

## Dissipative electro-elastic network model of protein electrostatics

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2012 Phys. Biol. 9 036004

(<http://iopscience.iop.org/1478-3975/9/3/036004>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 129.219.247.33

This content was downloaded on 27/08/2014 at 00:23

Please note that [terms and conditions apply](#).

# Dissipative electro-elastic network model of protein electrostatics

Daniel R Martin, S Banu Ozkan and Dmitry V Matyushov

Center for Biological Physics, Arizona State University, PO Box 871504, Tempe, AZ 85287-1504, USA

E-mail: [dmitrym@asu.edu](mailto:dmitrym@asu.edu)

Received 24 December 2011


Accepted for publication 22 March 2012

Published 4 May 2012

Online at [stacks.iop.org/PhysBio/9/036004](http://stacks.iop.org/PhysBio/9/036004)

## Abstract

We propose a dissipative electro-elastic network model to describe the dynamics and statistics of electrostatic fluctuations at active sites of proteins. The model combines the harmonic network of residue beads with overdamped dynamics of the normal modes of the network characterized by two friction coefficients. The electrostatic component is introduced to the model through atomic charges of the protein force field. The overall effect of the electrostatic fluctuations of the network is recorded through the frequency-dependent response functions of the electrostatic potential and electric field at the protein active site. We also consider the dynamics of displacements of individual residues in the network and the dynamics of distances between pairs of residues. The model is tested against loss spectra of residue displacements and the electrostatic potential and electric field at the heme's iron from all-atom molecular dynamics simulations of three hydrated globular proteins.

 Online supplementary data available from [stacks.iop.org/PhysBio/9/036004/mmedia](http://stacks.iop.org/PhysBio/9/036004/mmedia)

## 1. Introduction

Proteins in solution exist not as static structures but instead as dynamic entities interconverting between conformational sub-states on the time-scale reflecting the activation barriers involved in the transitions [1–3]. It is becoming increasingly clear that biomolecules are fluctuating machines, using dynamical equilibria between their conformational states to promote their function [2, 4]. The flexibility of the protein structure naturally leads to changes in both the charge distribution of the protein itself and in the polarization of the interfacial waters. Hydration water, being mostly a fast subsystem, follows the protein motions. At the same time, it provides an additional fluctuating thermal bath [1], broadening the configurational space accessible to the protein motions. The ability of the protein to visit many configurational basins, separated by activation barriers [5], results in a multitude of relaxation times and complex dynamics characterizing the hydrated protein [6]. A broad range of configurations accessible to proteins put an emphasis on fluctuations and require a shift of the focus from statistical averages and corresponding thermodynamic parameters, such as equilibrium Gibbs energies, to entire statistical distributions

of observables and their dynamics expressed through time correlation functions.

This requirement is certainly the case for the problem of protein electron transfer, where, according to the standard Marcus picture [7], both the average of the energy gap between the donor and acceptor energy levels of the electron and its fluctuation are required to determine the probability of electron tunneling. This is just one example when the knowledge of the entire fluctuation spectrum of the observables is of significant interest [8]. Other applications, such as electrostatic fluctuations reported by broad-band dielectric spectroscopy [9], THz absorption [10] and light scattering techniques [11], require either the entire fluctuation spectrum or a set of cumulants of the corresponding experimental observables.

With the general focus on the fluctuation spectrum of proteins, we propose here a model for the calculation of the statistics and dynamics of the electrostatic potential and electric field in proteins. Electrostatic interactions are clearly important for protein stability and function [12]. Electrostatic solvation and interactions affect the stability of folded proteins and their aggregation and crystallization [13–15]. Electrostatics might also play a significant role in promoting high-temperature flexibility of proteins through the coupling

of the charge distribution in the protein to fluctuations of the protein–water interface [16]. Finally, dielectric spectroscopy and light scattering/absorption [9–11] require modeling of the dynamics of the protein dipole and its hydration layer, also available from the electrostatic response.

The dynamics of electrostatic fluctuations span an enormous range of time-scales from sub-picoseconds to sub-microseconds, and possibly longer [17, 18, 9]. While spectroscopic techniques are capable of recording the ultra-fast fs-to-ps relaxation [19], the fluctuations of the protein–water interface recorded by Mössbauer spectroscopy and neutron scattering cover much longer time-scales from sub-nanoseconds to sub-microseconds [3]. Our present model aims at these long, and perhaps even longer, time-scales by coarse-graining the protein into an elastic network of beads, each representing a protein residue.

Elastic network models (ENMs) have consistently shown good performance in characterizing global, large-scale motions of proteins [20–30]. While the low-frequency portion of the spectrum of protein motions is mostly determined by the distribution of mass and molecular shape [31–33], improvements are still needed to account for deficiencies of networks when localized protein motions are involved [34–38]. Electrostatic interactions, on the other hand, are notoriously long-ranged. The electrostatic potential, slowly changing over a nanometer-size biomolecule, effectively averages the local structural variations out and is mostly sensitive to the global distribution of charge within the molecule. From this perspective, the combination of even a basic elastic network with the molecular charge distribution might be sufficient to describe the statistics of the potential fluctuations and their slow dynamics. We propose here a model combining the distribution of molecular charge from the standard atomic force fields with the dynamics and statistics of protein motions derived from an elastic network.

Much of the protein electrostatics is driven by the mobility of polar groups and overall dipoles of the residues. These dipolar motions, combined with displacements of ionized residues, are the main contributors to the calculated electrostatic response, although the complete set of the residue atomic charges also incorporates higher multipoles. However, coarse-graining of the protein into residue beads eliminates librations of polar groups, and this is a limitation of the present algorithm.

To test the model, we employ all-atom molecular dynamics (MD) simulations of three hydrated heme proteins. Previous analysis of these data has emphasized the importance of nanosecond (ns) relaxation modes in the electrostatic fluctuations of the protein matrix and of the protein–water interface [16, 39]. This time-scale is relevant to the biological function since heme proteins are typical components of biology’s energy chains transporting electrons on nanosecond-to-microsecond time-scales [8, 40]. The nanosecond fluctuations are also consistent with the dynamics of the global motions of the network of residues. An ENM naturally fits the physics of the problem. The present contribution is a first step in developing network models of the electrostatic response of hydrated proteins. Here, we focus on

the electrostatics of the protein matrix only, leaving the water component of the overall response to future studies.

## 2. Theory

### 2.1. Elastic network model

A coarse-grained ENM assigns a node to a collection of atoms, thus reducing the computational burden of an all-atom normal-mode analysis. A typical coarse-graining is done on the level of individual aminoacids (1-bead model [41]). The position of the node is defined by the coordinates of the  $C^\alpha$  atom. The springs connecting the nodes represent the bonded and non-bonded interactions within an accepted cutoff distance [20–22] or by a distance-dependent force constant [28, 37, 42, 43]. The ENM diagonalizes the Hessian matrix  $\mathbf{H}$  derived from a simplified Hookean potential suggested by Tirion [20]. This potential,  $E_{ij} = C(r_{ij} - r_{0,ij})^2/2$ , describes the elongation  $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$  between nodes  $i$  and  $j$  in the network characterized by one universal force constant  $C$  and the structural information stored in the equilibrium bead positions  $\mathbf{r}_{0,i}$  ( $r_{0,ij} = |\mathbf{r}_{0,i} - \mathbf{r}_{0,j}|$ ).

The Hessian  $H_{ij}^{\alpha\beta}$  is a  $3N \times 3N$  matrix representing the protein elastic energy as a quadratic form of the Cartesian displacements  $\delta r_i^\alpha = r_i^\alpha - r_{0,i}^\alpha$  of individual beads in the network

$$E = (C/2) \sum_{i,j} H_{ij}^{\alpha\beta} \delta r_i^\alpha \delta r_j^\beta. \quad (1)$$

Here,  $i$  and  $j$  run between 1 and  $N$ ,  $\alpha, \beta$  indicate the Cartesian projections, and summation over repeated Greek indices is assumed here and throughout. The Hessian is diagonalized by the unitary matrix  $\mathbf{U}$  producing  $3N$  eigenvalues  $\lambda_m$ . This standard linear algebra formalism [20, 21] yields the statistical correlator between the displacements of residues  $i$  and  $j$  in the network in thermal equilibrium with the surrounding thermal bath

$$\langle \delta r_i^\alpha \delta r_j^\epsilon \rangle = (\beta C)^{-1} \sum_m U_{mi}^{\gamma\alpha} \lambda_m^{-1} U_{mj}^{\gamma\epsilon}, \quad (2)$$

where  $\beta = 1/(k_B T)$  is the inverse temperature.

If the network is characterized by a universal force constant and a cutoff, it yields a bell-shaped density of vibrational states. A refinement of the network by adopting a stronger coupling between covalently bound neighbors splits the density of states into two maxima, in better agreement with all-atom calculations [44, 26, 45]. This approximation is adopted in our present calculations: the spring constant was multiplied by a constant factor  $\varepsilon = 100$  for covalent neighbors. In addition, a uniform cut-off radius of  $r_c = 15 \text{ \AA}$  was used in all calculations. The sensitivity of the algorithm to the choice of the cutoff is discussed below.

### 2.2. Overdamped network dynamics

The standard mechanical ENM outlined above obviously lacks dissipative dynamics. The equations of motion are harmonic, implying oscillatory time correlation functions. In contrast, most correlation functions of hydrated proteins observed by scattering [46, 9, 11] and relaxation [47] techniques are exponential, corresponding to the overdamped (Debye)

dynamics, or stretched-exponential [9]. Several approaches can be implemented to incorporate dissipative dynamics into the mechanical network of beads. Langevin equations of motion for the ENM potential, within the general framework of the Lamm–Szabo formalism [48], have been suggested [38, 49]. This approach, and some early suggestions [50, 51], still requires parameterization (done by fitting the rotational and translational diffusion coefficients from MD) and, in addition, doubles the size of the matrix to be diagonalized. We have adopted here a more straightforward formalism, not requiring additional computational resources.

Instead of a harmonically oscillating network, we have assumed for each normal mode  $\mathbf{q}_m$ , diagonalizing the network's Hessian, an overdamped motion described by the dissipative memory kernel  $\zeta(t)$  [52]. The equation of motion for such overdamped dynamics is

$$\int_0^t \zeta(t-t') \dot{\mathbf{q}}_m(t') dt' + \lambda_m \mathbf{q}_m = \mathbf{F}(t) + \mathbf{R}(t), \quad (3)$$

where  $\mathbf{F}(t) = \mathbf{F}_\omega e^{i\omega t}$  is an external oscillating force and  $\mathbf{R}(t)$  is a randomly fluctuating force. The latter satisfies the generalized fluctuation–dissipation relations [53, 54]

$$\langle \mathbf{R}(t) \rangle = 0, \quad \langle \mathbf{R}(t) \cdot \mathbf{R}(0) \rangle = k_B T \zeta(t). \quad (4)$$

The generalized Langevin equation is typically not derived and is introduced as a phenomenological framework to describe dissipative dynamics (see, however, [55, 56]). A certain separation of time-scales is required to define the random force  $\mathbf{R}(t)$  on the rhs in (3). It must be faster than the dissipative dynamics of  $\mathbf{q}_m$ . This requirement becomes a source of conceptual difficulties when applied to systems with broad distributions of relaxation rates, such as hydrated proteins [8, 57], when a separation between ‘fast’ and ‘slow’ becomes ill-defined. Practically, a continuous distribution of relaxation times leads to a dependence of the free energy along a ‘reaction coordinate’ on the observation window (the trajectory length in numerical simulations) [8, 58].

For the problem of hydrated proteins, the fast subsystem responsible for the random noise includes a thermal bath of water solvent, relaxing on  $\sim 1$  ps time-scale, and fast vibrational modes of the protein. The latter subsystem can be identified with the low-temperature vibrational density of states [59, 60] recorded by Mössbauer absorption and neutron scattering at temperatures below the dynamical transition of a protein [61, 62]. A clear separation between the dissipative and vibrational protein dynamics is recorded by inelastic neutron scattering of partially wetted protein powders. Inelastic scattering is purely vibrational and temperature-independent at energies exceeding  $\sim 4$  meV ( $\sim 1$  ps relaxation time), but becomes quasielastic and hydration and temperature dependent at lower energies [63, 64]. We therefore tentatively assign  $\omega_h \simeq 10^{12} \text{ s}^{-1}$  as the borderline frequency separating fast and slow (dissipative) modes of the protein [16]. The elastic network considered here aims at these slower modes, for which, based on experimental observations [63, 64], we postulate dissipative dynamics. Correspondingly, the harmonic energy in (1) should be understood as the harmonic expansion of a partial free energy obtained by integrating the system

density matrix, defined by the full system Hamiltonian, over the fast water and protein degrees of freedom. These fast degrees of freedom are therefore adiabatically eliminated from the problem and only slower dissipative modes enter the equations of motion in (3).

We note also that the harmonic network used here is restricted to only one equilibrium conformation around which the quadratic expansion of the potential function has been performed. This implies that the frequencies produced by the dissipative dynamics considered here must be limited from below by the rates of activated transitions between different conformational sub-states of the protein, separated by activation barriers. The dynamics are more complex, with a power-law dissipation kernel  $\zeta(\omega)$  [6, 65], when these activated transitions are included in the observation window.

From numerical simulations, the free energy surfaces of residue displacements are still quasi-harmonic on the time-scale  $\leq 100$  ns ( $\omega_l \simeq 0.01 \text{ ns}^{-1}$ ) [66, 58], with softening spring constants as the trajectories get longer. This is also the trajectory length of the simulations presented here, which currently limits the applicability of the model. To summarize, the present formulation aims at the range of frequencies  $\omega_l \leq \omega \leq \omega_h$  between the upper cutoff of the fast subsystem and the lower cutoff of the rate of transitions between conformational sub-states.

The distinction between equation (3) and the Langevin network [38, 49] is worth emphasizing here. In the latter, uncoupled Langevin dissipative equations are first assigned to each bead of the network. However, it was noted that dissipative equations for the beads are likely to become coupled when the Langevin dynamics of beads are consistently derived by integrating out the fast degrees of freedom in the equations of motion [67]. Given that the assignment of uncoupled Langevin equations to individual beads is phenomenological from the onset, one can introduce similar phenomenology at a different level of the theory. Here, in the spirit of the standard normal-mode analysis, we introduce uncoupled dissipative equations for the normal modes in (3). This is still a phenomenological assumption requiring further testing, but it reduces to the standard normal-mode analysis in the static limit.

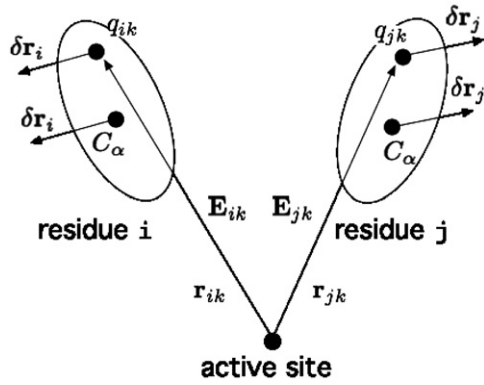
When the Laplace–Fourier transform [52] is applied to (3), the displacement response function for the collective mode  $\mathbf{q}_m$  follows. It is given as a scalar function connecting the average displacement to the average external field [53, 68],  $\langle \mathbf{q}_m(\omega) \rangle = \chi_m(\omega) \mathbf{F}_\omega$ , where  $\langle \mathbf{q}_m(t) \rangle = \langle \mathbf{q}_m(\omega) \rangle e^{i\omega t}$ . From equation (3), one gets

$$\chi_m(\omega) = [i\omega \tilde{\zeta}(\omega) + \lambda_m]^{-1}, \quad (5)$$

where  $\tilde{\zeta}(\omega)$  is the Laplace–Fourier transforms of the friction kernel  $\zeta(t)$ . Extending this procedure to all normal modes of the network, one obtains the response function of the bead displacements

$$\chi_{ij}^{\alpha\beta}(\omega) = C^{-1} \sum_m U_{mi}^{\gamma\alpha} [\lambda_m + i\omega \zeta(\omega)]^{-1} U_{mj}^{\gamma\beta}. \quad (6)$$

In equation (6), the response function  $\chi_{ij}(\omega)$ , which is a rank-2 tensor, represents the displacement of residue  $i$  of the protein due to an oscillatory force with frequency  $\omega$  applied



**Figure 1.** Sketch illustrating the application of the DENM to the calculation of the electro-elastic response function at an atomic position in the active site of the protein. The elastic displacements of residues  $i$  and  $j$  create fluctuations of the electric fields by atomic charges  $q_{ik}$  and  $q_{jk}$ . Electric fields  $\mathbf{E}_{ik}$  and  $\mathbf{E}_{jk}$  are produced by the corresponding charges  $q_{ik}$  and  $q_{jk}$  at the active site. The overall electric field of residue  $i$  at the active site,  $\mathbf{E}_{0i}$ , is obtained by summation of all  $\mathbf{E}_{ik}$  produced by the charges  $q_{ik}$  of the residue.

to residue  $j$  and propagated through the network to residue  $i$ . This response function is the basis of the dissipative electro-elastic network model (DENM) proposed here. At  $\omega = 0$ , equation (6) returns the standard result of the ENM. The connection between  $\chi_{ij}^{\alpha\beta}(0)$  and the statistical average given by equation (2) is stipulated by the fluctuation–dissipation theorem [52] requiring

$$\text{Re}[\chi_{ij}^{\alpha\beta}(0)] = \beta \langle \delta r_i^\alpha \delta r_j^\beta \rangle. \quad (7)$$

### 2.3. Electrostatic potential response function

The ENM has shown good performance in describing low-frequency structural/mechanical deformations of individual proteins and biomolecular assemblies [29]. Our goal here is to supplement these low-frequency motions with atomic partial charges to model the dynamics and statistics of electrostatic fluctuations at active sites of proteins. Coarse-grained electrostatics of proteins has been addressed in a number of recent papers [69–71], mostly dealing with long-range protein assembly and interactions [72]. These approaches also involve coarse-graining of the charge distribution [70, 71] or generating effective charges to reproduce the protein’s external field [73]. Given that we are interested in the electrostatic fluctuations at protein active sites, which are more sensitive to a local charge distribution, atomic charges from force-field potentials are more consistent with our purpose.

The electrostatic potential or field at a chosen point inside the protein are clearly affected by all partial atomic charges. For instance, when an atom with coordinate  $\mathbf{r}_0$  is chosen in the active site of the protein (Fe of the heme in our calculations below), one gets for the electrostatic potential at this point

$$\phi_0 = \sum_i \sum_k \frac{q_{ik}}{|\mathbf{r}_{ik} - \mathbf{r}_0|}, \quad (8)$$

where  $q_{ik}$  is the charge of atom  $k$  of residue  $i$  with the coordinate  $\mathbf{r}_{ik}$  (figure 1). In order to calculate the electrostatic potential and electric field, we therefore need all charges  $q_{ik}$  and vectorial

positions  $\mathbf{r}_{ik}$  produced by normal-mode displacements  $\delta \mathbf{r}_{ik}$  relative to equilibrium configuration  $\mathbf{r}_{ik,0}$ . The elastic network, however, coarse-grains the problem to equilibrium positions of beads  $\mathbf{r}_{i,0}$  and their normal-mode displacements  $\delta \mathbf{r}_{im}$  for each normal mode  $m$  with the eigenvalue  $\lambda_m$ . The information about librations of residues is lost when the network of beads replaces the protein and therefore is not present in our description of the protein electrostatics. For instance, librations of residue dipoles are absent from our modeling and only the displacements of dipoles are included.

Our approach starts with solving the standard ENM problem to obtain a set of displacements  $\delta \mathbf{r}_{im}$ . Each of these displacements is then applied to all charges  $q_{ik}$  of residue  $i$ , where the index  $k$  runs over all atoms of residue  $i$  (figure 1). Displacements  $\delta \mathbf{r}_{im}$  in turn produce dipole moments  $\delta \boldsymbol{\mu}_{ik}^{(m)} = q_{ik} \delta \mathbf{r}_{im}$ . The alteration of the electrostatic potential at the atom  $\mathbf{r}_0$  of the active site is

$$\delta \phi_0 = \sum_{m,i} \delta \mathbf{r}_{im} \cdot \mathbf{E}_{0i}, \quad (9)$$

where  $\mathbf{E}_{0i}$  is the electric field produced by all charges  $q_{ik}$  of residue  $i$  at the position  $\mathbf{r}_0$  of the active site atom. Its  $\alpha$ -Cartesian projection is given by the relation

$$E_{0i}^\alpha = \sum_k \frac{q_{ik} (r_0^\alpha - r_{ik}^\alpha)}{|\mathbf{r}_0 - \mathbf{r}_{ik}|^3}, \quad (10)$$

where  $\mathbf{r}_{ik}$  is the position of charge  $q_{ik}$ .

In the linear-response formalism [52], the displacements  $\delta \mathbf{r}_{im}$  are caused by the force acting on residue  $i$  from residue  $j$  in response to the field perturbation produced by the oscillatory probe charge  $q_0(t) = q_0 e^{i\omega t}$  placed at  $\mathbf{r}_0$ . The propagation of the oscillating force through the network, with the resulting oscillating residue displacements, is described by the response function given by equations (5) and (6). The result is

$$\delta r_i^\alpha(\omega) = \sum_j \chi_{ij}^{\alpha\beta}(\omega) \delta F_j^\beta(\omega), \quad (11)$$

where  $\delta \mathbf{F}_j(\omega) = -q_0 \mathbf{E}_{0j}$ . The frequency-dependent potential response function  $\chi_\phi(\omega)$  connects the potential to the charge perturbation,  $\delta \phi_0(\omega) = \chi_\phi(\omega) q_0$ . From the previous arguments, one finally finds

$$\chi_\phi(\omega) = - \sum_{i,j} E_{0j}^\alpha \chi_{ij}^{\alpha\beta}(\omega) E_{0i}^\beta. \quad (12)$$

### 2.4. Electrostatic field response function

The potential response function  $\chi_\phi(\omega)$  in (12) describes the oscillating electrostatic potential produced by the protein matrix in response to placing an oscillatory charge at the position in the active site where the potential is recorded (Fe ion in our MD simulations). Similarly, the response function of the electric field considers the field produced by the protein in response to an oscillating dipole moment  $\mathbf{m}_0(t) = \mathbf{m}_0 e^{i\omega t}$  placed in the active site. The deformation of the protein matrix induced by this probe dipole results in the electric field  $\mathbf{E}_0(\omega)$  at the same site. It can be found by summation over the contributions from all individual dipoles  $\delta \boldsymbol{\mu}_{ik}(\omega)$  arising from residues’ atomic charges

$$\mathbf{E}_0(\omega) = \sum_{ik} \mathbf{T}_{ik} \cdot \delta \boldsymbol{\mu}_{ik}(\omega). \quad (13)$$



Here,  $\mathbf{T}_{ik}$  is the dipolar tensor connecting the position of the active site with the charge  $q_{ik}$  of residue  $i$ . The dipole moment  $\delta\boldsymbol{\mu}_{ik}(\omega)$  at the position of  $q_{ik}$  is caused by the residue displacement  $\delta\mathbf{r}_{ik}(\omega)$  and can be expressed in terms of the displacement response function

$$\delta\boldsymbol{\mu}_{ik}(\omega) = q_{ik} \sum_j \boldsymbol{\chi}_{ij}(\omega) \cdot \delta\mathbf{F}_j(\omega), \quad (14)$$

where  $\delta\mathbf{F}_j(\omega) = \sum_k q_{jk} \mathbf{T}_{jk} \cdot \mathbf{m}_\omega$  is the force caused by the dipole  $\mathbf{m}_\omega$  at residue  $j$ . We therefore obtain a linear relation between the dipole  $\mathbf{m}_\omega$  placed at the active site and the electric field  $\mathbf{E}_0(\omega)$  produced by the protein in response to this perturbation. The proportionality coefficient between the external perturbation and the response is the response function given by the rank-2 tensor

$$\chi_E^{\alpha\beta}(\omega) = \sum_{i,j,k,l} q_{ik} T_{ik}^{\alpha\gamma} \chi_{ij}^{\gamma\delta}(\omega) T_{jl}^{\delta\beta} q_{jl}. \quad (15)$$

Here, as above, the summation is done over the repeated Greek indices denoting the Cartesian components of the corresponding tensors. We will be mostly interested in the trace of the tensor

$$\chi_E(\omega) = \chi_E^{\alpha\alpha}(\omega). \quad (16)$$

## 2.5. Response and correlation functions

The response functions calculated by the DENM formalism need to be related to time correlation functions supplied by the MD trajectories. The dynamics of a general dynamical variable  $\delta X(t) = X(t) - \langle X \rangle$  is characterized by the time self-correlation function

$$C_X(t) = \langle \delta X(t) \delta X(0) \rangle. \quad (17)$$

The normalized correlation function  $S_X(t) = C_X(t) / \langle (\delta X)^2 \rangle$  is related to the response function by the fluctuation–dissipation equation which forms the basis of our analysis [74]

$$\chi_X(\omega) = \beta \langle (\delta X)^2 \rangle [1 + i\omega S_X(\omega)]. \quad (18)$$

Here,  $S_X(\omega)$  is the Laplace–Fourier transform of  $S(t)$  [52].

Since our main focus is on variances of physical properties, we will define the generalized compliance for the variable  $X$  as follows:

$$\lambda_X = \beta \langle (\delta X)^2 \rangle / 2 = \int_0^\infty \chi_X''(\omega) (d\omega / \pi \omega). \quad (19)$$

The same property can be calculated from the  $\omega = 0$  value of the real part of the response function

$$\lambda_X = (1/2) \chi_X'(0). \quad (20)$$

In the case of residue displacements considered as variable  $X$  ( $X = r$ ),  $\lambda_r \propto C^{-1}$  is proportional to the inverse force constant of the network springs. As for the compliance of a macroscopic body, defined as the inverse of stiffness,  $\lambda_r$  depends on the shape and boundary conditions, in contrast to elastic moduli representing material properties.

For the elastic network, the generalized compliance reports on how the softness of the network affects the variance of the variable of interest. In the case of  $X = \phi$  reporting on the potential  $\phi$  at the position of the heme's iron,  $\lambda = e^2 \lambda_\phi$  is the reorganization energy of a half-redox reaction

corresponding to changing the oxidation state of the heme [7, 75]. The standard definition of half-reaction reorganization energy involves, instead of one atom, the distribution of the electronic density of the transferred electron over a few atoms of the active site. We simplify this problem here by assuming all electron charge  $e$  localized at the centroid of this charge density, the iron atom of the heme. The electrostatic potential at a single atom is a well-defined physical property, but its variance is only an approximate representation of the observable reorganization energy of changing the redox state [7].

We want to gain insight into the dissipative dynamics of the elastic network and its comparison with the dynamics calculated from MD simulations. The loss function  $\chi_X''(\omega)$  provides direct access to both the set of characteristic relaxation frequencies and their relative contributions. However, it does not weigh the low and high frequencies as they contribute to the variance in (19). Therefore, in addition to the loss function, we will consider the function

$$\alpha_X(\omega) = \frac{2}{\pi \omega} \frac{\chi_X''(\omega)}{\chi_X'(0)}. \quad (21)$$

This function is normalized,  $\int_0^\infty \alpha_X(\omega) d\omega = 1$ , and tends to the characteristic relaxation time  $\langle \tau \rangle$  in the  $\omega \rightarrow 0$  limit

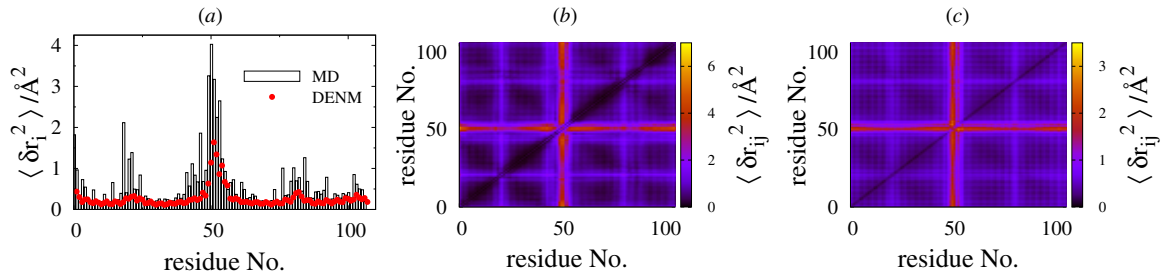
$$\lim_{\omega \rightarrow 0} \alpha_X(\omega) = (2/\pi) \langle \tau \rangle, \quad (22)$$

where

$$\langle \tau \rangle = \int_0^\infty S_X(t) dt. \quad (23)$$

## 2.6. Proteins used in the case studies and the simulation protocol

Three heme proteins, oxidized form of myoglobin (metmyoglobin, metMB, 1YMB), reduced state of cytochrome B562 (cytB, 256B), and reduced state of bovine heart cytochrome *c* (cytC, 2B4Z), were simulated by all-atom MD. The parameters of the proteins were taken from the CHARMM27 force field [76], and NAMD [77] was used for the trajectories production. Each protein was placed at the center of a cubic box with the side length of  $\simeq 108$  Å and solvated with TIP3P waters [78]. The number of waters used in simulations were 32891 (metMB), 33268 (cytB) and 33189 (cytC). The VMD [79] plugin Autoionize was used to neutralize the simulation cell by adding  $\text{Na}^+$  and  $\text{Cl}^-$  ions to bring the ionic strength of the solution to 0.1. Particle-mesh Ewald with the grid resolution  $< 1$  Å was used for electrostatic interactions, and all other non-bonded interactions were calculated within 12 Å cutoff. Following energy minimization, each protein was simulated for 5 ns in an NPT ensemble at  $P = 1$  atm and  $T = 300$  K. Temperature and pressure were controlled by using the Langevin dynamics with the damping coefficient of  $5 \text{ ps}^{-1}$ . The production NVE trajectory was started at the end of each 5 ns trajectory. The NVE ensemble is required for the proper sampling of the long-time dynamics since we found that using NPT and NVT ensembles artificially accelerates the observables relaxing on sub-nanosecond to nanosecond time-scale [80]. The integration step was 2 fs and simulation frames for the



**Figure 2.** Residue mean-square displacements (msd's)  $\langle \delta \mathbf{r}_i^2 \rangle$  (a) and variance-covariance matrix  $\langle \delta \mathbf{r}_i \cdot \delta \mathbf{r}_j \rangle$  of residues'  $C^\alpha$  carbons calculated from MD (b) and from DENM (c). The diagonals in (b) and (c) represent the residue msd's  $\langle \delta \mathbf{r}_i^2 \rangle$ . The network force constant in the DENM calculations is  $k_B T/C = 1 \text{ \AA}^2$ . The DENM values indicated by red squares in (a) need to be multiplied by about a factor of 2 in order to obtain the best-fit agreement with the MD data. They are separated for a better visibility in the plot.

analysis were saved each 0.05 ps. The simulation trajectories were 65 ns (metMB), 100 ns (cytC) and 123 ns (cytB).

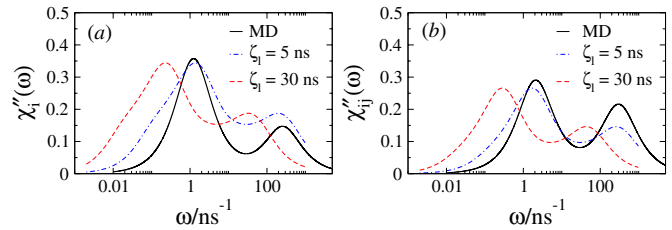
The force field standard parameters of the heme in the reduced state are taken from CHARMM27 [76]. The charge of iron is  $q_{\text{Fe}} = 0.24$  in the reduced state. No standard parameters are available for the oxidized heme required for the simulations of metmyoglobin. These force field parameters were taken from Autenrieth *et al* [81]. The atomic charge of oxidized Fe is  $q_{\text{Fe}} = 1.34$  in this parametrization. The iron atom is five-coordinate in its oxidized state. In order to allow the iron to shift out of the porphyrin plane, the harmonic force constant of the bond between Fe and N $_{\epsilon 2}$  of histidine 93 was adjusted to  $65 \text{ kcal mol}^{-1}$  [82, 83].

### 3. Results and discussion

#### 3.1. Statistics and dynamics of residue displacements

Consistent with many previous studies, we have found that elastic networks are capable of reproducing the basic pattern of the distribution of mean-square displacements (msd's) along the protein backbone. The left panel in figure 2 compares residue msd's from MD with DENM calculations, while two other panels show the maps of variance-covariance matrices. Overall, there is a good agreement between the alteration in residue displacements along the backbone, and corresponding cross-correlations, calculated from the two sets of data. However, the network force constant required for the best fit of the msd from MD simulations,  $\approx 0.3 \text{ kcal (mol \AA}^2)^{-1}$ , is somewhat lower than the value of  $\approx 1 \text{ kcal (mol \AA}^2)^{-1}$  typically adopted from fitting the B-factors from crystallography [20–22]. This value is also about a factor of 2 lower than  $0.6 \text{ kcal (mol \AA}^2)^{-1}$  adopted below to globally fit the variances of the field and electrostatic potential fluctuations from MD data for all three proteins studied here.

The fitting of the force constant to the absolute msd's from MD makes the network too soft. This outcome is expected since the elastic network cuts high-frequency vibrations from the density of states and thus underestimates the absolute magnitudes of the residue displacements [29]. It appears more consistent to use for the purpose of fitting only the portion of the msd related to protein's global motions. This component can in fact be separated from the vibrational part in the temperature dependence of the msd since global motions of the protein become observable at high temperatures, above



**Figure 3.** Loss spectra  $\chi''(\omega)/\chi'_i(0)$  of individual residues ( $i = 21$  (ASP)) (a) and of pairs of residues  $\chi''_{ij}(\omega)/\chi'_{ij}(0)$  ( $i = 21$  (ASP) and  $j = 84$  (VAL)) (b). The results have been obtained from MD (solid lines) and DENM (dashed lines). The calculations are done for cytB with the residues used in the calculations shown in figure 4. The DENM calculations are performed with  $\zeta_l = 30 \text{ ns}$  used for the electrostatic calculations in figure 5 (red dashed line) and with  $\zeta_l = 5 \text{ ns}$  (blue dash-dotted line). The rest of DENM parameters are  $\zeta_h = 0.006\zeta_l$ ,  $a = 0.35$ ,  $\varepsilon = 100$ .

the temperature of the protein dynamical transition [61]. This high-temperature portion of the protein msd is about a half of the total magnitude [84], which is close to the factor required to reconcile the force constants from electrostatic variances and msd's.

Figure 3 shows the dynamics of positions  $\mathbf{r}_i(t)$  of individual residues within the network and the dynamics of distances between residues  $\mathbf{r}_{ij}(t) = \mathbf{r}_i(t) - \mathbf{r}_j(t)$ . It reports the loss functions  $\chi''_i(\omega)$  and  $\chi''_{ij}(\omega)$  calculated from the self-correlation functions of individual residues

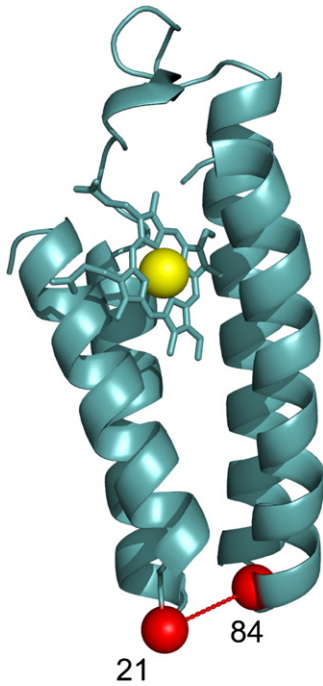
$$C_i(t) = \langle \delta \mathbf{r}_i(t) \cdot \delta \mathbf{r}_i(0) \rangle \quad (24)$$

and self-correlation functions of distances between the residues

$$C_{ij}(t) = \langle \delta \mathbf{r}_{ij}(t) \cdot \mathbf{r}_{ij}(0) \rangle. \quad (25)$$

These latter correlation functions are experimentally available from FRET [85] recording the evolution of the distance between two residues tagged with chromophores.

The comparison of  $\chi''_i(\omega)$  from DENM and MD is illustrated in figure 3(a) for only one residue,  $i = 21$  (ASP), belonging to the cytB protein. The pair of residues with  $i = 21$  (ASP) and  $j = 84$  (VAL) is used to calculate  $C_{ij}(t)$  in figure 3(b). The position of the pair in the backbone of cytB is shown in figure 4. Examples of calculations for several additional pairs of residues from cytB can be found in the Supplementary Data available at [stacks.iop.org/PB/9/036004/mmedia](http://stacks.iop.org/PB/9/036004/mmedia).



**Figure 4.** Sketch of cytb showing the residues (red) used to produce the correlation functions  $C_i(t)$  ( $i = 21$  (ASP)) and  $C_{ij}(t)$  ( $i, j = 21, 84$  (VAL), equations (24) and (25)) and corresponding loss spectra  $\chi''_i(\omega)$  and  $\chi''_{ij}(\omega)$  shown in figure 3. The residues used in producing  $\chi''_{ij}(\omega)$  are connected by the red solid line. The red spheres representing the two chosen residues are centered at their corresponding  $C^\alpha$  carbons. The Fe atom of the heme is rendered as a yellow sphere.

Both correlation functions,  $C_i(t)$  and  $C_{ij}(t)$ , report the dynamics on roughly two time-scales, seen as two peaks in their loss functions (figure 3). The two peaks represent, with varying weights, the faster local dynamics of individual residues and the slower global dynamics of the network.

The elastic network with single-exponential, Debye dynamics does not capture two time-scales of the dynamics from MD simulations. When the Debye memory kernel  $\zeta(\omega) = \zeta_0$  is used in the response function  $\chi_m(\omega)$  in (5) and (6), the loss function shows only one relaxation peak, slightly deformed from a simple Debye form by the distribution of the network's eigenvalues  $\lambda_m$ . A single friction coefficient evidently misses the fact that energy dissipation decreases with increasing frequency of vibrations. Since the protein network is characterized by two spring constants, lower for non-covalent neighbors within the cutoff and higher for covalent neighbors, it must possess at least two characteristic friction coefficients.

Friction kernels typically used in applications are phenomenological [52] and we also cannot offer a consistent theoretical formalism capturing the complex dynamics of residue displacements. We have therefore adopted a phenomenological response function best describing the MD results. Instead of searching for a functional form of  $\zeta(\omega)$  reproducing MD simulations, we have resorted to a two-Debye form for the entire response function of bead displacements in equation (5):

$$\chi_m(\omega) = \frac{a}{i\omega\zeta_h + \lambda_m} + \frac{1-a}{i\omega\zeta_l + \lambda_m}. \quad (26)$$

Here, the amplitude  $a$  represents the relative weight of the high-frequency dissipation with the high-frequency friction  $\zeta_h$ . Correspondingly,  $\zeta_l$  is the low-frequency friction, which is larger in magnitude. It is obvious that equation (26) satisfies the static limit  $\chi_m(0) = \lambda_m^{-1}$ .

The two-Debye form of  $\chi_m(\omega)$  is directly applied to calculate the response function of residue displacements

$$\chi_{ij}(\omega) = \chi_{ij}^{\alpha\alpha}(\omega), \quad (27)$$

where the function  $\chi_{ij}^{\alpha\beta}(\omega)$  is given by (6). The loss spectra of residue displacements show two peaks and qualitatively agree with the simulations. The positions and relative heights of the peaks can be adjusted by choosing  $\zeta_{h,l}$  and the amplitude  $a$  in (26), as is illustrated in figure 3(a) by dashed and dash-dotted lines. The model can therefore be well parameterized to reproduce the dynamics of displacements of a small subset of residues. However, it fails to discriminate between the dynamics of different residues across the protein backbone.

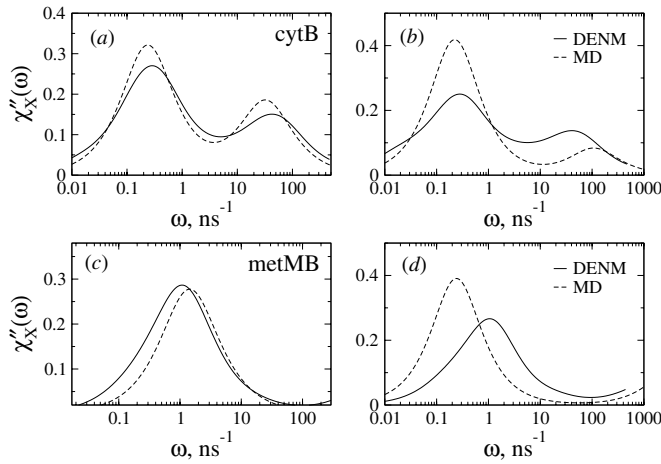
The loss functions calculated from MD for a number of residues (see the Supplementary Data available at [stacks.iop.org/PB/9/036004/mmedia](http://stacks.iop.org/PB/9/036004/mmedia)) display a similar two-peak pattern, but the weights and positions of their peaks vary among the residues. In contrast, the dynamics of displacements of individual residues in the elastic network are mostly driven by the global motions of the entire network. As a result, there is little difference between the loss functions of individual residues calculated with DENM. The same statement applies to the dynamics of distances between residues shown in figure 3(b). While one can reproduce the loss function of a chosen pair of residues by adjusting the parameters of  $\chi_m(\omega)$  in (26), these parameters do not translate into all pairs in the network and instead can only be viewed as an average representation of the pairs dynamics.

The fast component of the dynamics, prominent in the loss function  $\chi''_X(\omega)$ , is less significant in the function  $\alpha_X(\omega)$  due to the  $1/\omega$  scaling of  $\chi''_X(\omega)$  in (21). The function  $\alpha_X(\omega)$  is more relevant to problems related to the generalized compliance as defined by (19). The correct representation of the fast dynamics is therefore of lesser importance for these types of problems. The  $1/\omega$  scaling difference between  $\chi''_X(\omega)$  and  $\alpha_X(\omega)$  explains the relative success of the network models in reproducing the pattern of residue msd's (figure 2). Those are obtained by frequency integration of  $\alpha_r(\omega)$  and are less sensitive to the details of local dynamics of individual beads in the network represented by high-frequency peaks of their loss functions.

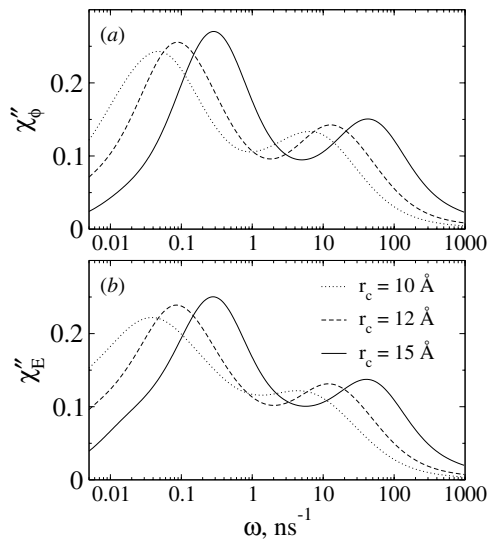
### 3.2. Dynamics of the electrostatic field and potential

Essentially all time correlation functions obtained here from MD simulations show a pattern that is roughly represented by a two-exponential decay, requiring in some cases a third decaying exponent for a better mathematical fit. The relative weights of the fast and slow components of the correlation functions differ depending on the observable. The loss function of the electric field  $\chi''_E(\omega)$  at the position of protein's heme is mostly single-exponential, with the low-frequency peak much exceeding the high-frequency one. This pattern is less uniform for the loss function of the electrostatic potential, showing





**Figure 5.** Loss spectra  $\chi''_X(\omega)/\chi'_X(0)$  (a,c) and  $\chi''_E(\omega)/\chi'_E(0)$  (b,d) for cytB (a,b) and metMB (c,d). The results are from MD trajectories (dashed lines) and from the DENM calculations (solid lines). The DENM calculations were performed with the two-Debye  $\chi_m(\omega)$  in (26). The two-Debye relaxation parameters are  $\zeta_l = 30$  ns,  $\zeta_h = 0.006\zeta_l$ ,  $a = 0.35$  for cytB and  $\zeta_l = 10$  ns,  $\zeta_h = 0.0002\zeta_l$ ,  $a = 0.35$  for metMB. The elastic network is defined with  $k_B T/C = 1 \text{ \AA}^2$ ,  $\varepsilon = 100$  and the cutoff radius of  $15 \text{ \AA}$ .



**Figure 6.** Loss spectra  $\chi''_\phi(\omega)/\chi'_\phi(0)$  (a) and  $\chi''_E(\omega)/\chi'_E(0)$  (b) for cytB at different cutoff radii  $r_c$  of the elastic network as indicated in the plot.

different weights of the low- and high-frequency peaks in  $\chi''_\phi(\omega)$  among the proteins [39]. Figure 5 shows the loss spectra  $\chi''_X(\omega)/\chi'_X(0)$  calculated from the DENM and MD for cytB and metMB proteins. Similarly to the case with the residue displacements, the low-frequency friction coefficient in (26) needs adjustment to reproduce the position of the main loss peak: the value of  $\zeta_l = 30$  ns was adopted for cytB and  $\zeta_l = 10$  ns was taken for metMB. Still the set of parameters taken to reproduce the loss spectrum of the electrostatic potential does not perform as well when applied to the electric field (cf figures 5(c) and (d)).

Figure 6 shows the dependence of the loss spectra on the cutoff radius  $r_c$  applied to the network of springs. There is a consistent shift of the loss peaks to slower dynamics with

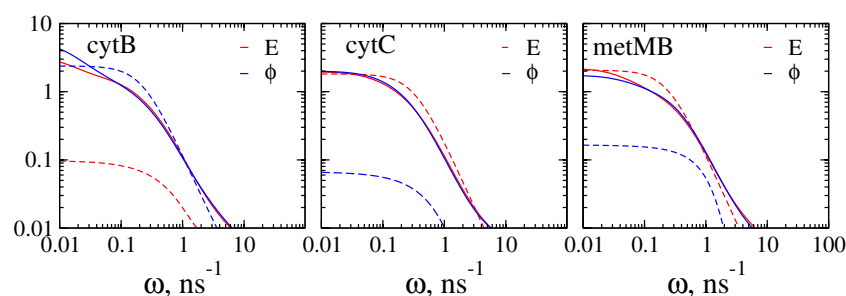
**Table 1.** Generalized compliances (equation (19)) for the electrostatic potential fluctuations  $\lambda_\phi$  (also known as reorganization energies of redox half-reaction  $\lambda = e^2\lambda_\phi$ , kcal mol<sup>-1</sup>) and for the electric force,  $F = eE$  ( $\lambda_F$ , kcal (mol  $\text{\AA}^2$ )<sup>-1</sup>) for the three proteins used in the case study; the network is assigned the force constant of  $C = 0.6$  kcal (mol  $\text{\AA}^2$ )<sup>-1</sup> for non-covalent springs and  $\varepsilon C$ ,  $\varepsilon = 100$ , for covalently bound neighbors. The cutoff radius is  $15 \text{ \AA}$ .

Protein	$e^2\lambda_\phi$ (DENM)	$e^2\lambda_\phi$ (MD)	$\lambda_F$ (DENM)	$\lambda_F$ (MD)
MetMB	85	228	436	243
CytB	115	136	9	17
CytC	141	159	18	83

decreasing  $r_c$ , but the resulting loss spectra are only slightly deformed. The formalism is thus stable to the cutoff variation and the cutoff distance can be used as an additional parameter to improve the agreement with MD results.

The relative contribution of the fast dynamics component is reduced in the spectral function  $\alpha_X(\omega)$  describing the frequency variation of the generalized compliance (equations (19) and (21)). Figure 7 compares  $\alpha_X(\omega)$  from DENM with MD data for all three globular proteins studied here. In order to show the ability of DENM to describe the entire set of MD data, a single set of network and friction parameters has been assigned to all three proteins. As expected, the performance of DENM is better for  $\alpha_X(\omega)$ . The average relaxation time  $\langle\tau\rangle$ , given by the  $\omega \rightarrow 0$  limit of  $\alpha_X(\omega)$  (equation (22)), is well reproduced by the network calculations. The deviations for some observables correspond to the cases when the dynamics are strongly dominated by its high-frequency component. For these cases, the fit can be improved by adjusting the weight  $a$  of the fast component in (26) to a higher value.

The loss functions  $\chi_X(\omega)$  shown in figures 3, 5 and 7 are normalized with the corresponding values of  $\chi'_X(0)$  and therefore are not affected by the magnitude of the spring constant  $C$  describing the non-covalent interactions within the cutoff distance  $r_c = 15 \text{ \AA}$ . The absolute values of the variances are, however, proportional to  $k_B T/C$ , as for instance in (2). We have adopted in our calculations the value of  $k_B T/C = 1 \text{ \AA}^2$ , which is equivalent to  $C \simeq 0.6$  kcal (mol  $\text{\AA}^2$ )<sup>-1</sup>. The results of calculating the generalized compliances  $\lambda_X$  for electrostatic observables (equations (19) and (20)) with this choice of the force constant are summarized in table 1. The agreement is semi-quantitative at best. While the value of the force constant can clearly be adjusted in each particular case, this change does not propagate into equally good agreement between  $\lambda_\phi$  and  $\lambda_F$ , where  $F = eE$  is the electrostatic force acting on the unit charge placed at Fe of the heme. The requirement to reproduce electrostatic forces might be more stringent than to describe the electrostatic potential because of a shorter range and higher sensitivity to the local structure of the former. Since the statistical variances listed in table 1 are not affected by the modeling of the dissipative dynamics, the properties of the network itself need to be further fine-tuned. The inclusion of softening of the residue displacements by electrostatic water fluctuations [16] appears to be the first avenue for the model improvement. This addition will allow one to accommodate for



**Figure 7.** Functions  $\alpha_X(\omega)$  (ns) for cytB, cytC and metMB from MD (dashed lines) and DENM (solid lines). The results for the electrostatic potential ( $\alpha_\phi(\omega)$ , blue) and electric field ( $\alpha_E(\omega)$ , red) are shown. A single set of network parameters is used for all three heme proteins:  $\zeta_l = 30$  ns,  $\zeta_h = 0.006\zeta_l$ ,  $a = 0.35$ ,  $\varepsilon = 100$ .

the effect of the interface, in addition to the motions dictated by the shape and mass distribution, which elastic models capture in the first place [31–33]. Incorporating the water response will effectively produce a more heterogeneous distribution of the network force constants.

#### 4. Conclusions

Network models of biomolecules are coarse-grained representations designed to describe their large-scale collective conformational motions [29]. They capture the basic topology and packing of residues in a biopolymer and robustly reproduce conformational global motions driven by the distribution of mass and shape [31–33]. Electrostatic properties is yet another area where coarse-graining of biomolecules might be efficient. The long range of Coulomb interactions effectively averages the details of the local structure out, suggesting that one can potentially describe the statistics and long-time dynamics of electrostatic fluctuations by global motions of charges assigned to the elastic network. The demonstration that this approach is in principle consistent with all-atom MD simulations is the main result of this study. Two components are critical to our approach: force-field atomic charges distributed at the residues of the network and two-exponential, overdamped dynamics assigned to each normal mode diagonalizing the network Hessian.

Harmonic mechanical motions of the network of beads clearly do not incorporate dissipative dynamics of biomolecules in solution [50, 51, 38] and much still needs to be done to achieve a physically consistent description of the protein dynamics. The elastic network employed here assigns weaker springs between all non-covalent neighbors within a cutoff radius and stronger springs between covalent neighbors. Correspondingly, two global friction coefficients are assigned to each normal mode diagonalizing the network Hessian. This phenomenological model qualitatively captures two-peak loss spectra of residue displacements and qualitatively similar loss spectra of the electrostatic potential and electric field. However, the characteristic adjustable friction coefficients used for the electrostatic fluctuations are not transferrable to the network displacements. A new set

of parameters is needed when each property is considered separately.

#### Acknowledgments

This research was supported by the National Science Foundation (DVM, CHE-0910905). CPU time was provided by the National Science Foundation through TeraGrid resources (TG-MCB080116N).

#### References

- [1] Fenimore P W, Frauenfelder H, McMahon B H and Young R D 2004 Bulk-solvent and hydration-shell fluctuations, similar to  $\alpha$  and  $\beta$ -fluctuations in glasses, control protein motion and functions *Proc. Natl Acad. Sci.* **101** 14408–13
- [2] Henzler-Wildman K and Kern D 2007 Dynamic personalities of proteins *Nature* **450** 964
- [3] Frauenfelder H, Chen G, Berendzen J, Fenimore P W, Jansson H, McMahon B H, Stroer I R, Swenson J and Young R D 2009 A unified model of protein dynamics *Proc. Natl Acad. Sci.* **106** 5129–34
- [4] Marlow M S, Dogan J, Frederick K K, Valentine K G and Joshua Wand A 2010 The role of conformational entropy in molecular recognition by calmodulin *Nature Chem. Biol.* **6** 352–8
- [5] Wolynes P G, Onuchic J N and Thirumalai D 1995 Navigating the folding routes *Science* **267** 1619–20
- [6] Min W, English B P, Luo G, Cherayil B J, Kou S C and Xie X S 2005 Fluctuating enzymes: lessons from single-molecule studies *Acc. Chem. Res.* **38** 923–31
- [7] Marcus R A and Sutin N 1985 Electron transfer in chemistry and biology *Biochim. Biophys. Acta* **811** 265–322
- [8] LeBarb D N and Matyushov D V 2010 Protein–water electrostatics and principles of bioenergetics *Phys. Chem. Chem. Phys.* **12** 15335
- [9] Khodadadi S, Roh J H, Kisliuk A, Mamontov E, Tyagi M, Woodson S A, Briber R M and Sokolov A P 2010 Dynamics of biological macromolecules: not a simple slaving by hydration water *Biophys. J.* **98** 1321–6
- [10] Ebbinghaus S, Kim S J, Heyden M, Yu X, Heugen U, Gruebele M, Leitner D M and Havenith M 2007 An extended dynamical hydration shell around proteins *Proc. Natl Acad. Sci.* **104** 20749–52
- [11] Perticaroli S, Comez L, Paolantoni M, Sassi P, Lupi L, Fioretto D, Paciaroni A and Morresi A 2010 Broadband depolarized light scattering study of diluted protein aqueous solutions *J. Phys. Chem. B* **114** 8262–69

- [12] Warshel A, Sharma P K, Kato M and Parson W W 2006 Modeling electrostatic effects in proteins *Biochim. Biophys. Acta* **1764** 1647–76
- [13] Richardson J S and Richardson D C 2002 Natural  $\beta$ -sheet proteins use negative design to avoid edge-to-edge aggregation *Proc. Natl Acad. Sci.* **99** 2754
- [14] Lawrence M S, Phillips K J and Liu D R 2007 Supercharging proteins can impart unusual resilience *J. Am. Chem. Soc.* **129** 10110
- [15] Pace C N, Grimsley G R and Scholtz J M 2009 Protein ionizable groups: pK values and their contribution to protein stability and solubility *J. Biol. Chem.* **284** 13285
- [16] Matyushov D V and Morozov A Y 2011 Electrostatics of the protein–water interface and the dynamical transition in proteins *Phys. Rev. E* **84** 011908
- [17] Andreatta D, Pérez J L, Kovalenko S A, Ernsting N P, Murphy C J, Coleman R S and Berg M A 2005 Power-law solvation dynamics in DNA over six decades in time *J. Am. Chem. Soc.* **127** 7270
- [18] Tripathy J and Beck W F 2010 Nanosecond-regime correlation time scales for equilibrium protein structural fluctuations of metal-free cytochrome c from picosecond time-resolved fluorescence spectroscopy and the dynamic Stokes shift *J. Phys. Chem. B* **114** 15958–68
- [19] Zhang L, Wang L, Kao Y-T, Qiu W, Yang Y, Okobiah O and Zhong D 2007 Mapping hydration dynamics around a protein surface *Proc. Natl Acad. Sci.* **104** 18461
- [20] Tirion M M 1996 Large amplitude elastic motions from single-parameter, atomic analysis *Phys. Rev. Lett.* **77** 1905–8
- [21] Atilgan A R, Durell S R, Jernigan R L, Demirel M C, Keskin O and Bahar I 2001 Anisotropy of fluctuation dynamics of proteins with an elastic network model *Biophys. J.* **80** 505–15
- [22] Tama F and Sanejouand Y H 2001 Conformational change of proteins arising from normal mode calculations *Protein Eng.* **14** 1–6
- [23] Tama F, Valle M, Frank J and Brooks C L 2003 Dynamic reorganization of the functionally active ribosome explored by normal mode analysis and cryo-electron microscopy *Proc. Natl Acad. Sci. USA* **100** 9319–23
- [24] Tama F and Brooks C L 2005 Diversity and identity of mechanical properties of icosahedral viral capsids studied with elastic network normal mode analysis *J. Mol. Biol.* **345** 299–314
- [25] Lu M, Poon B and Ma J 2006 A new method for coarse-grained elastic normal-mode analysis *J. Chem. Theory Comput.* **2** 464–71
- [26] Moritsugu K and Smith J C 2007 Coarse-grained biomolecular simulation with reach: realistic extension algorithm via covariance hessian *Biophys. J.* **93** 3460–9
- [27] Lu M and Ma J 2008 A minimalist network model for coarse-grained normal mode analysis and its application to biomolecular x-ray crystallography *Proc. Natl Acad. Sci.* **105** 15358–63
- [28] Riccardi D, Cui Q and Phillips G N 2009 Application of elastic network models to proteins in the crystalline state *Biophys. J.* **96** 464–75
- [29] Bahar I, Lezon T R, Yang L-W and Eyal E 2010 Global dynamics of proteins: bridging between structure and function *Ann. Rev. Biophys.* **39** 23–42
- [30] Romo T D and Grossfield A 2011 Validating and improving elastic network models with molecular dynamics simulations *Proteins: Struct. Funct. Bioinformatics* **79** 23–34
- [31] Halle B 2002 Flexibility and packing in proteins *Proc. Natl Acad. Sci.* **99** 1274–9
- [32] Lu M and Ma J 2005 The role of shape in determining molecular motions *Biophys. J.* **89** 2395–401
- [33] Tama F and Brooks C L 2006 Symmetry, form, and shape: guiding principles for robustness in macromolecular machines *Annu. Rev. Biophys. Biomol. Struct.* **35** 115–33
- [34] Micheletti C, Carloni P and Maritan A 2004 Accurate and efficient description of protein vibrational dynamics: comparing molecular dynamics and Gaussian models *Proteins* **55** 635–45
- [35] Petrone P and Pande V S 2006 Can conformational change be described by only a few normal modes? *Biophys. J.* **90** 1583–93
- [36] Lyman E, Pfaendtner J and Voth G A 2008 Systematic multiscale parametrization of heterogeneous elastic network models of proteins *Biophys. J.* **95** 4183
- [37] Hinsen K and Kneller G R 1999 A simplified force field for describing vibrational protein dynamics over the whole frequency range *J. Chem. Phys.* **111** 10766–9
- [38] Miller B T, Zheng W, Venable R M, Pastor R W and Brooks B R 2008 Langevin network model of myosin *J. Phys. Chem. B* **112** 6274–81
- [39] Matyushov D V 2011 Nanosecond Stokes shift dynamics, dynamical transition, and gigantic reorganization energy of hydrated heme proteins *J. Phys. Chem. B* **115** 10715–24
- [40] LeBard D N and Matyushov D V 2009 Energetics of bacterial photosynthesis *J. Phys. Chem. B* **113** 12424–37
- [41] Tozzini V 2010 Minimalist models for proteins: a comparative analysis *Q. Rev. Biophys.* **43** 333–71
- [42] Hinsen K 2008 Structural flexibility in proteins: impact of the crystal environment *Bioinformatics* **24** 521–8
- [43] Gerek Z N and Ozkan S B 2011 Change in allosteric network affects binding affinities of PDZ domains: analysis through perturbation response scanning *PLoS Comput. Biol.* **7** e1002154
- [44] Ming D and Wall M E 2005 Allostery in a coarse-grained model of protein dynamics *Phys. Rev. Lett.* **95** 198103
- [45] Gerek Z N, Keskin O and Ozkan S B 2009 Identification of specificity and promiscuity of PDZ domain interactions through their dynamic behavior *Proteins* **77** 796–811
- [46] Smith J C 1991 Protein dynamics: comparison of simulations with inelastic neutron scattering experiments *Q. Rev. Biophys.* **24** 227
- [47] Cametti C, Marchetti S, Gambi C M C and Onori G 2011 Dielectric relaxation spectroscopy of lysozyme aqueous solutions: analysis of the delta-dispersion and the contribution of the hydration water *J. Phys. Chem. B* **115** 7144–53
- [48] Lamm G and Szabo A 1986 Langevin modes of macromolecules *J. Chem. Phys.* **85** 7334
- [49] Essiz S G and Coalson R D 2009 Dynamic linear response theory for conformational relaxation of proteins *J. Phys. Chem. B* **113** 10859–69
- [50] Ansari A 1999 Langevin modes analysis of myoglobin *J. Chem. Phys.* **110** 1774–80
- [51] Erkip A and Erman B 2004 Dynamics of large-scale fluctuations in native proteins. Analysis based on harmonic inter-residue potentials and random external noise *Polymer* **45** 641–8
- [52] Hansen J P and McDonald I R 2003 *Theory of Simple Liquids* (Amsterdam: Academic)
- [53] Kubo R 1966 The fluctuation–dissipation theorem *Rep. Prog. Phys.* **29** 255–84
- [54] Zwanzig R 2001 *Nonequilibrium Statistical Mechanics* (Oxford: Oxford University Press)
- [55] Nordholm S and Zwanzig R 1975 A systematic derivation of exact generalized Brownian motion theory *J. Stat. Phys.* **13** 347–71
- [56] Schweizer K S 1989 Microscopic theory of the dynamics of polymeric liquids: general formulation of a mode–mode-coupling approach *J. Chem. Phys.* **91** 5802–21

- [57] Karplus M 2000 Aspects of protein reaction dynamics: deviations from simple behavior *J. Phys. Chem. B* **104** 11–27
- [58] Pontiggia F, Colombo G, Micheletti C and Orlandi H 2007 Anharmonicity and self-similarity of the free energy landscape of protein G *Phys. Rev. Lett.* **98** 048102
- [59] Achterhold K, Keppler C, Ostermann A, van Bürck U, Sturhahn W, Alp E E and Parak F G 2002 Vibrational dynamics of myoglobin determined by the phonon-assisted Mössbauer effect *Phys. Rev. E* **65** 051916
- [60] Achterhold K and Parak F G 2003 Protein dynamics: determination of anisotropic vibrations at the haem iron myoglobin *J. Phys.: Condens. Matter* **15** S1683–92
- [61] Gabel F, Bicoût D, Lehnert U, Tehei M, Weik M and Zaccai G 2002 Protein dynamics studied by neutron scattering *Q. Rev. Biophys.* **35** 327–67
- [62] Doster W 2010 The protein–solvent glass transition *Biochim. Biophys. Acta* **1804** 3–14
- [63] Diehl M, Doster W, Petry W and Schober H 1997 Water-coupled low-frequency modes of myoglobin and lysozyme observed by inelastic neutron scattering *Biophys. J.* **73** 2726–32
- [64] Marconi M, Cornicchi E, Onori G and Paciaroni A 2008 Comparative study of protein dynamics in hydrated powders and in solution: a neutron scattering investigation *Chem. Phys.* **345** 224–9
- [65] Min W, Luo G, Cherayil B J, Kou S C and Sunney Xie X 2005 Observation of a power-law memory kernel for fluctuations within a single protein molecule *Phys. Rev. Lett.* **94** 198302
- [66] Janežič D, Venable R M and Brooks B R 1995 Harmonic analysis of large systems: III. Comparison with molecular dynamics *J. Comput. Chem.* **16** 1554–66
- [67] Soheilifard R, Makarov D E and Rodin G J 2011 Rigorous coarse-graining for the dynamics of linear systems with applications to relaxation dynamics in proteins *J. Chem. Phys.* **135** 054107
- [68] Henery R J 1971 The generalized Langevin equation and the fluctuation–dissipation theorems *J. Phys. A: Gen. Phys.* **4** 685–94
- [69] Skepö M, Linse P and Arnebrant T 2006 Coarse-grained modeling of proline rich protein I (prp-1) in bulk solution and adsorbed to a negatively charged surface *J. Phys. Chem. B* **110** 12141–8
- [70] Pizzitutti F, Marchi M and Borgis D 2007 Coarse-graining the accessible surface and the electrostatics of proteins for protein–protein interactions *J. Chem. Theory Comput.* **3** 1867–76
- [71] Leherter L and Vercauteren D P 2009 Coarse point charge models for proteins from smoothed molecular electrostatic potentials *J. Chem. Theory Comput.* **5** 3279–98
- [72] Dong F, Olsen B and Baker N A 2008 Computational methods for biomolecular electrostatics *Methods in Cell Biology* vol 84 (Amsterdam: Elsevier) p 843
- [73] Berardi R, Muccioli L, Orlandi S, Ricci M and Zannoni C 2004 Mimicking electrostatic interactions with a set of effective charges: a genetic algorithm *Chem. Phys. Lett.* **389** 373–8
- [74] Chaikin P M and Lubensky T C 1995 *Principles of Condensed Matter Physics* (Cambridge: Cambridge University Press)
- [75] LeBard D N and Matyushov D V 2008 Redox entropy of plastocyanin: developing a microscopic view of mesoscopic solvation *J. Chem. Phys.* **128** 155106
- [76] MacKerell A D *et al* 1998 All-atom empirical potential for molecular modeling and dynamics studies of proteins *J. Phys. Chem. B* **102** 3586–616
- [77] Phillips J C, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel R D, Kale L and Schulten K 2005 Scalable molecular dynamics with NAMD *J. Comput. Chem.* **26** 1781–802
- [78] Jorgensen W L, Chandrasekhar J, Madura J D, Impey R W and Klein M L 1983 Comparison of simple potential functions for simulating liquid water *J. Chem. Phys.* **79** 926–35
- [79] Humphrey W, Dalke A and Schulten K 1996 VMD—visual molecular dynamics *J. Mol. Graph.* **14** 33–8
- [80] LeBard D N and Matyushov D V 2010 Ferroelectric hydration shells around proteins: electrostatics of the protein–water interface *J. Phys. Chem. B* **114** 9246–58
- [81] Autenrieth F, Tajkhorshid E, Baudry J and Luthney-Schulten Z 2004 Classical force field parameters for the heme prosthetic group of cytochrome *c* *J. Comput. Chem.* **25** 1613
- [82] Meuwly M, Becker O M, Stote R and Karplus M 2002 No rebinding to myoglobin: a reactive molecular dynamics study *Biophys. Chem.* **98** 183
- [83] Zhang Y and Straub J E 2009 Diversity of solvent dependent energy transfer pathways in heme proteins *J. Phys. Chem. B* **113** 825
- [84] Zaccai G 2000 How soft is a protein? A protein dynamics force constant measured by neutron scattering *Science* **288** 1604–7
- [85] Weiss S 2000 Measuring conformational dynamics of biomolecules by single molecule fluorescence spectroscopy *Nature Struct. Biol.* **7** 724–9