Coarse-Grained Biomolecular Simulation with REACH: Realistic Extension Algorithm via Covariance Hessian

Kei Moritsugu*[†] and Jeremy C. Smith*[†]

*Center for Molecular Biophysics, University of Tennessee/Oak Ridge National Laboratory, Oak Ridge, Tennessee; and [†]Computational Molecular Biophysics, Interdisciplinary Center for Scientific Computing (IWR), University of Heidelberg, Heidelberg, Germany

ABSTRACT   Coarse-graining of protein interactions provides a means of simulating large biological systems. Here, a coarse-graining method, REACH, is introduced, in which the force constants of a residue-scale elastic network model are calculated from the variance-covariance matrix obtained from atomistic molecular dynamics (MD) simulation. In test calculations, the C_α-atoms variance-covariance matrices are calculated from the ensembles of 1-ns atomistic MD trajectories in monomeric and dimeric myoglobin, and used to derive coarse-grained force constants for the local and nonbonded interactions. Construction of analytical model functions of the distance-dependence of the interresidue force constants allows rapid calculation of the REACH normal modes. The model force constants from monomeric and dimeric myoglobin are found to be similar in magnitude to each other. The MD intra- and intermolecular mean-square fluctuations and the vibrational density of states are well reproduced by the residue-scale REACH normal modes without requiring rescaling of the force constant parameters. The temperature-dependence of the myoglobin REACH force constants reveals that the dynamical transition in protein internal fluctuations arises principally from softening of the elasticity in the nonlocal interactions. The REACH method is found to be a reliable way of determining spatiotemporal protein motion without the need for expensive computations of long atomistic MD simulations.

INTRODUCTION

Internal dynamics is essential for protein function, including ligand binding to macromolecules and mechanical structural change (1–7). Among the theoretical methods used to determine protein dynamics at atomic detail are molecular dynamics (MD) simulation using an anharmonic molecular mechanics potential function, and normal mode analysis (NMA), in which the potential energy is approximated as harmonic, leading to representation of the motions as a combination of vibrational modes (8–10).

Use of the harmonic approximation considerably simplifies the potential energy surface, and NMA has thus been applied to a wide range of biological macromolecules (e.g., 11–17). However, the difficulty remains of high computational costs for large systems. To overcome this difficulty, a further simplification of NMA, the elastic network model (ENM), has been proposed and applied in various studies such as structural changes of viruses (18), the dynamical behavior of GroEL-GroES complex on ATP binding (19), the motions of domain-swapped proteins (20), a database analysis of structural changes on ligand binding (21), and structural rearrangement on protein:protein interaction (22). In ENM, the protein molecule is represented by point residues centered, for example, at the C_α atoms, and connected by harmonic springs (23–27). ENM enables collective vibrational motions to be rapidly calculated. Its simplest form uses one force-constant parameter, assuming all the force constants between residue pairs within a cutoff distance, r_{cut}, to have the same value, k_c. In an extended form, the additional

force-constant parameters for the interactions within α-helices or the intradomain interaction were also introduced (28,29). The magnitude of k_c can be determined by comparing the atomic fluctuations, the N-H bond order parameter (30,31), or the coordinate probability distribution (28), from the elastic network normal modes with those from experiment or MD simulation. In more sophisticated approaches, coarse-grained force fields have been determined by iteratively matching the internal coordinate fluctuations (32) or the forces of the coarse-grained sites (33) to all-atom MD. In these methods, therefore, k_c is determined a posteriori and ENM in this form cannot predict the magnitude of protein fluctuations. Furthermore, k_c has been found to be strongly dependent on r_{cut} (23).

To make ENM more widely applicable, in this study a new method is introduced, REACH (Realistic Extension Algorithm via Covariance Hessian), in which the residue interaction force constants are obtained directly from the atomic-detail variance-covariance matrix calculated using MD simulation. In this way, physically based atomic MD force fields can be projected onto interresidue force constants.

The methodology derived here calculates the REACH force constants by relating the harmonic-approximated potential energy of the ENM to the Hessian (second-derivative) matrix and then to the variance-covariance matrix. As a test case, MD trajectories of myoglobin are applied to calculate REACH force constants for this protein. Associated model functions are also constructed, allowing analytical calculation of the distance dependence of the local (virtual bonds, angles, and torsions) and the nonbonded interactions, enabling straightforward application in elastic network normal mode calculation. To examine whether the protein vibrational motions simulated by atomistic MD are reproduced by

the REACH method the mean-square fluctuation and vibrational density of states are calculated and compared.

Next, anticipating that a major use of coarse-grained model will be in the dynamics of protein complexes, a myoglobin dimer is examined using the REACH model. Force constants and their model functions in dimeric myoglobin are calculated and compared with the monomer results. The mean-square fluctuations of the intramolecular and intermolecular motions are separately determined so as to characterize protein:protein interactions.

The temperature dependence of protein dynamics has been extensively studied using both experiments and computer simulations. The dynamical ''glass'' transition phenomenon in protein fluctuations at $T_g \sim 180-220$ K (34–39) has been characterized in terms of softening of the effective potential wells (40). As the force constant determines the curvature of the potential energy surface, analyzing force constants from MD at different temperatures provides a direct method for examining temperature-dependent protein dynamics. Here, as an application of the methodology, the REACH force constants are calculated from myoglobin MD trajectories at various temperatures and compared. The results shed light on the length-scales important in determining protein dynamical transition phenomena.

The REACH method introduced here is a true multiscale method in that detailed information from atomic scale MD is mapped onto a coarse-grained model of macromolecule systems, allowing, in principle, extension of length scales examined by an order of magnitude.

THEORY AND METHODS

Dynamical simulations

Molecular dynamics simulations of myoglobin were performed. The simulation details of the monomer are described in the literature (41,42). Briefly, the simulations were performed at constant temperature and pressure (1 atm) conditions (the NPT ensemble) in an explicit water box using the CHARMM program (43) and CHARMM 22 all-atom potential function (44). Production runs were performed for 10 ns for the monomer and dimer simulations at 300 K. Shorter runs, of 1 ns, were also performed for the monomer at 12 temperatures ranging from 150 K to 280 K. The atomic coordinates were saved every 50 fs for analysis.

The dimer MD simulations were carried out using starting coordinate from Protein Data Bank (PDB) (45) entry 1A6G (46). A rectangular primary simulation box was built of dimensions 100 Å × 75 Å × 60 Å. To explicitly solvate the protein 12,902 TIP3P water molecules (47) and 26 chloride counterions were placed in the box, leading to an electrically neutral system of 43,780 atoms. The simulations were performed using the same potential function and simulation protocol conditions for monomeric myoglobin. Production runs were performed for 5 ns and the atomic coordinates were saved every 50 fs for analysis.

REACH: Calculation of force constants

We now summarize the basic concept of the REACH model, together with the method for calculating the corresponding model parameters from MD trajectories. The coarse-grained model was built using only C_α atom coordinates. The mass of each pseudo-atom was defined as the sum of atomic masses comprising the corresponding residue. Heteroatoms such as ions and the heme group in myoglobin were not included.

Given the coordinates of the residues in the system, the potential energy is (23)

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

where N is the number of residues, d_{ij} (d_{ij}^0) is the distance between the dynamical (equilibrium) coordinates of C_α atoms i and j, and k_{ij} is the force constant for the spring between i and j.

Given the displacement vector in the $3N$ dimensional representation, i.e., $\mathbf{r} - \mathbf{r}_0 = (x_1 - x_1^0, y_1 - y_1^0, z_1 - z_1^0, \cdots, x_N - x_N^0, y_N - y_N^0, z_N - z_N^0)^t$, Eq. 1 becomes

$$E = \frac{1}{2}(\mathbf{r} - \mathbf{r}_0)^t \mathbf{K}(\mathbf{r} - \mathbf{r}_0), \qquad (2)$$

where the force constant (Hessian) matrix $\mathbf{K}$ is

$$\mathbf{K} = \begin{pmatrix} \mathbf{K}_{11} & \cdots & \mathbf{K}_{1N} \\ \vdots & \ddots & \vdots \\ \mathbf{K}_{N1} & \cdots & \mathbf{K}_{NN} \end{pmatrix}, \quad \mathbf{K}_{ii} = \begin{pmatrix} \sum_j k_{ij}^{xx} & \sum_j k_{ij}^{xy} & \sum_j k_{ij}^{xz} \\ \sum_j k_{ij}^{yx} & \sum_j k_{ij}^{yy} & \sum_j k_{ij}^{yz} \\ \sum_j k_{ij}^{zx} & \sum_j k_{ij}^{zy} & \sum_j k_{ij}^{zz} \end{pmatrix},$$

$$\mathbf{K}_{ij} = \begin{pmatrix} -k_{ij}^{xx} & -k_{ij}^{xy} & -k_{ij}^{xz} \\ -k_{ij}^{yx} & -k_{ij}^{yy} & -k_{ij}^{yz} \\ -k_{ij}^{zx} & -k_{ij}^{zy} & -k_{ij}^{zz} \end{pmatrix} \quad (i \neq j), \qquad (3)$$

where $k_{ij}^{st} = k_{ij}(s_i^0 - t_j^0)^2/(d_{ij}^0)^2$ and $s,t \in \{x, y, z\}$. The force constant between i and j is then derived as $k_{ij} = -tr(\mathbf{K}_{ij}) = k_{ij}^{xx} + k_{ij}^{yy} + k_{ij}^{zz}$.

Making the harmonic approximation under the equilibrium condition at constant temperature, T, allows the Hessian matrix to be calculated from the variance-covariance matrix, $\mathbf{C} = (C_{nm}) = (\langle (r_n - \langle r_n \rangle)(r_m - \langle r_m \rangle) \rangle)$ (48), as

$$\mathbf{K} = k_B T \mathbf{C}^{-1}, \qquad (4)$$

where k_B is the Boltzmann constant. The value k_{ij} can then be calculated using Eqs. 3 and 4 as

$$k_{ij} = -k_B T \, tr(\mathbf{C}_{ij}^{-1}). \qquad (5)$$

We note again that this derivation is exact only in the harmonic approximation.

Calculation of the Hessian matrix directly from $\mathbf{C}$ (in Eq. 4) leads to numerical errors arising from the calculation of the inverse matrix. To overcome this difficulty $\mathbf{C}$ was diagonalized and then $\mathbf{K}$ calculated from the eigenvalues and eigenvectors of $\mathbf{C}$. $\mathbf{C}$ is diagonalized as

$$\mathbf{CV} = \mathbf{V}\Delta, \qquad (6)$$

where $\mathbf{V} = (\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots)$ is the associated eigenvector matrix and $\Delta_{ij} = (\lambda_i \delta_{ij})$ is the eigenvalue matrix.

$\mathbf{K}$ is then derived as follows:

$$\mathbf{K} = k_B T \mathbf{V} \Delta^{-1} \mathbf{V}^t$$
$$K_{ij} = k_B T \sum_n v_{in} \lambda_n^{-1} v_{jn}. \qquad (7)$$

Inspection of Eq. 7 shows that numerical error can originate from the six modes (translation and rotation modes) with eigenvalues numerically close to zero (in such cases λ_n^{-1} becomes infinitely large). These modes were thus omitted from the calculation of $\mathbf{K}$.

In Eq. 1 k_{ij} of any pair of atoms is assumed to be positive and thus the off-diagonal Hessian (in Eq. 3) to be negative. However, in general, some

off-diagonal components in a Hessian matrix may be positive even if the potential energy is positive-definite harmonic. For example, the Hessian matrix, $\begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix}$, has semi-positive (nonnegative) eigenvalues of 0, $2 - \sqrt{2}$, and $2 + \sqrt{2}$. Thus, the assumption in Eq. 1 is stronger than the standard harmonic approximation, as some off-diagonal elements of $\mathbf{H}$ calculated from the MD-derived $\mathbf{C}$ may be positive.

The anharmonicity in atomistic MD simulations may lead to the force constants deviating from the ''ideal'' harmonic approximation quantities, but the contribution will likely be small, i.e., this is not the main origin of negative force constants. Inspection of the distribution of $k(r)$ in Results (Figs. 1 a and 3 a) suggests that the errors can be assumed to be random, i.e., the MD-derived force constant, k_{MD}, can be written as

$$k_{MD} = k + k_{err}, \qquad (8)$$

where k_{err} is assumed to be random. The results also show that most k_{MD} are positive and distribute around the average, $\langle k \rangle$, with a width of $|k_{err}|$ when the pairwise distance, r, is small. k_{MD} are close to zero at large r, and therefore, the pairs are then considered to have no interaction.

$\mathbf{K}$, derived from the MD trajectories via $\mathbf{C}$, is the effective force constant representing the interaction between any given residue and the environment, the latter including both the other implicit atoms and the surrounding water molecules. In an explicit representation including both the degrees of freedom of the subsystem (e.g., C_α atoms), $\mathbf{r}_s$, and the environment system, $\mathbf{r}_e$, the effective force constant is derived as follows.

The potential energy in Eq. 2 can be decomposed as (47)

$$E = \frac{1}{2}(\mathbf{r}_s^t \quad \mathbf{r}_e^t)\begin{pmatrix} \mathbf{K}_{ss} & \mathbf{K}_{se} \\ \mathbf{K}_{es} & \mathbf{K}_{ee} \end{pmatrix}\begin{pmatrix} \mathbf{r}_s \\ \mathbf{r}_e \end{pmatrix}. \qquad (9)$$

When r_e is projected onto r_s using the condition $(\partial/\partial r_e)E = 0$, i.e., $\mathbf{r}_e = -\mathbf{K}_{ee}^{-1}\mathbf{K}_{es}\mathbf{r}_s$, the effective potential energy in Eq. 2 becomes

$$E = \frac{1}{2}\mathbf{r}_s^t \mathbf{K}^{eff}\mathbf{r}_s = \frac{1}{2}\mathbf{r}_s^t(\mathbf{K}_{ss} - \mathbf{K}_{se}\mathbf{K}_{ee}^{-1}\mathbf{K}_{es})\mathbf{r}_s. \qquad (10)$$

In the force constant derived from the atomic-resolution MD trajectories with explicit water molecules, the solvent contributions can thus be included according to the form of Eq. 10, i.e., using the protein-water ($\mathbf{K}_{se}$) and water-water ($\mathbf{K}_{ee}$) interactions.

In this study each 1-ns MD trajectory was used to calculate the variance-covariance matrix: i.e., the full trajectories were separated into 1-ns trajectories, allowing calculating of $\mathbf{C}$ from each of the 1-ns trajectories and averaging of all the matrices derived. The time length of 1 ns is long enough to characterize the vibrational component of protein fluctuation, which arises from the harmonic potential, but is not so long that the intramolecular contribution is small compared to the slow, diffusive motion. Since normal modes represent vibrational motions on an effective harmonic potential, and since the lowest frequency of the atomistic normal modes is 3.84 cm^{-1} (the period is ∼10 ps; results not shown), the MD time length of 1 ns is a suitable choice. From $\mathbf{C}$, the Hessian matrices were separately calculated using Eq. 4 and averaged. This procedure leads to an accurate estimation of $\mathbf{H}$, particularly the off-diagonal components corresponding to the correlation of two residues. The force constants were then calculated by combining with Eqs. 5–7.

To separate the internal and external motions, the best fit to a reference protein structure was used to eliminate the translational and rotational motions. If the whole protein structure is used as the reference structure, a segment, e.g., one α-helix, may undergo external motion. Such external motion leads to errors in estimating covariances of atom pairs within the segment and thus errors in the corresponding force constants. To overcome this difficulty the system was separated into segments of 20 residues (test calculations using 10 ∼ 30 residues led to only small differences (∼2%) in the force constants derived; results not shown) and the submatrices of $\mathbf{C}$ were individually calculated by best-fitting each segment, thus allowing the local interactions within the segment, such as the virtual 1–2 (between
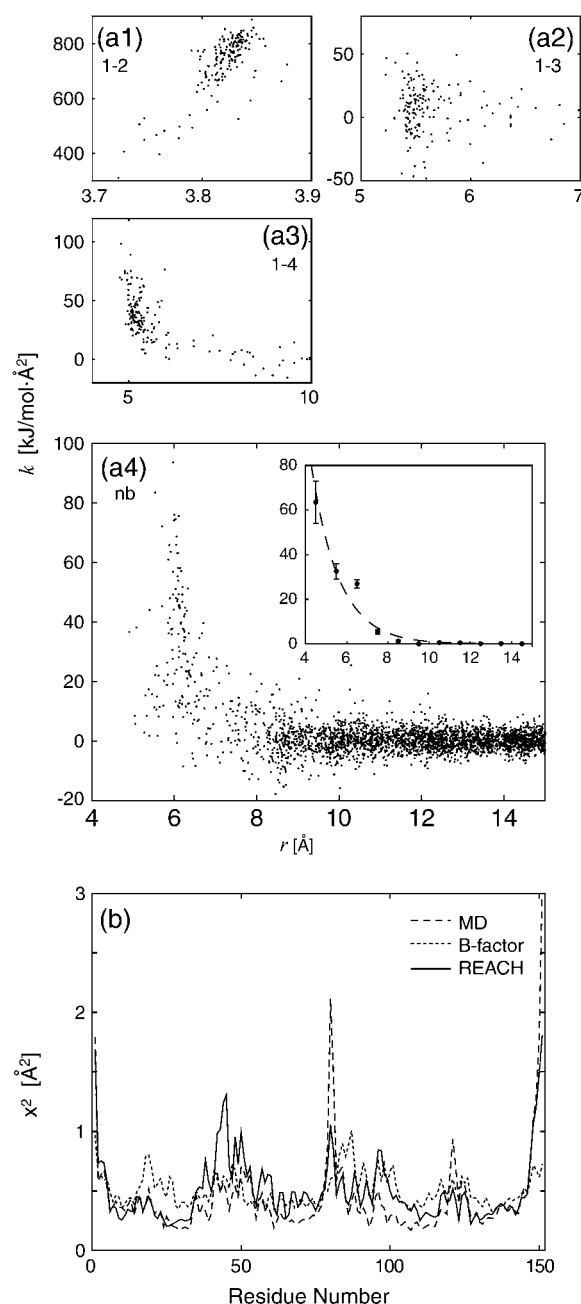

FIGURE 1   (a) REACH force constants, k, of virtual (a1) 1-2, (a2) 1-3, (a3) 1-4, and (a4) nonbonded interactions in the myoglobin monomer as a function of pairwise distance, r. In the inset of a4 the averages of k within bins of 1 Å width are shown together with the corresponding fitted model function. (b) Mean-square fluctuation of each residue in myoglobin monomer derived from MD (*dashed line*), B-factor (*dotted line*), and REACH normal modes (*solid line*).

residues i and $i+1$), 1–3 (between residues i and $i+2$), and 1–4 (between residues i and $i+3$) force constants, to be obtained.

Model parameter functions are useful for obtaining a simplified understanding of protein dynamics and for convenient application in coarse-grained MD simulation. Their functional forms were constructed here for each interaction type, depending on the shape of distributions. The parameters were calculated by fitting the distributions to the model functions.

Calculation of mean-square fluctuation from normal modes and the contribution of intermolecular motions

From the set of eigenvalues, $\{\lambda_i\}$, and eigenvectors, $\{v_i\}$, the mean-square fluctuation of residue n, x_n^2, was calculated at temperature, T, as

$$x_n^2 = \frac{k_B T}{m_n} \sum_i v_{n,i}^2 / \lambda_i, \tag{11}$$

where $v_{n,i}$ is the displacement of residue n in mode i. To calculate the contribution to x_n^2 from intermolecular motions the $\{v_i\}$ in Eq. 11 were replaced by $\{\tilde{v}_i\}$, such that the intramolecular contributions are subtracted as

$$\tilde{v}_i = v_i - \sum_j v_{1,j}(v_i \cdot v_{1,j}) - \sum_j v_{2,j}(v_i \cdot v_{2,j}), \tag{12}$$

where $v_{1,j}$ ($v_{2,j}$) is the normal-mode eigenvector calculated from the subsystem of the single monomer 1 (2). Note that these monomer eigenvectors (the number of residues in the monomer is N_{mono}) are given the full dimensions corresponding to the number of degrees of freedom of the dimer (i.e., $6N_{mono}$) by adding 0 to the elements corresponding to the other monomer (i.e., the number of 0 is $3N_{mono}$).

RESULTS

Myoglobin monomer

In an initial set of calculations the variance-covariance matrices were calculated from the MD trajectories of monomeric myoglobin and subsequently the force constants derived using Eqs. 4−7. Analytical model parameter functions were fitted to these force constants and the resulting REACH model used for calculating normal modes. The mean-square fluctuations and vibrational densities of states thus obtained are compared with those from the MD so as to examine whether the reduced normal-mode model can reproduce these aspects of protein internal dynamics obtained from a full all-atom simulation.

In Fig. 1 a the force constants between the protein residue pairs are shown as a function of the pairwise distance, r. As expected, the distribution of the virtual 1-2 bond interaction force constant, k_{12}, is quantitatively different from the others: k_{12} is much larger and has a much narrower distribution width, centered at $r \sim 3.82$ Å, the typical nearest-neighbor C_α distance. The values k_{13} and k_{14} also have relatively narrow distribution widths, centered at $r \sim 5.5$ and 5.0 Å, respectively, which are again characteristic lengths for these local interactions, especially within the α-helices dominating the myoglobin structure. In contrast, the nonbonded interaction, k_{nb}, has a broader distribution at $r > 4$ Å. At $r > 10$ Å, k_{nb} are close to zero, meaning that the corresponding residues

exert nearly no net forces on each other, and validating the conventional "one force-constant parameter model within a cutoff distance" (23–26) and a decaying force constant parameter function used empirically in the coarse-grained MD (27,50).

Model parameter functions were constructed by fitting to the distributions in Fig. 1 a. A constant value model was assumed for each of the virtual bond (1-2, 1-3, and 1-4) interactions (k_{12}, k_{13}, k_{14}) and a single exponential decay [$k_{nb}(r) = a \exp(-br)$] for the nonbonded interaction. A nonlinear fitting procedure (51) was applied to the exponential model so as to avoid the overweighting of relatively smaller values (i.e., at larger r). The resulting values and standard-deviation errors are given in Table 1.

The functional forms above resemble those of another force field recently derived for coarse-grained MD simulation (51). However, this method is improved in three respects: First, in Trylska et al. (50), the pair distribution function, $g(r)$, was calculated for atom pairs using conformational ensembles in the PDB and then the effective potential energy derived using Boltzmann inversion, i.e., $v(r) = -k_B T \ln g(r)$. The effective potential energy derived was then decomposed into several harmonic/Morse potential functions corresponding to local (1-2, 1-3, and so on) and nonlocal interactions. The associated force constants were derived from the curvatures of the potential functions. Therefore, in this method the averaged distribution for all pairs of each interaction type (e.g., 1-2, 1-3 interactions) defines one equilibrium distance, $d_{ij}^0 = d^0$, in contrast to Eq. 1, in which each interaction has a distinct d_{ij}^0. In principle, under the potential energy of Eq. 1, averaging distributions with different d_{ij}^0 leads to broadening of the potential energy wells and thus error in k. Second, the force parameters derived by the Boltzmann inversion were again iteratively modified so that test MD simulations of a coarse-grained model reproduce the all-atom MD fluctuations. This iterative procedure may preclude a direct mapping of atomistic MD results onto coarse-grained model and may include system dependence. Third, in Trylska et al. (50), only pairwise distances are considered and thus correlations through intermediate residues are not well taken into account (i.e., the correlations between residue i and k and between j and k lead to a correlation between i and j, even if i and j are not directly correlated). In contrast, the full-dimensional matrix representations of $\mathbf{C}$ and $\mathbf{H}$ in this study eliminate such indirect correlations which may cause errors in force-constant estimation.

In Fig. 1 b, the mean-square fluctuation, x^2, derived from REACH using the model force constants, are compared with

TABLE 1 REACH force constants derived from MD trajectories

	k_{12}	k_{13}	k_{14}	a_{intra}	b_{intra}	a_{inter}	b_{inter}
Monomer	735 ± 9.7	5.59 ± 2.0	31.9 ± 2.1	1980 ± 260	0.749 ± 0.031		
Dimer	712 ± 24	6.92 ± 2.3	32.0 ± 2.4	2560 ± 530	0.800 ± 0.034	1630 ± 17	0.772 ± 0.17

Note that $k_{intra}(r) = a_{intra} \exp(-b_{intra}r)$ and $k_{inter}(r) = a_{inter} \exp(-b_{inter}r)$. Units of kJ/ Å², except for b, which is in Å^{-1}.

those from crystallographic B-factors and from the MD trajectories. The MD and REACH fluctuations are similar in average magnitude to each other and to the B-factor amplitudes. Furthermore, in all three plots the regions of the protein with higher or lower fluctuations are similar. However, the correlation coefficients of the B-factor with both computational models are low (0.443 and 0.479 for the MD and REACH methods, respectively). This may be due in part to differences in the crystallographic and simulation environments (e.g., crystal packing lowers the fluctuation, especially at N- and C-termini), to the limited timescale probed in this dynamical model, and to experimental errors. The agreement between the MD and REACH x^2 is much better (correlation coefficient: 0.786), showing that REACH using the MD-derived model residue force constants reproduces well x^2 derived directly from the all-atom MD trajectories.

The results using the REACH method can be compared with the results of previous simplified normal-mode methods. In the simplest version, using one parameter with a cutoff length (18–27,29,30), the force constants of pairs with $r < r_c$ are assumed to be a constant, k_{const}. k_{const} is adjusted a posteriori so as to fit the average magnitude of experimental residue fluctuations. Therefore, the method in Tirion (23) can predict only ratios of residue fluctuations (which is nevertheless useful for determining, for example, dynamic domains (52) and separating static from flexible parts of protein molecule). In contrast, this method can predict the magnitudes of fluctuations using the model force constants determined from MD trajectories. Furthermore, this method has a significantly larger correlation coefficient with MD fluctuations than the method in Tirion (23) (which was 0.560 with $r_c = 12$ Å), due partly to the separation of the force constants into the local and nonbonded interactions (49).

Finally, the frequencies derived from the REACH calculation are compared with those from the atomistic MD. For this purpose, the vibrational spectra and the velocity autocorrelation functions were calculated and compared. The vibrational density of states along each REACH eigenvector was calculated from the MD trajectories and compared with the REACH frequency. To do this, 2^{14} MD trajectory frames, separated by 50 fs (total length of ~0.82 ns), were used for calculating the velocity autocorrelation function and its Fourier transform so as to derive the vibrational density of states, $g(\omega)$. In Fig. 2 a are shown the results of selected normal modes: 1, 100, 400. Modes 1 and 100 possess broad peaks. The positions of the MD vibrational peaks are at somewhat higher frequency than the corresponding REACH frequency, but, given the simplicity of the method, the agreement is satisfactory. In contrast, the high-frequency mode (mode 400) has a much broader peak at $200-300$ cm^{-1} in the REACH vibrational spectrum, due to the expected poor representation of the localized, high-frequency modes by the coarse-grained model.

For the purpose of comparing the frequencies of all the MD and REACH modes, the velocity autocorrelation func-
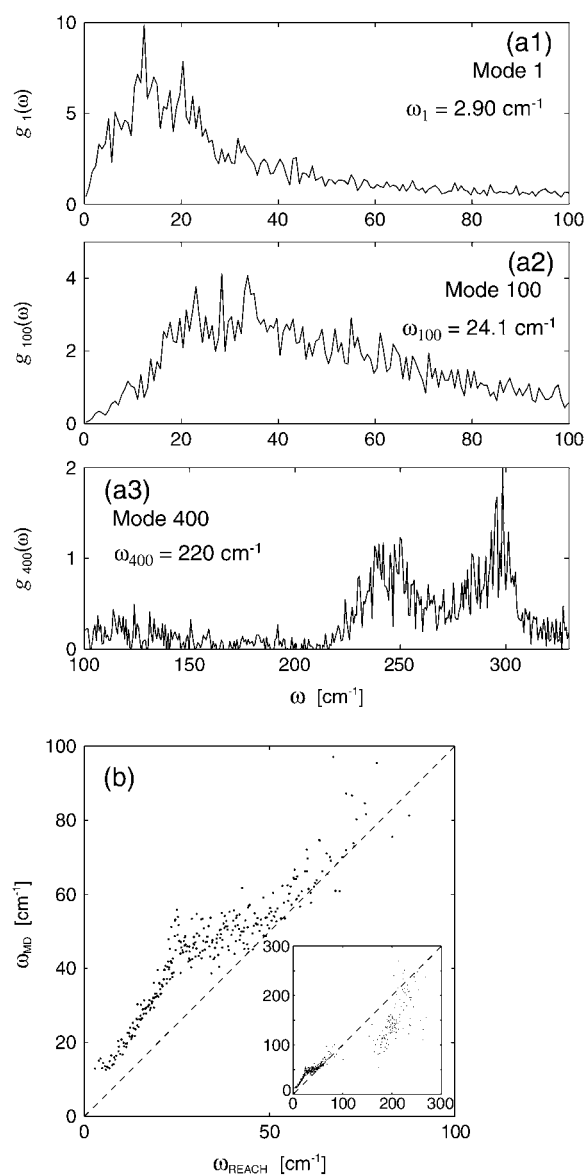

FIGURE 2 (*a*) Vibrational density of states from MD trajectory of myoglobin monomer along selected normal modes, (*a*1) mode 1; (*a*2) mode 100; and (*a*3) mode 400. The associated REACH normal-mode frequencies are also shown. In panel *b*, MD-derived frequencies are plotted as a function of REACH normal-mode frequencies. See text for details.

tions along each REACH mode was calculated from the MD trajectories and fitted to a model function, allowing the characteristic MD frequency to be derived. To do this, a model using the Langevin equation was applied to obtain the mode frequency and friction (39). Fig. 2 *b* shows the comparison between the REACH normal-mode frequency, ω_{REACH}, and the MD-derived frequency, ω_{MD}. As seen in Fig. 2 *a* for modes <100 cm^{-1}, ω_{MD} is slightly higher than ω_{REACH} but there is a good correlation, i.e., the modes with larger ω_{MD} have also larger ω_{REACH}. In contrast, the correlation is low for $\omega >100$ cm^{-1}. Thus, the simplified normal mode method reproduces well the frequencies of the larger

amplitude, lower-frequency motions ($<100 \, \text{cm}^{-1}$) simulated by atomistic MD.

Myoglobin dimer: intra- and intermolecular motions

The force constants for dimeric myoglobin were calculated from the corresponding MD trajectories and the associated model functions of the intra- and intermolecular force constants were separately derived. Using the REACH normal modes calculated from the derived model function force constants, the mean-square fluctuations were calculated and decomposed into intra- and intermolecular components.

In Fig. 3 a, k between the pairs of protein residues is plotted against r. The distributions for the virtual 1-2, 1-3, 1-4 bonds and intramolecular interactions are close to those found for the myoglobin monomer (see Fig. 1), although the dimer distributions are wider, due probably to the shorter MD sampling length of 5 ns. This similarity suggests that the intramolecular force constants derived from MD simulation using atomistic force fields are independent of protein association state. The distribution of intermolecular interaction force constants in Fig. 3 $a5$ appears at first glance to be random, due to the relatively low correlation of the C_α atom pairs at the nearest distance of $r \sim 5.5$ Å. However, the averages at $r = 5.5-10$ Å lead to statistically meaningful, nonzero values, which were used for fitting (see Fig. 3 $a6$).

Associated model functions were constructed by fitting the model functions used in the monomer calculation to the corresponding distributions. The intra- and intermolecular force constants were both modeled in single exponential form, but with different fitting parameters, i.e., $k_{\text{intra}}(r) = a_{\text{intra}} \exp(-b_{\text{intra}} r)$ and $k_{\text{inter}}(r) = a_{\text{inter}} \exp(-b_{\text{inter}} r)$. The results, together with the standard-deviation errors, are given in Table 1. The model parameters of the 1-2, 1-3, 1-4, and intramolecular interactions in the myoglobin dimer are found to be also similar to those of monomer, as are their distributions. The intermolecular force constants are slightly smaller than the corresponding intramolecular values (see Fig. 3 $a6$).

Using the derived model parameter functions the REACH normal modes of the myoglobin dimer were calculated, and from these, x^2. Fig. 3 $b1$ shows an excellent correspondence between REACH and MD, with a high correlation coefficient of 0.835. The average x^2 magnitude of REACH is slightly smaller (average over residues is 1.26 Å^2) than that from the MD (1.46 Å^2), due partially again to the fact that the 1-ns MD dynamics of the fully solvated myoglobin dimer includes diffusive motions absent from the effective harmonic potential approximated by NMA.

The dimer motion can be separated into the internal motions within two monomers (intramolecular motions) and the external motions (translation and rotation) between two monomers (intermolecular motions). In addition to the overall motions, x^2 from the intramolecular motions calculated
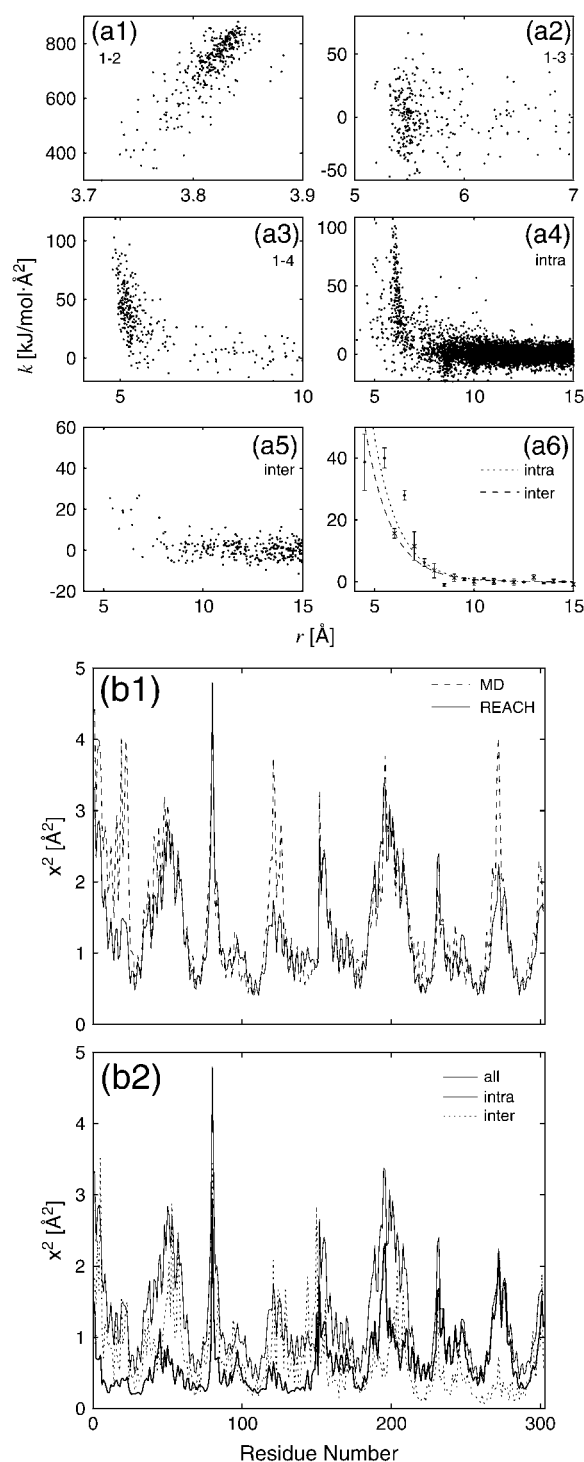

FIGURE 3 (a) REACH force constants, k, of virtual ($a1$) 1-2, ($a2$) 1-3, ($a3$) 1-4 interactions and ($a4$) nonbonded intramolecular and ($a5$) intermolecular interactions in myoglobin dimer are shown as a function of pairwise distance, r. In $a6$, the averages of k (intra, *dot* and inter, *cross*) within bins of 1 Å width are shown together with the associated fitted curves (intra, *dotted curve* and inter, *dashed curve*). ($b1$) Mean-square fluctuation of each residue in myoglobin dimer derived from MD (*dashed curve*) and REACH (*solid curve*). ($b2$) REACH mean-square fluctuation from the contributions of overall (*thin solid curve*), intramolecular motion (*thick solid curve*), and intermolecular motion (*dotted curve*).

using Eqs. 11 and 12 are also in good agreement (0.630 Å^2 from REACH and 0.581 Å^2 from MD). In Fig. 3 $b2$, the REACH x^2 contribution from the intramolecular and intermolecular motion, x^2_{intra} and x^2_{inter}, are shown. The presence of variation with residue number indicates that not only translation (fluctuations constant with residue number) but also whole-molecule rotation contribute to the intermolecular rigid-body motion. This correlates with the overall mean-square fluctuations, indicating, as expected, that the residues far from the molecular center (i.e., near the surface) undergo both larger internal fluctuation and larger rigid-body rotation.

Temperature dependence of the force constants in myoglobin monomer

In this section, as an application of this REACH method, the force constants and associated model functions are calculated from the atomistic MD trajectories of the myoglobin monomer at various temperatures, and the temperature dependence of the REACH model parameters analyzed. The atomistic MD force field used in this study has no temperature dependence, i.e., the same all-atom force field was applied to the MD simulations at all temperatures. In principle, temperature-dependent atomistic force fields could be used to parameterize the REACH model, which would then be expected to produce a more accurate description of the temperature-dependent dynamics.

In Fig. 4, a–d, the average force constants for the 1-2, 1-3, 1-4, and nonbonded interactions are plotted as a function of temperature. The 1-2 and 1-3 interactions undergo transitions at T_g ~170 K, respectively ($k_{12}(T) = 847$ for $T < T_g$ and $-0.955\,T + 1020$ for $T > T_g$, and $k_{13}(T) = 29.0$ for $T < T_g$ and $-0.155T + 55.3$ for $T > T_g$, in units of kJ/Å^2). Below T_g, these interactions are constant, consistent with harmonic dynamics. Above T_g, k decreases linearly, indicating softening of the effective harmonic potential curvature (40). In contrast, the 1-4 and nonbonded interactions decrease linearly without a transition ($k_{14}(T) = -0.0541T + 43.7$ kJ/Å^2).

In Fig. 4 e, the exponent, b, in the nonbonded force constant model is shown as a function of temperature. $1/b = l$ is the correlation length over which the force constant decays to $1/e$. b is found to increase linearly with T ($b(T) = -0.0011T + 0.41$ kJ/Å^2), i.e., l decreases with temperature, implying increased random, short-range motion at higher temperatures. This randomness can originate from, for example, Langevin-type friction (41,42).

The mean-square fluctuations, x^2, were calculated from the REACH normal modes in combination with the temperature-dependent model force constants derived above. As a, the coefficient in the nonbonded function, has little temperature dependence (data not shown); the average of a over all the temperatures was used as a constant, i.e., $a(T) = \langle a \rangle = 1040$ kJ/Å^2. Fig. 4 f plots x^2 versus T and again shows good agreement between the REACH and MD results. The transition at ~170 K is thus reproduced using reduced REACH
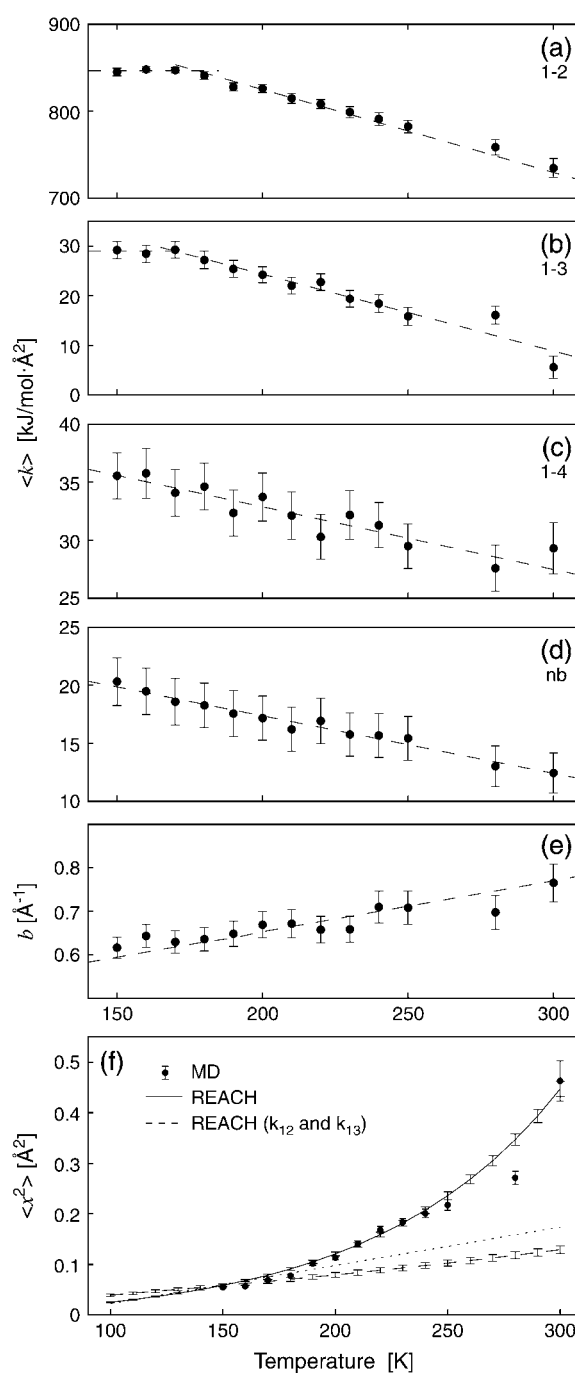

FIGURE 4 Average of force constants, k, in myoglobin monomer as a function of temperature for virtual (a) 1-2, (b) 1-3, (c) 1-4, and (d) nonbonded interactions. In case of the nonbonded interactions the force constants at $r < 10$ Å were averaged over all the pairs. In panel e, the exponent in the nonbond force constant model, b (see text), is plotted. Linear fits are also shown, as dashed lines. (f) Mean-square fluctuation, x^2, from REACH normal modes calculated using temperature-dependent model force constants. Dotted line is linear fit for lower temperature data. MD-derived x^2 is also plotted, as dots. x^2 from REACH with the same force constants but with temperature-independent 1-4 and nonbonded force constants is shown as dashed curve. See text for details.

normal modes with model force constants that have very simple temperature dependences.

Finally, the question arises as to the effects on x^2 of the temperature-dependent transitions in k_{12} and k_{13} seen in Fig. 4, a and b. To determine this, x^2 was calculated from the REACH using the temperature-dependent k_{12} and k_{13} but imposing temperature-independent k_{14} and k_{nb}, fixed at their values at 150 K, i.e., $k_{14}(T) = k_{14}(150\ K)$ and $k_{nb}(T) = k_{nb}(150\ K)$. The results in Fig. 4 f show that, although the 1-2 and 1-3 interactions exhibit temperature-dependent force constant transitions, these interactions contribute little to the overall x^2 transition. Rather, the transition in x^2 arises from the 1-4 and nonbonded interactions, the force constants of which decrease linearly with temperature such that the associated fluctuations become increasingly dominant at $T > 170$ K.

CONCLUDING REMARKS

In this article, a novel method, REACH, is presented for calculating effective coarse-grained dynamical modes in protein systems. The REACH elastic network model uses the variance-covariance matrix obtained from atomistic MD trajectories to derive the required residue interaction force constants. The force constants plotted as a function of pairwise distance present distinct distributions for the 1-2, 1-3, 1-4, and nonbonded interactions. Fitted analytical model functions further simplify the force field required for calculating REACH normal modes. The simplified potential function and force-field model reproduces well the mean-square fluctuations from atomic resolution MD, including both intramolecular motions of the myoglobin monomer and intermolecular motions in the dimer system. The distributions and model functions of the force constants are closely similar in monomeric and dimeric myoglobin.

The internal protein dynamics can be conveniently characterized using the force constants derived. For example, a temperature dependence is seen in the average virtual 1-2 and 1-3 bond force constants. The interactions undergo a transition: at $T < T_g$ they are constant, corresponding to harmonic potentials, whereas at $T > T_g$ they decrease linearly, indicating softening of the effective harmonic potentials and the resulting increase of internal fluctuation. In the 1-4 and nonbonded interactions there are no such transitions observed. The dynamical transition in the mean-square fluctuation, x^2, which was observed in inelastic neutron scattering experiments (34,38,53), is reproduced using these simple temperature-dependent model force constants obtained in this study. However, the local 1-2 and 1-3 interactions contribute little to the x^2 transition. Rather, the linear decrease of the nonlocal interactions leads to an approximately quadratic increase in x^2 by softening the elasticity in the global protein modes.

Comparison of the REACH normal-mode frequencies with the corresponding MD-derived vibrational spectra shows that the normal modes at $\omega < 100$ cm^{-1} have slightly lower frequencies than the MD, but that the correspondence is satisfactory given the simplicity of the model. For $\omega > 100$ cm^{-1}, the REACH modes are inaccurate as expected, due to the coarse-grained nature of the model. This threshold frequency (~ 100 cm^{-1}) is reasonable because the frequency of the vibration of adjacent residue pairs, which is the highest frequency in residue-scale model, is $\omega = \sim 140$ cm^{-1} when $k_{12} = 735$ [kJ/mol Å^2] and the average residual mass for myoglobin $\langle m \rangle = 106$ a.u. are used.

The mass definition of the coarse-grained residue determines the timescale of the vibrational motion. Here, the mass was defined as that of the sum of the atoms in the residue. A slight underestimation of the vibrational frequencies resulted, which may imply overestimation of the mass, although the increased roughness of the PES in atomic simulation may also contribute somewhat to the observed frequency shift. A more refined determination of the mass and iterative estimations of the force constant model functions would improve the amplitudes and frequencies of vibrational motions calculated with REACH. Furthermore, in coarse-grained systems, the contribution of the degrees of freedom other than C$_\alpha$ atoms can be represented dynamically in the form of a (generalized) Langevin equation (54), as well as in the potential energy perturbation, i.e., the effective Hessian discussed in Eq. 10.

This method has similarities with an approach presented in Chu and Voth (32) in which the averaged values and fluctuations of the effective internal coordinates were used to define a CG model matched to all-atom MD. As with REACH, this method is self-consistently multiscale in that only information from MD is used in the derivation. This is in contrast to other methods that use experimental data (e.g., x-ray B factors) (23–27). A difference, however, is that the REACH method derives residue-interaction force constants directly from the MD covariance matrix. Thus, REACH directly analytically maps the atomistic MD onto the CG model without any iterative steps, whereas in Chu and Voth (32) iterations are performed to match the required fluctuations. Therefore, arguably, the REACH method is more direct. However, as REACH does not involve iterative fitting it would not be expected to yield as close agreement with the MD fluctuations.

The REACH method requires only fitted analytical functions for the distance dependence of the various force constants used. These can be obtained from relatively short MD simulations of a small number of systems. MD studies have indicated that the atomic-detail covariance matrix does not converge in ~ 10 ns (55,56). However, a recent study demonstrated that essential conformational subspace from principal component analysis converges in a few nanoseconds and it reproduces well the space explored by a protein over ~ 200 ns timescale (57), suggesting that the force constants calculated from nanosecond MD simulations may be useful in describing protein dynamics on ~ 100 ns timescale.

Iterative procedures used in previous ENM studies are not necessary, allowing a direct mapping of atomistic MD results onto the coarse-grained model. The use of fluctuation covariance for calculating the force constants enables the equilibrium distance to be introduced into pair interactions, and avoids indirect pair correlations and force-constants error estimation when calculated using the pair distribution function.

The simple, REACH concept of calculating force-constant parameters using MD trajectories enables convenient implementation in MD simulation packages: this work is now in progress. Future work will also examine possible secondary and tertiary structure dependence of the REACH force constants before application to the study of large biological systems.

REFERENCES

1. Klotz, I. M. 1985. Ligand-receptor interactions: facts and fantasies. *Q. Rev. Biophys.* 18:227–259.

2. Benkovic, S. J., C. A. Fierke, and A. M. Naylor. 1988. Insights into enzyme function from studies on mutants of dihydrofolate reductase. *Science.* 239:1105–1110.

3. Frauenfelder, H., S. G. Sligar, and P. G. Wolynes. 1991. The energy landscapes and motions of proteins. *Science.* 254:1598–1603.

4. Lian, L. Y., I. L. Barsukov, M. J. Sutcliffe, K. H. Sze, and G. C. Roberts. 1994. Protein-ligand interactions: exchange processes and determination of ligand conformation and protein-ligand contacts. *Methods Enzymol.* 239:657–700.

5. Lamb, M. L., and W. L. Jorgensen. 1997. Computational approaches to molecular recognition. *Curr. Opin. Chem. Biol.* 1:449–457.

6. Frauenfelder, H., F. Parak, and R. D. Young. 1988. Conformational substates in proteins. *Annu. Rev. Biophys. Biophys. Chem.* 17:451–479.

7. Wang, W., O. Donini, C. M. Reyes, and P. A. Kollman. 2001. Biomolecular simulations: recent developments in force fields, simulations of enzyme catalysis, protein-ligand, protein-protein, and protein-nucleic acid noncovalent interactions. *Annu. Rev. Biophys. Biomol. Struct.* 30:211–243.

8. Go, N., T. Noguti, and T. Nishikawa. 1983. Dynamics of a small globular protein in terms of low-frequency vibrational modes. *Proc. Natl. Acad. Sci. USA.* 80:3670–3696.

9. Brooks, B. R., 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.

10. Levitt, M., C. Sander, and P. S. Stern. 1985. Protein normal-mode dynamics: trypsin inhibitor, crambin, ribonuclease and lysozyme. *J. Mol. Biol.* 181:423–447.

11. Brooks, B. R., 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.

12. Gibrat, J. F., and N. Go. 1990. Normal mode analysis of human lysozyme: study of the relative motion of the two domains and characterization of the harmonic motion. *Proteins Struct. Funct. Genet.* 8:258–279.

13. Seno, Y., and N. Go. 1990. Deoxymyoglobin studied by the conformational normal mode analysis. 1. Dynamics of globin and the heme-globin interaction. *J. Mol. Biol.* 216:95–109.

14. Perahia, P., 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.

15. Mouawad, L., and D. Perahia. 1996. Motions in hemoglobin studied by normal mode analysis and energy minimization: evidence for the existence of tertiary T-like, quaternary R-like intermediate structures. *J. Mol. Biol.* 258:393–410.

16. Delarue, M., and Y. H. Sanejouand. 2002. Simplified normal mode analysis of conformational transitions in DNA-dependent polymerases: the elastic network model. *J. Mol. Biol.* 320:1011–1024.

17. Tama, F., and Y. H. Sanejouand. 2001. Conformational change of proteins arising from normal mode calculations. *Protein Eng.* 14:1–6.

18. Tama, F., and C. L. Brooks 3rd. 2005. Diversity and identity of mechanical properties of icosahedral viral capsids studied with elastic network normal mode analysis. *J. Mol. Biol.* 345:299–314.

19. Bahar, I., and A. J. Rader. 2005. Coarse-grained normal mode analysis in structural biology. *Curr. Opin. Struct. Biol.* 15:586–592.

20. Kundu, S., and R. L. Jernigan. 2004. Molecular mechanism of domain swapping in proteins: an analysis of slower motions. *Biophys. J.* 86:3846–3854.

21. Yang, L. W., X. Liu, C. J. Jursa, M. Holliman, A. J. Rader, H. A. Karimi, and I. Bahar. 2005. iGNM: a database of protein functional motions based on Gaussian network model. *Bioinformatics.* 21:2978–2987.

22. Tobi, D., and I. Bahar. 2005. Structural changes involved in protein binding correlate with intrinsic motions of proteins in the unbound state. *Proc. Natl. Acad. Sci. USA.* 102:18908–18913.

23. Tirion, M. M. 1996. Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Phys. Rev. Lett.* 77:1905–1908.

24. 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.

25. Haliloglu, T., I. Bahar, and B. Erman. 1998. Gaussian dynamics of folded proteins. *Phys. Rev. Lett.* 79:3090–3093.

26. 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.

27. Hinsen, K. 1998. Analysis of domain motions by approximate normal mode calculations. *Proteins Struct. Funct. Genet.* 33:417–429.

28. Ming, D., and M. E. Wall. 2005. Allostery in a coarse-grained model of protein dynamics. *Phys. Rev. Lett.* 95:198103.

29. Song, G., and R. L. Jernigan. 2006. An enhanced elastic network model to represent the motions of domain-swapped proteins. *Proteins.* 63:197–209.

30. Haliloglu, T., and I. Bahar. 1999. Structure-based analysis of protein dynamics: comparison of theoretical results for hen lysozyme with x-ray diffraction and NMR relaxation data. *Proteins.* 37:654–667.

31. Ming, D., and R. Bruschweiler. 2006. Reorientational contact-weighted elastic network model for the prediction of protein dynamics: comparison with NMR relaxation. *Biophys. J.* 90:3382–3388.

32. Chu, J. W., and G. A. Voth. 2006. Coarse-grained modeling of actin filament derived from atomistic-scale solutions. *Biophys. J.* 90:1572–1582.

33. Zhou, J., I. F. Thorpe, S. Izvekov, and G. A. Voth. 2007. Coarse-grained peptide modeling using a systematic multiscale approach. *Biophys. J.* 92:4289–4303.

34. Parak, F., and H. Formanek. 1971. Untersuchung des schwingungsanteils und des kristallgitterfehleranteils des temperaturfaktors in myoglobin durch vergleich von Moessbauer-absorptionsmessungen mit Roentgenstrukturdaten. *Acta Crystallogr. A.* 27:573–578.

35. Iben, I. E., D. Braunstein, W. Doster, H. Frauenfelder, M. K. Hong, J. B. Johnson, S. Luck, P. Ormos, A. Schulte, P. J. Steinbach, A. H. Xie, and R. D. Young. 1989. Glassy behavior of a protein. *Phys. Rev. Lett.* 62:1916–1919.

36. Doster, W., S. Cusack, and W. Petry. 1989. Dynamical transition of myoglobin revealed by inelastic neutron scattering. *Nature.* 337:754–756.

37. Smith, J. C., K. Kuczera, and M. Karplus. 1990. Dynamics of myoglobin: comparison of simulation results with neutron scattering spectra. *Proc. Natl. Acad. Sci. USA.* 87:1601–1605.

38. Tournier, A. L., and J. C. Smith. 2003. Principal components of the protein dynamical transition. *Phys. Rev. Lett.* 91:208106.

39. Hayward, J. A., and J. C. Smith. 2002. Temperature dependence of protein dynamics: computer simulation analysis of neutron scattering properties. *Biophys. J.* 82:1216–1225.

40. Zaccai, G. 2000. How soft is a protein? A protein dynamics force constant measured by neutron scattering. *Science.* 288:1604–1607.

41. Moritsugu, K., and J. C. Smith. 2005. Langevin model of the temperature and hydration dependence of protein vibrational dynamics. *J. Phys. Chem. B.* 109:12182–12194.

42. Moritsugu, K., and J. C. Smith. 2006. Temperature-dependent protein dynamics: a simulation-based probabilistic diffusion-vibration Langevin description. *J. Phys. Chem. B.* 110:5807–5816.

43. Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* 4:187–217.

44. MacKerell, A. D., Jr., D. Bashford, R. L. Bellott, R. L. Jr. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. I. I. I. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B.* 102:3586–3616.

45. Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, and H. Wessig. 2000. The Protein Data Bank. *Nucleic Acids Res.* 28:235–242.

46. Vojtechovsky, J., K. Chu, J. Berendzen, R. M. Sweet, and I. Schlichting. 1999. Crystal structures of myoglobin-ligand complexes at near-atomic resolution. *Biophys. J.* 77: 2153–2174.

47. Jorgensen, W. D., J. Chandrasekhar, and J. D. Madura. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.

48. Karplus, M., and J. N. Kushick. 1981. Method for estimating the configurational entropy of macromolecules. *Macromolecules.* 14:325–332.

49. Zheng, W., and B. R. Brooks. 2005. Probing the local dynamics of nucleotide-binding pocket coupled to the global dynamics: myosin versus kinesin. *Biophys. J.* 89:167–178.

50. Trylska, J., V. Tozzini, and A. McCammon. 2005. Exploring global motions and correlations in the ribosome. *Biophys. J.* 89:1455–1463.

51. Flannery, B. P., S. A. Teukolsky, and W. T. Vetterling. 1998. Numerical Recipes in C. Cambridge University Press, Cambridge, UK.

52. Hayward, S. 1999. Structural principles governing domain motions in proteins. *Proteins.* 36:425–435.

53. Tournier, A. L., J. Xu, and J. C. Smith. 2003. Translational hydration water dynamics drives the protein glass transition. *Biophys. J.* 85: 1871–1875.

54. Izvekov, S., and G. A. Voth. 2006. Modeling real dynamics in the coarse-grained representation of condensed phase systems. *J. Chem. Phys.* 125:151101.

55. Meinhold, L., and J. C. Smith. 2005. Correlated dynamics determining x-ray diffuse scattering from a crystalline protein revealed by molecular dynamics simulation. *Phys. Rev. Lett.* 95:218103.

56. Meinhold, L., and J. C. Smith. 2007. Protein dynamics from x-ray crystallography: anisotropic, global motion in diffuse scattering patterns. *Proteins.* 66:941–953.

57. Lange, O. F., and H. Grubmuller. 2006. Can principal components yield a dimension reduced description of protein dynamics on long time scales? *J. Phys. Chem. B.* 110:22842–22852.