence differences between the two domains, His-123 is in a more strained conformation than His-18.

Although the two pseudosymmetric heme-binding sites in cytochrome c_4 both appear in the rigidity profile calculated for the whole protein, this is not necessarily the case for proteins built up from structurally independent domains. For such proteins, as we and others have already noted, hinge motions between the domains will lead to low B-factors (53) and to high force constants (29) for residues belonging to the interdomain region. To avoid these values dominating the force constant profile, it is sufficient to divide the protein into its constituent domains and carry out the force-constant calculations domain by domain (that is, limiting the pseudoatom distances constituting the measured fluctuations to residues within the chosen domain). For the hemoproteins studied here, this situation arises for the two domain structures of the diheme cytochrome *c* peroxidase, cytochrome cd_1 nitrite reductase, and hemopexin. We illustrate the results in the case of the latter protein in Fig. 5. Hemopexin consists of two β-propeller domains joined by a 20-residue linker. When the protein is treated as a single unit, the highest force constants (shown in red in Fig. 5 *a*) correspond to regions close to the linker between the two domains. In contrast, once each domain is analyzed separately, the most rigid regions involve residues within the core of each domain (Fig. 5 *b*) and, in particular, as we will see shortly, heme-associated residues.

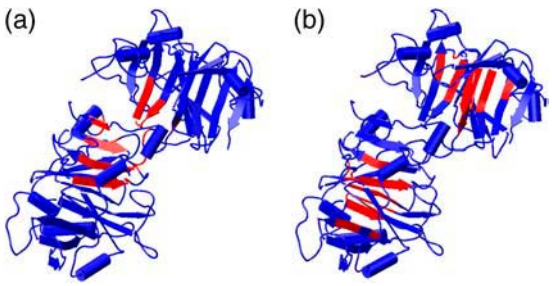

FIGURE 5 Cartoon representation of hemopexin, in which residues with force constants <50 kcal mol^{-1} Å^{-2} are represented in blue and residues with force constants >50 kcal mol^{-1} Å^{-2} are in red. (*a*) Before domain separation. (*b*) After domain separation.

We begin by discussing cytochrome cd_1 nitrite reductase, which contains a noncovalent d_1-heme group with a hexacoordinated iron that is ligated by Tyr-25 and His-200, and a covalent *c*-heme with a hexacoordinated iron bound by two axial histidine ligands (His-17 and His-69). The d_1-heme is located in the core of a β-propeller structure formed by residues 135–567, and is the site of nitrite and oxygen reduction, whereas the domain of the *c*-heme has a cytochrome *c*-type structure and is the site of electron entry from donors (see Fig. 6 *b*). A first rigidity profile (Fig. 6 *a*, *upper curve*) of the protein, obtained without domain separation, shows a series of peaks that all correspond to residues whose side chains are directed toward the d_1-heme group, and, in particular, His-345 and His-388, which are involved in the reduction reaction (see Fig. 6 *c*). In the proposed mechanism of nitrite and oxygen reduction by the enzyme, Tyr-25 must be displaced from its position as an axial ligand to allow access to the iron for substrates. It is perhaps therefore reasonable that Tyr-25 is found to be more flexible than the other functional residues of the d_1 domain, as seen in the upper curve of Fig. 6 *a*.

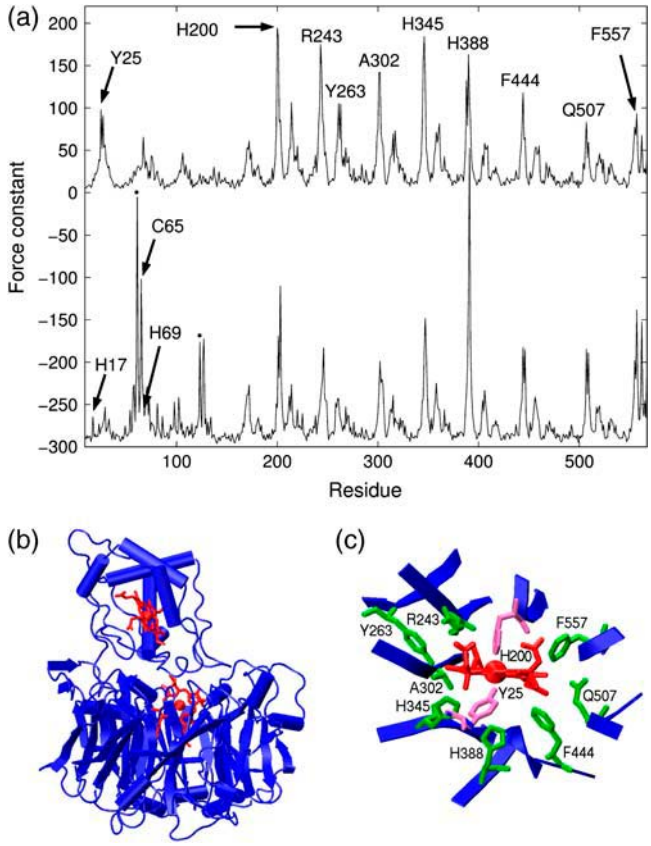

FIGURE 6 (*a*) Force-constant profile for cytochrome cd_1 before (*upper line*), and after (*lower line*) domain separation (with a vertical offset of −300 kcal mol^{-1} Å^{-2}). (*b*) A cartoon representation of the protein with the heme groups in red. (*c*) A close-up view of the d_1-heme domain, with rigid heme-binding residues in green and iron ligands in purple.

However, none of the main rigidity peaks in the upper curve of Fig. 6 *a* are related to residues belonging to the *c*-type domain. This situation changes if we separate the protein into its two domains following the definitions given in the MMDB database (54). The resulting curve shown in the lower part of Fig. 6 *a* reveals two additional rigid areas for residues in the segments 55–65 and 120–130 located in the *c* domain. There is a notable peak for Cys-65, which is covalently bound to the *c*-heme; however, the ligating residues His-17 and His-69 still do not have particularly high force constants. This may be explained by the fact that the *c*-heme domain undergoes major conformational rearrangements upon reduction (55). In particular, a His-17–Met-106 heme-ligand switch has been described, which acts together with concerted movements of a loop of the *c* domain (residues 99–116) and of Tyr-25 in the d_1 domain. Note that the domain separation also leads to several other rigidity peaks (indicated by the black dots in the lower curve of Fig. 6 *a*) that do not correspond to functional residues and that will be discussed later.

We now turn to cytochrome *c* peroxidase, which contains two covalent *c*-heme groups that can be distinguished by their redox potentials: The high-potential (HP) heme functions as an electron transfer center and is covalently bound to residues Cys-197 and Cys-200, whereas the low-potential (LP) heme corresponds to the peroxidatic center (see Fig. 7 *b*) and is covalently bound by Cys-51 and Cys-154. In the inactive form of the enzyme, a methionine (Met-275) and a histidine (His-201) residue coordinate the HP heme iron, whereas the LP heme iron is coordinated by two histidine residues (His-55 and His-71; see Fig. 7 *c*). Activation of the enzyme results in the loss of one of the histidine ligands (His-71) of the LP heme, allowing substrate binding. The rigidity profile obtained for the inactive protein before domain separation (Fig. 7 *a*, *upper curve*) shows mainly peaks corresponding to functional residues of the LP heme domain, including Pro-81 and Phe-93. His-71, which must release the heme iron to allow peroxide access, belongs to a flexible loop (residues 68–78) and indeed has a much smaller force constant than the other histidine ligand (His-55). If the force constants are recalculated after separation into three domains following the definition in the MMDB database, new rigidity peaks appear in the profile (Fig. 7 *a*, *lower curve*), among which we can find the HP heme iron ligands His-201 and Met-275. Once again, several nonfunctional residues with high force constants are seen in Fig. 7 *a* (*black dots*). Their significance will be discussed later.

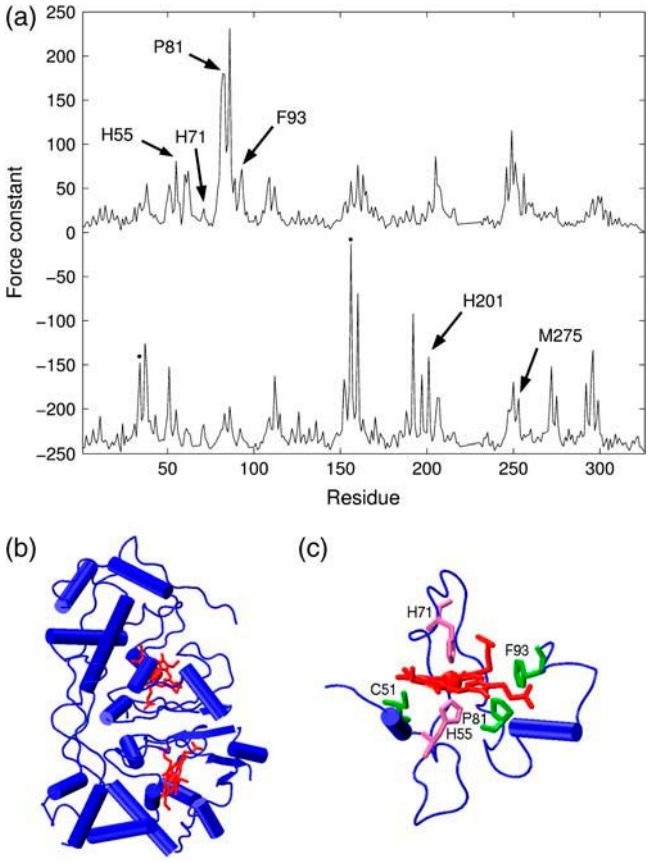

FIGURE 7 (*a*) Force-constant profile for cytochrome *c* peroxidase before (*upper line*) and after (*lower line*) domain separation (with a vertical offset of −250 kcal mol^{-1} Å^{-2}). (*b*) A cartoon representation of the protein with the heme groups in red. (*c*) A close-up view of the LP domain, with the rigid heme-binding residues in green and iron ligands in purple.

Distinguishing between different conformations of the same protein

The increased resolution of the Zacharias representation has been used to compare two conformations of diheme cytochrome *c* peroxidase: its inactive (IN) form, where the peroxidatic heme in the low-potential domain is coordinated by His-55 and His-71, and its active (OUT) form, where the

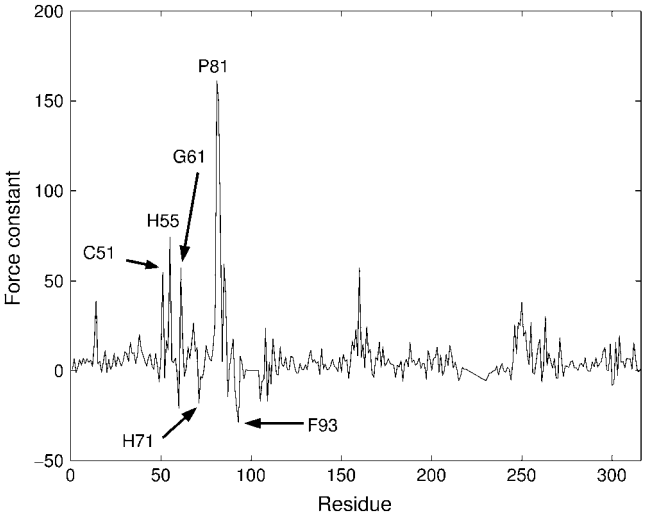

FIGURE 8 Changes in the force constant profile for cytochrome *c* peroxidase when changing from the inactive form to the active form.

release of the distal histidine (His-71) allows substrate binding. Fig. 8 shows the variations in residue rigidity after activation of the protein. As might be expected, the main changes concern residues that are involved in the LP heme domain. In particular, we note a decrease in rigidity of the distal histidine, which no longer coordinates the heme iron, and of Phe-93, which interacts with a propionate of the LP heme in the inactive form, but not in the active form. The residues with a significant increase in their force constant are the heme-binding Cys-51 and the proximal histidine (His-55), as well as residues that interact with His-55 via hydrogen bonds, Gly-61 (via its carbonyl group) and Pro-81, thus stabilizing its orientation. Protein activation thus overall appears to generate a more rigid catalytic site.

Hemopexin: a hemoprotein with flexible functional residues?

Hemopexin is a glycoprotein that sequesters free heme from the bloodstream and releases it upon interaction with a specific surface receptor on liver cells. It comprises two β-propeller domains joined by a 20-residue linker, with the heme-binding site located between these two propeller domains in a pocket constituted by the interdomain peptide (see Fig. 9 *b*). Two histidines coordinate the heme iron, His-213 from the linker peptide and His-265. As can be seen in Fig. 9 *a*, hemopexin is the only protein in our study in which the heme-binding residues (His-213 and His-265) do not belong to a rigid area of the molecule. In fact, both these residues have extremely low force constants (whether or not we make a domain separation before calculating the rigidity profile). This unusual result correlates very well with the biological function of hemopexin, as it is also the only protein in our study that can release its heme group. Paoli et al. (51) propose that heme release results from disruption of the heme-binding pocket through movements of the domains and/or the linker peptide. The latter, to which His-213 belongs, forms the periphery of the heme-binding site and should therefore be flexible to allow disruption of the complex, in good agreement with our calculated profile. Note in the lower curve of Fig. 9 *a* (after domain separation), that the most significant rigidity peaks in this protein belong to regularly spaced residues forming the core of the β-propeller domains.

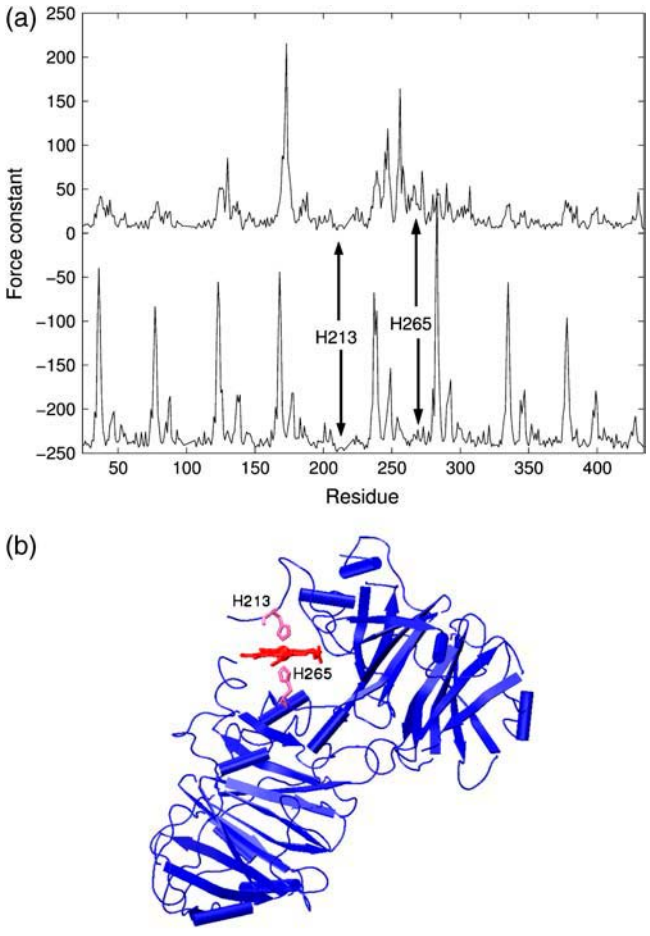

FIGURE 9 (*a*) Force-constant profile for hemopexin before (*upper line*) and after (*lower line*) domain separation (with a vertical offset of −250 kcal mol^{-1} Å^{-2}). (*b*) A cartoon representation of the protein with the heme group in red and the iron binding histidines in purple.

DISCUSSION

We have seen from the results presented above that, with the exception of the heme transport protein hemopexin, virtually all the residues that interact directly with heme groups are outstandingly rigid. There are, however, a number of significant peaks in the force-constant profiles that are not associated with the heme groups. The example of the residues belonging to the helices D and J of horseradish peroxidase have already been mentioned in the section entitled ''Hemoproteins with a single binding site''. A second example concerns the hemoproteins containing *c*-type heme groups. Both cytochrome cd_1 nitrite reductase and diheme cytochrome *c* peroxidase belong to this category and show additional peaks (after domain separation), as indicated by the black dots added to the lower curves of Figs. 6 *a* and 7 *a*, respectively. It turns out that these residues are associated with the folding nucleus of cytochrome *c* identified in a previous work (56). Cytochrome *c* proteins show only seven conserved positions across all their subfamilies. Three of these form the typical cytochrome *c* fingerprint with the heme-binding residues Cys *i*, Cys $i + 3$, and His $i + 4$. The four remaining residues occupy positions (j, $j + 4$) and (k, $k + 3$) in two almost perpendicular helices, forming a conserved network of atomic contacts that are especially strong between the aromatic groups in positions ($j + 4$) and ($k + 3$). For cytochrome cd_1 nitrite reductase and diheme cytochrome *c* peroxidase, the corresponding residues are Ala-57, Tyr-61, Met-123, Tyr-126, and Gly-34, Phe-38, Ile-156, and Tyr-159. The conservation of these residues

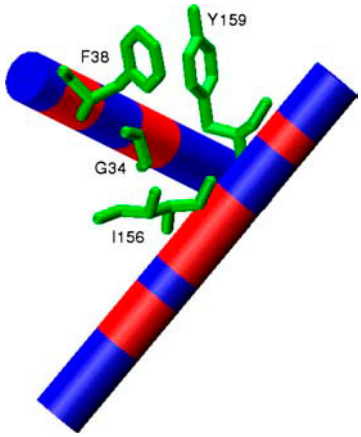

FIGURE 10 A close-up view of the folding nucleus of cytochrome *c* peroxidase. Helices are shown in a cartoon representation with the same color code as in Fig. 5. The four conserved residues are plotted in green.

among highly diverged cytochromes suggests that they are of special importance for protein folding and this group of four amino acids has been described as a folding nucleus, that is, a core from which the protein grows toward its fully folded state. A close-up view of the *c*-type folding nucleus in diheme cytochrome *c* peroxidase is represented in Fig. 10. Our determination of high force constants for these residues supports the conclusions of other groups that not only functional residues are associated with low flexibility (31). The results for the cytochrome *c* family also suggest that it would be worth studying other proteins where a folding nucleus has been identified, such as the globins (57). In passing, we remark that diheme cytochrome c_4 belongs to the same protein family. Its folding nucleus is also easily detected in force-constant terms, after domain separation, in the second domain of the protein, but the corresponding residues in the first, and globally more flexible, domain are not clearly distinguished from their neighbors.

Before looking into the mechanism underlying the appearance of unusually rigid residues, we pose the question of whether the results presented here for heme-binding proteins apply to families of proteins containing other prosthetic groups or to other classes of functional residues. This question is presently under study in our group, but the number of proteins that need to be investigated makes it impossible to give a full answer at the moment. We can, however, cite some preliminary calculations, carried out on a small number of cases, which suggest that the results on heme proteins may be more generally applicable. The first calculations concern two non-heme ion-binding proteins, transferrin (58), a monomeric iron-transporter with two metal binding sites, and rubrerythrin (59), a dimeric protein with a bound iron and a binuclear iron-zinc center that has a number of enzymatic activities. The force-constant profiles of both transferrin and rubrerythrin show that the residues directly bound to the metal ions, along with a number of closely neighboring residues, are again associated with unusually high force constants. In the case of transferrin, although the metal ions are eventually released from the protein this does not weaken the force constants of the interacting residues, as we saw in the case of the heme transport protein hemopexin (see Results, fourth section). This difference may be related to the fact that transferrin releases iron not by a simple domain movement, but by a pH-controlled mechanism that modifies the ionization state of the iron-binding residues and thus significantly changes the local environment of the metal-binding centers (60). The second example concerns proteins containing prosthetic groups other than hemes. Among the proteins we are studying, bacteriochlorophyll *a* protein (61), containing noncovalently bound chlorophyll rings, again shows the expected correlation between high force constants and residues directly associated with, or close to, the prosthetic groups. In contrast, for the oxidized form of flavodoxin (62), containing a noncovalently bound flavin mononucleotide cofactor, and for kinesin (63), containing ADP, the residues interacting with the bound groups do not have particularly high force constants, presumably because, like hemopexin, both the bound molecules must eventually be released by the proteins. The quantitative results for this group of proteins are available in Supplementary Material. Although these few examples are only anecdotal, they encourage us to believe that our findings may indeed be more general and that residue-by-residue force constants can be a useful guide to functional properties.

We now turn to the interesting question of how a protein actually succeeds in rigidifying a small number of specific residues. Given the results already presented, it should be possible to find an explanation within the model used for all our calculations, namely a coarse-grained pseudoatom representation with elastic network interactions between the pseudoatoms. The most obvious reason for rigidity within such a model would be that rigid residues lie within dense regions of the protein, that is, in our terms, that they have an above average number of neighbors within the elastic network. As Fig. 11 *a* shows, this is true. The group of unusually rigid residues always has well beyond the average numbers of neighbors. (The group of rigid residues is defined here as those with force constants at least one standard deviation above the average value for each protein studied. This group contains all the rigid residues associated with the heme groups or with folding nuclei, plus closely neighboring residues that also have high force constants.) However, Fig. 11 *a* also shows that there are many other residues that have a large number of neighbors but do not belong to the highly rigid group. It turns out that this special group is distinguished not only by having many neighbors, but also because each of these special residues is surrounded by other residues with high force constants. This is demonstrated in Fig. 11 *b*. The protein thus appears to build a ''support structure'' around its key residues which keeps them firmly in place within the overall protein architecture. A macro-

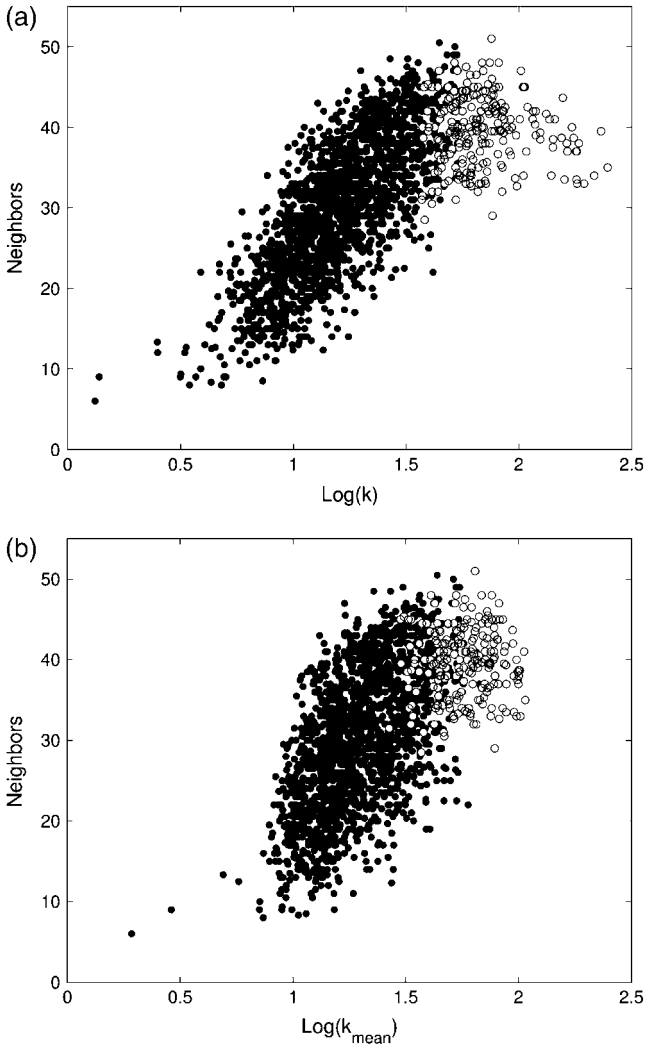

FIGURE 11 (*a*) Average number of neighbors per residue (using a 9-Å cutoff) versus the log of the force constant of the corresponding residue. (*b*) Average number of neighbors per residue versus the log of the average force constant, k_{MEAN}, of these neighbors. Open circles correspond to the set of unusually rigid residues *i* having force constants $k_i > \langle k \rangle + \sigma(k)$.

scopic structural analogy would be the superposed sets of flying buttresses in cathedrals, which support the roof vaults, which in turn lead to the keystones.

Lastly, from a more methodological point of view, we note that the definition of our force constants means that they refer to the rigidity of a given residue with respect to the overall protein structure. Although the resulting values show some correlation with the inverse of the fluctuations characterized by B-factors, the latter, being largely determined by the local packing density around a given residue within the protein structure (64), do not seem to be able to pick out functional residues with the same precision. We have also checked for correlations with NMR-derived order parameters in the case of flavodoxin (62), but find no significant correlation, since the main-chain NH groups show little variation and the side-chain mobilities seem to reflect only local hops between rotameric states, as described in a recent article analyzing such movements during molecular dynamics simulations (65). In contrast, a very preliminary study of cytochrome *c* (see Supplementary Material), suggests that there is some correlation between the folding units (''foldons'') identified by hydrogen exchange experiments (66) and our data, the groups of residues along the primary sequence with high force constants being generally associated with early folding units. This comparison, however, needs to be confirmed for other proteins.

CONCLUSIONS

Enzyme activity is now widely regarded as being dependent not only on structure, but also on mechanical properties. Although the existence, and indeed the necessity, of atomic fluctuations within proteins is well established, our understanding of how local variations in fluctuations play a biological role in specific regions, such as catalytic sites, is still very limited. It is therefore of interest to develop a simple approach for quantifying flexibility on a residue-by-residue basis. To this end, we extended our earlier method of determining force constants describing the displacement of any chosen residue within the framework of an overall protein structure, first by adopting a more refined protein representation that takes into account side-chain conformations and, second, by greatly speeding up the calculations with the help of Brownian dynamics simulations.

We applied this method to a set of six hemoproteins having a variety of structures and functions, paying particular attention to the mechanical properties of the residues located around the active heme groups. With very few exceptions, it was shown that the ligands chelating the heme group are unusually rigid, with force constants well above those of other residues. In the case of proteins consisting of more than one structural domain, it is, however, generally necessary to perform calculations on a domain-by-domain basis, to avoid the interdomain hinge-region force constants dominating the internal properties of each domain. The only significant exception to the high-force constant rule involves hemopexin, which is a heme transport protein and consequently must release the heme group to perform its function. Preliminary studies of other proteins containing metal-binding centers or prosthetic groups other than hemes suggest that these general results may apply to other protein families.

The overall results suggest that active-site residues in hemoproteins must be not only carefully positioned, but also strongly held in place for the protein to carry out its biological function. The exceptions to this rule also seem to be correlated with biological function since they concern heme-associated residues that have low force constants but are also known to be mobile during the active cycle of the associated protein. This is the case for residues such as Tyr-25 in cytochrome cd_1 and His-71 in cytochrome *c* peroxidase, which

must disconnect from their heme groups to allow substrate access and exhibit much smaller force constants than the corresponding permanently linked residues (His-200 and His-55, respectively). Another striking example is provided by the heme transport protein hemopexin, where the residues around the prosthetic group must be able to move to allow heme uptake and release.

We also showed that the more refined protein representation we employed here is capable of analyzing the subtle mechanical changes linked to conformational changes within a single protein structure. Thus, we were able to detect differences in the rigidity profiles of the active and inactive forms of cytochrome *c* peroxidase that correlate well with the known role of the various active-site residues in this enzyme's function. The same method will hopefully also be applicable to studies of the mechanical consequences of point mutations and thus to a better understanding of their functional consequences.

Lastly, our formulation of residue-level force constants showed that rigid residues correspond not simply to residues with large numbers of neighbors (that is, those within dense regions of the protein), but also to residues that are in turn surrounded with other rigid residues. The protein thus seems to build up mechanical stability in its key sites by a support structure that can involve a much larger portion of its overall architecture. The key residues discussed here are those holding, or interacting with, a prosthetic group, but also a small number of residues involved in folding nuclei. Other classes of proteins certainly merit study and, in line with other recent studies, we believe that it should be possible in this way to detect key residues that have a variety of important biological functions.

SUPPLEMENTARY MATERIAL

An online supplement to this article can be found by visiting BJ Online at http://www.biophysj.org.

The authors wish to thank Fabien Cailliez for help in comparing our results with earlier force-constant calculations.

S.S.-M. wishes to thank the CNRS for a one-year postdoctoral grant during which part of this work was carried out.

REFERENCES

1. Orengo, C. A., J. E. Bray, D. W. Buchan, A. Harrison, D. Lee, F. M. Pearl, I. Sillitoe, A. E. Todd, and J. M. Thornton. 2002. The CATH protein family database: a resource for structural and functional annotation of genomes. *Proteomics.* 2:11–21.
2. Daniel, R. M., R. V. Dunn, J. L. Finney, and J. C. Smith. 2003. The role of dynamics in enzyme activity. *Annu. Rev. Biophys. Biomol. Struct.* 32:69–92.
3. Parak, F. 2003. Physical aspects of protein dynamics. *Rep. Prog. Phys.* 66:103–129.
4. Nolde, S. B., A. S. Arseniev, V. Y. Orekhov, and M. Billeter. 2002. Essential domain motions in barnase revealed by MD simulations. *Proteins.* 46:250–258.
5. Pang, A., Y. Arinaminpathy, M. S. Sansom, and P. C. Biggin. 2003. Interdomain dynamics and ligand binding: molecular dynamics simulations of glutamine binding protein. *FEBS Lett.* 550:168–174.
6. Karplus, M., and J. Kuriyan. 2005. Molecular dynamics and protein function. *Proc. Natl. Acad. Sci. USA.* 102:6679–6685.
7. Brooks, B., and M. Karplus. 1983. Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor. *Proc. Natl. Acad. Sci. USA.* 80:6571–6575.
8. Brooks, B., and M. Karplus. 1985. Normal modes for specific motions of macromolecules: application to the hinge-bending mode of lysozyme. *Proc. Natl. Acad. Sci. USA.* 82:4995–4999.
9. Perahia, D., and L. Mouawad. 1995. Computation of low-frequency normal modes in macromolecules: improvements to the method of diagonalization in a mixed basis and application to hemoglobin. *Comput. Chem.* 19:241–246.
10. Bahar, I., A. R. Atilgan, and B. Erman. 1997. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Fold. Des.* 2:173–181.
11. Haliloglu, T., I. Bahar, and B. Erman. 1997. Gaussian dynamics of folded proteins. *Phys. Rev. Lett.* 79:3090–3093.
12. Tama, F., F. X. Gadea, O. Marques, and Y. H. Sanejouand. 2000. Building-block approach for determining low-frequency normal modes of macromolecules. *Proteins.* 41:1–7.
13. Bahar, I., A. R. Atilgan, M. C. Demirel, and B. Erman. 1998. Vibrational dynamics of folded proteins: significance of slow and fast motions in relation to function and stability. *Phys. Rev. Lett.* 80:2733–2736.
14. Hinsen, K., A. Thomas, and M. J. Field. 1999. Analysis of domain motions in large proteins. *Proteins.* 34:369–382.
15. Tama, F., W. Wriggers, and C. L. Brooks 3rd. 2002. Exploring global distortions of biological macromolecules and assemblies from low-resolution structural information and elastic network theory. *J. Mol. Biol.* 321:297–305.
16. Atilgan, A. R., S. R. Durell, R. L. Jernigan, M. C. Demirel, O. Keskin, and I. Bahar. 2001. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys. J.* 80:505–515.
17. Ming, D., Y. Kong, M. A. Lambert, Z. Huang, and J. Ma. 2002. How to describe protein motion without amino acid sequence and atomic coordinates. *Proc. Natl. Acad. Sci. USA.* 99:8620–8625.
18. Kim, M. K., G. S. Chirikjian, and R. L. Jernigan. 2002. Elastic models of conformational transitions in macromolecules. *J. Mol. Graph. Model.* 21:151–160.
19. Navizet, I., R. Lavery, and R. L. Jernigan. 2004. Myosin flexibility: structural domains and collective vibrations. *Proteins.* 54:384–393.
20. Pandey, B. P., C. Zhang, X. Yuan, J. Zi, and Y. Zhou. 2005. Protein flexibility prediction by an all-atom mean-field statistical theory. *Protein Sci.* 14:1772–1777.
21. Yuan, Z., J. Zhao, and Z. X. Wang. 2003. Flexibility analysis of enzyme active sites by crystallographic temperature factors. *Protein Eng.* 16:109–114.
22. Jones, W. C., T. M. Rothgeb, and F. R. Gurd. 1976. Nuclear magnetic resonance studies of sperm whale myoglobin specifically enriched with ^{13}C in the methionine methyl groups. *J. Biol. Chem.* 251:7452–7460.
23. Richarz, R., K. Nagayama, and K. Wuthrich. 1980. Carbon-13 nuclear magnetic resonance relaxation studies of internal mobility of the polypeptide chain in basic pancreatic trypsin inhibitor and a selectively reduced analogue. *Biochemistry.* 19:5189–5196.
24. Nicholson, L. K., L. E. Kay, D. M. Baldisseri, J. Arango, P. E. Young, A. Bax, and D. A. Torchia. 1992. Dynamics of methyl groups in proteins as studied by proton-detected ^{13}C NMR spectroscopy. Application to the leucine residues of staphylococcal nuclease. *Biochemistry.* 31:5253–5263.
25. Krishna, M. M., L. Hoang, Y. Lin, and S. W. Englander. 2004. Hydrogen exchange methods to study protein folding. *Methods.* 34: 51–64.

26. Frauenfelder, H., and B. McMahon. 1998. Dynamics and function of proteins: the search for general concepts. *Proc. Natl. Acad. Sci. USA.* 95:4795–4797.
27. Zaccai, G. 2000. How soft is a protein? A protein dynamics force constant measured by neutron scattering. *Science.* 288:1604–1607.
28. Heller, W. T., D. Vigil, S. Brown, D. K. Blumenthal, S. S. Taylor, and J. Trewhella. 2004. C subunits binding to the protein kinase A RI alpha dimer induce a large conformational change. *J. Biol. Chem.* 279: 19084–19090.
29. Navizet, I., F. Cailliez, and R. Lavery. 2004. Probing protein mechanics: residue-level properties and their use in defining domains. *Biophys. J.* 87:1426–1435.
30. Bartlett, G. J., C. T. Porter, N. Borkakoti, and J. M. Thornton. 2002. Analysis of catalytic residues in enzyme active sites. *J. Mol. Biol.* 324: 105–121.
31. Yang, L. W., and I. Bahar. 2005. Coupling between catalytic site and collective dynamics: a requirement for mechanochemical activity of enzymes. *Structure.* 13:893–904.
32. Chapman, S. K., S. Daff, and A. W. Munro. 1997. Heme: the most versatile redox centre in biology? *Struct. Bond.* 88:39–70.
33. Zacharias, M. 2003. Protein-protein docking with a reduced protein model accounting for side-chain flexibility. *Protein Sci.* 12:1271–1282.
34. Bastard, K., C. Prevost, and M. Zacharias. 2005. Accounting for loop flexibility during protein-protein docking. *Proteins.* 62:959–969.
35. Zacharias, M. 2005. ATTRACT: protein-protein docking in CAPRI using a reduced protein model. *Proteins.* 60:252–256.
36. Doruker, P., R. L. Jernigan, and I. Bahar. 2002. Dynamics of large proteins through hierarchical levels of coarse-grained structures. *J. Comput. Chem.* 23:119–127.
37. Tozzini, V. 2005. Coarse-grained models for proteins. *Curr. Opin. Struct. Biol.* 15:144–150.
38. Tirion, M. M. 1996. Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Phys. Rev. Lett.* 77:1905–1908.
39. Berman, H. M., T. Battistuz, T. N. Bhat, W. F. Bluhm, P. E. Bourne, K. Burkhardt, Z. Feng, G. L. Gilliland, L. Iype, S. Jain, P. Fagan, J. Marvin, D. Padilla, V. Ravichandran, B. Schneider, N. Thanki, H. Weissig, J. D. Westbrook, and C. Zardecki. 2002. The Protein Data Bank. *Acta Crystallogr. D Biol. Crystallogr.* 58:899–907.
40. Gabdoulline, R. R., and R. C. Wade. 2002. Biomolecular diffusional association. *Curr. Opin. Struct. Biol.* 12:204–213.
41. Wade, R. C., M. E. Davis, B. A. Luty, J. D. Madura, and J. A. McCammon. 1993. Gating of the active site of triose phosphate isomerase: Brownian dynamics simulations of flexible peptide loops in the enzyme. *Biophys. J.* 64:9–15.
42. Kamiya, Y., and C. A. Reynolds. 1999. Brownian dynamics simulations of the β_2-adrenergic receptor extracellular loops: evidence for helix movement in ligand binding? *J. Mol. Struct. THEOCHEM.* 469: 229–232.
43. Ermak, D. L., and J. A. McCammon. 1978. Brownian dynamics with hydrodynamic interactions. *J. Chem. Phys.* 69:1352–1360.
44. Pastor, R. W., R. Venable, and M. Karplus. 1988. Brownian dynamics simulation of a lipid chain in a membrane bilayer. *J. Chem. Phys.* 89: 1112–1127.
45. Rotne, J., and S. Prager. 1969. Variational treatment of hydrodynamic interaction in polymers. *J. Chem. Phys.* 50:4831–4837.
46. Halle, B., and M. Davidovic. 2003. Biomolecular hydration: from water dynamics to hydrodynamics. *Proc. Natl. Acad. Sci. USA.* 100: 12135–12140.
47. Kadziola, A., and S. Larsen. 1997. Crystal structure of the dihaem cytochrome c_4 from *Pseudomonas stutzeri* determined at 2.2 Å resolution. *Structure.* 5:203–216.
48. Gajhede, M., D. J. Schuller, A. Henriksen, A. T. Smith, and T. L. Poulos. 1997. Crystal structure of horseradish peroxidase C at 2.15 Å resolution. *Nat. Struct. Biol.* 4:1032–1038.
49. Frolow, F., A. J. Kalb, and J. Yariv. 1994. Structure of a unique twofold symmetric haem-binding site. *Nat. Struct. Biol.* 1:453–460.
50. Dias, J. M., T. Alves, C. Bonifacio, A. S. Pereira, J. Trincao, D. Bourgeois, I. Moura, and M. J. Romao. 2004. Structural basis for the mechanism of Ca(2+) activation of the di-heme cytochrome c peroxidase from *Pseudomonas nautica* 617. *Structure.* 12:961–973.
51. Paoli, M., B. F. Anderson, H. M. Baker, W. T. Morgan, A. Smith, and E. N. Baker. 1999. Crystal structure of hemopexin reveals a novel high-affinity heme site formed between two β-propeller domains. *Nat. Struct. Biol.* 6:926–931.
52. Fulop, V., J. W. Moir, S. J. Ferguson, and J. Hajdu. 1995. The anatomy of a bifunctional enzyme: structural basis for reduction of oxygen to water and synthesis of nitric oxide by cytochrome cd1. *Cell.* 81:369–377.
53. Bahar, I., and R. L. Jernigan. 1999. Cooperative fluctuations and subunit communication in tryptophan synthase. *Biochemistry.* 38:3478–3490.
54. Chen, J., J. B. Anderson, C. DeWeese-Scott, N. D. Fedorova, L. Y. Geer, S. He, D. I. Hurwitz, J. D. Jackson, A. R. Jacobs, C. J. Lanczycki, C. A. Liebert, C. Liu, T. Madej, A. Marchler-Bauer, G. H. Marchler, R. Mazumder, A. N. Nikolskaya, B. S. Rao, A. R. Panchenko, B. A. Shoemaker, V. Simonyan, J. S. Song, P. A. Thiessen, S. Vasudevan, Y. Wang, R. A. Yamashita, J. J. Yin, and S. H. Bryant. 2003. MMDB: Entrez's 3D-structure database. *Nucleic Acids Res.* 31:474–477.
55. Nurizzo, D., F. Cutruzzola, M. Arese, D. Bourgeois, M. Brunori, C. Cambillau, and M. Tegoni. 1999. Does the reduction of c heme trigger the conformational change of crystalline nitrite reductase? *J. Biol. Chem.* 274:14997–15004.
56. Ptitsyn, O. B. 1998. Protein folding and protein evolution: common folding nucleus in different subfamilies of *c*-type cytochromes? *J. Mol. Biol.* 278:655–666.
57. Ptitsyn, O. B., and K. L. Ting. 1999. Non-functional conserved residues in globins and their possible role as a folding nucleus. *J. Mol. Biol.* 291:671–682.
58. Hall, D. R., J. M. Hadden, G. A. Leonard, S. Bailey, M. Neu, M. Winn, and P. F. Lindley. 2002. The crystal and molecular structures of diferric porcine and rabbit serum transferrins at resolutions of 2.15 and 2.60 Å, respectively. *Acta Crystallogr. D Biol. Crystallogr.* 58:70–80.
59. Sieker, L. C., M. Holmes, I. Le Trong, S. Turley, B. D. Santarsiero, M. Y. Liu, J. LeGall, and R. E. Stenkamp. 1999. Alternative metal-binding sites in rubrerythrin. *Nat. Struct. Biol.* 6:308–309.
60. Hemadi, M., P. H. Kahn, G. Miquel, and J. M. El Hage Chahine. 2004. Transferrin's mechanism of interaction with receptor 1. *Biochemistry.* 43:1736–1745.
61. Tronrud, D. E., M. F. Schmid, and B. W. Matthews. 1986. Structure and X-ray amino acid sequence of a bacteriochlorophyll a protein from *Prosthecochloris aestuarii* refined at 1.9 Å resolution. *J. Mol. Biol.* 188:443–454.
62. Liu, W., P. F. Flynn, E. J. Fuentes, J. K. Kranz, M. McCormick, and A. J. Wand. 2001. Main chain and side chain dynamics of oxidized flavodoxin from *Cyanobacterium anabaena*. *Biochemistry.* 40:14744–14753.
63. Kull, F. J., E. P. Sablin, R. Lau, R. J. Fletterick, and R. D. Vale. 1996. Crystal structure of the kinesin motor domain reveals a structural similarity to myosin. *Nature.* 380:550–555.
64. Halle, B. 2002. Flexibility and packing in proteins. *Proc. Natl. Acad. Sci. USA.* 99:1274–1279.
65. Best, R. B., J. Clarke, and M. Karplus. 2005. What contributions to protein side-chain dynamics are probed by NMR experiments? A molecular dynamics simulation analysis. *J. Mol. Biol.* 349:185–203.
66. Krishna, M. M., Y. Lin, L. Mayne, and S. W. Englander. 2003. Intimate view of a kinetic protein folding intermediate: residue-resolved structure, interactions, stability, folding and unfolding rates, homogeneity. *J. Mol. Biol.* 334:501–513.
67. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14:33–38, 27–38.