Solvated dissipative electro-elastic network model of hydrated proteins

Daniel R. Martin and Dmitry V. Matyushov

Citation: The Journal of Chemical Physics **137**, 165101 (2012); doi: 10.1063/1.4759105
View online: http://dx.doi.org/10.1063/1.4759105
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/137/16?ver=pdfcov
Published by the AIP Publishing

Solvated dissipative electro-elastic network model of hydrated proteins

Daniel R. Martin and Dmitry V. Matyushov[a)]
Center for Biological Physics, Arizona State University, PO Box 871504, Tempe, Arizona 85287-1504, USA

(Received 13 July 2012; accepted 26 September 2012; published online 22 October 2012)

Elastic network models coarse grain proteins into a network of residue beads connected by springs. We add dissipative dynamics to this mechanical system by applying overdamped Langevin equations of motion to normal-mode vibrations of the network. In addition, the network is made heterogeneous and softened at the protein surface by accounting for hydration of the ionized residues. Solvation changes the network Hessian in two ways. Diagonal solvation terms soften the spring constants and off-diagonal dipole-dipole terms correlate displacements of the ionized residues. The model is used to formulate the response functions of the electrostatic potential and electric field appearing in theories of redox reactions and spectroscopy. We also formulate the dielectric response of the protein and find that solvation of the surface ionized residues leads to a slow relaxation peak in the dielectric loss spectrum, about two orders of magnitude slower than the main peak of protein relaxation. Finally, the solvated network is used to formulate the allosteric response of the protein to ion binding. The global thermodynamics of ion binding is not strongly affected by the network solvation, but it dramatically enhances conformational changes in response to placing a charge at the active site of the protein. [http://dx.doi.org/10.1063/1.4759105]

I. INTRODUCTION

Folding a globular protein in water largely places polar/ionized residues to its surface, while moving the non-polar residues to its core. The resulting structure is not unique, and a number of conformations with close energy minima always exist. Conformational changes are required for function. They are achieved by either populating the existing (quasi)stable states (sampling of pre-existing equilibria[1–4]) or by shifting the existing minimum-energy conformation to a new configuration minimum upon perturbation, such as ligand binding (induced fit mechanism[5,6]).

Conformational transitions involve several types of free energy penalty. The free energy of elastic deformation relative to the native structure, involving large-scale alteration of the protein shape, is the most prominent penalty.[7] This is not the only long-ranged component of the overall protein's thermodynamics since electrostatic interactions are also involved in several ways. Changing the protein conformation not only alters the interactions between its atomic charges, but also, to a significant extent, the free energy of solvation of these charges by hydration water. Water clearly affects the flexibility of proteins.[8] As a fast, highly polar subsystem, it follows adiabatically the large-scale protein motions, continuously stretching and loosening the protein structure by strong protein-water solvation forces. It lowers the barriers of transitions between the local minima of the rugged landscape at the energy bottom of the native basin of attraction,[9,10] accelerating the rate of conformational changes.[11] When dried, proteins stiffen and their relaxation time, as probed by dielectric spectroscopy, increases by about six orders of magnitude.[12]

The goal of this paper is to develop an efficient computational algorithm to include hydration in calculations of conformational flexibility of large protein complexes. Our starting point is to coarse-grain the protein into an elastic network of beads, a formalism known as elastic network model (ENM).[13–15] These types of models aim at calculating the elastic energy of deformation near the equilibrium structure and the directions of normal-mode displacements corresponding to the slowest normal-mode vibrations. The typical coarse-graining is achieved on the scale of a single residue by replacing it with a single rigid bead. The beads are then connected by elastic springs physically capturing the connectivity, shape, and packing of the residues in the folded protein structure.

The ENM coarse-graining of the elastic energy has proven to be very successful.[16–22] Global elastic deformations of a protein are mostly affected by its shape and mass distribution.[23–25] Electrostatics is another good candidate for coarse-graining. Coulomb forces are long-ranged and effectively average out the variations of the local structure. The final outcome for the free energy of electrostatic interactions is mostly determined by the overall density and distribution of the protein charge and dipolar polarization of hydration water. This physical reality is addressed by generalized Born solvation models, which design fast computational algorithms to calculate the free energy of electrostatic solvation.[26]

The problem addressed here is twofold. First, we want to re-normalize the elastic network by water's hydration. Given the large free energy of hydration of the surface residues, the network of beads is expected to be softer at the interface. We achieve this goal here by integrating out the dipolar polarization of the hydration water using formalisms developed in the liquid-state theory of polar liquids.[27,28] The result is an analytical model, a solvated dissipative electro-elastic network model (sDENM), which assigns lower force constants to springs attached to the ionized surface residues. The second issue is the calculation of the response functions related

a)Electronic mail: dmitrym@asu.edu.

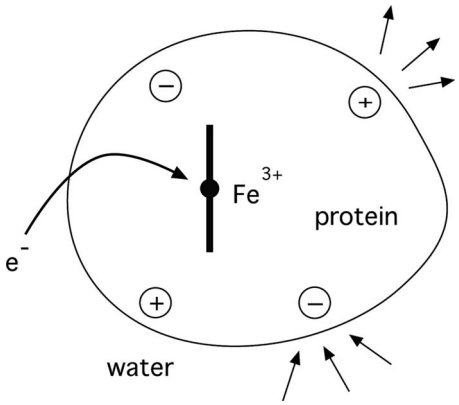

FIG. 1. Cartoon displaying the electrostatic perturbation induced by a half redox reaction of transferring electron to the heme in the protein used as an example. The elastic deformation of the protein not only shifts the ionized surface residues, shown by charges at the surface, but also results in adiabatic movements of water dipoles solvating them (shown by arrows).

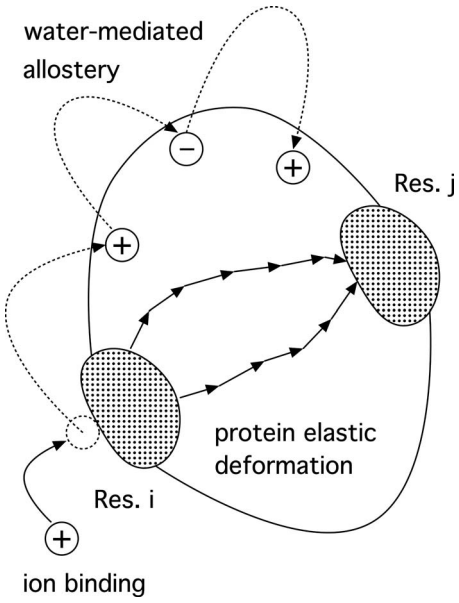

FIG. 2. Cartoon showing propagation of piezoelectric perturbation caused by binding an ion to residue *i* and producing a displacement of residue *j*. The electric force exerted by the ion is propagated throughout the protein as an elastic deformation indicated by chains of arrows. In contrast, the displacements of ionized surface residues are propagated as water-mediated, dipole-correlated surface motions. Ionized surface residues combine into a global, correlated net for transmitting signals, which does not necessarily require a specific binding site.

to problems affected by the protein electrostatics. Here, we consider three types of problems: (i) electrostatic response to a probe charge or dipole placed inside the protein, (ii) response of the protein to a uniform external field (dielectric spectroscopy), and (iii) elastic response at a given site of the protein to altering the charge state of a distant residue (allosteric action).

The first type of problems appear in redox reactions involving proteins[29] and in optical and IR spectroscopy[30–32] when the position of a spectral line is affected by the local electric field. In redox reactions, the electron is transferred by tunneling from an electron donor to an active site (shown as protein heme in Fig. 1). The dynamics of these processes, and the activation barrier required to produce resonant conditions for electron tunneling, can be calculated from the electrostatic response function $\chi_\phi(\omega)$. It arises not only from an elastic deformation, shifting the atomic charges of the protein, caused by transferring the electron, but also from the change in the positions and orientations of water dipoles hydrating the surface residues (Fig. 1). For spectroscopic applications, it is the dipole moment of the chromophore that is altered by light absorption. The corresponding response functions $\chi_E(\omega)$ is the one of the electric field acting on the chromophore's dipole and responsible for spectral solvatochromism.[33]

The second problem addressed here is the dynamic susceptibility of the protein to a uniform external electric field produced in the dielectric spectroscopy experiment.[34] Dielectric spectroscopy of partially wet protein powders has identified a number of generic relaxation peaks, the assignment of which has been problematic.[12,35] The solvent-renormalized network model developed here results in three relaxation peaks of the protein assigned to fast backbone vibrations (fastest), the low-frequency dissipative motions altering the shape of the protein (main peak), and the slow motions of highly solvated charged residues with significant solvent exposure (slowest).

Finally, the last type of problems considered here is the allosteric response.[2,4,36] It specifies the alteration in the structure of the protein produced by a perturbation at a distant site.[3] The perturbation can be achieved by localized ligand binding, often carrying a charge (Fig. 2). Allosteric signaling often involves oligomeric proteins, although single-domain proteins also display allostery.[2,37,38] Given that two equilibrium conformations are involved, two equilibrium sets of atomic coordinates need to be considered for a full description of allosteric signaling. The barrier to the transition between the two equilibrium structures is the free energy of the protein elastic deformation, which can be approximated as crossing of two harmonic elastic wells.[7,39–42] The elastic free energy is quadratic as a function of global normal-mode displacements, but can involve anharmonic effects and will also change its functional form to a linear function when localized (cracking) excitations, corresponding to local unfolding events, are produced.[7]

In the harmonic approximation, the lowest elastic barrier is reached along the lowest curvature path on the free energy surface against deformation, i.e., the lowest frequency of the elastic vibration. The allosteric pathways are therefore often associated with the lowest frequencies of harmonic motions near the two equilibrium structures.[9,43] The dissipative dynamics of these motions can therefore be explored in the framework of response functions referring to a single equilibrium conformation.[44] The formulation of such response functions in the framework of a dissipative electro-elastic network is our purpose here.

We calculate the dynamics of displacement of a distant residue in response to changing the charge at the binding site of an allosteric protein. The main question here is how water modifies the response. We find that solvation of ionized residues provides a potential mechanism alternative to the typically anticipated elastic propagation of the perturbation.

Despite differences in packing and connectivity of residues in different regions of a folded protein, elastic response tends to be non-specific, spreading out over the entire volume of the protein (Fig. 2). In contrast, a net of ionized residues potentially provides an alternative, surface-bound propagation of the perturbation by water-mediated allostery. The water-mediated cross-coupling between the displacements of surface charges scales as r^{-3}. Therefore, an alteration of the charge at a binding site can propagate large distances over the network of surface residues, instead of, or an addition to, the bulk elastic deformation. The allosteric response can then be channeled to a site where a conformational change is required for function.

II. MODEL

The Hamiltonian of the protein hydrated by polar water can be generally written in the following form:

$$H = E(\mathbf{R}) - \sum_{i,j} \mathbf{E}_{ij} \cdot \mathbf{m}_j. \tag{1}$$

Here, $E(\mathbf{R})$ is the solvent-unperturbed Hamiltonian of the protein depending on the manifold of atomic coordinates $\mathbf{R}$. The potential energy $E(\mathbf{R})$ includes all bonded and non-bonded interactions of the protein in the ionization state of the aqueous solution. Further, $\mathbf{E}_{ij}$ is the electric field acting from residue i of the protein on dipole moment $\mathbf{m}_j$ of water. Both the protein coordinates $\mathbf{R}$ and the water dipoles $\mathbf{m}_j$ fluctuate with the instantaneous configuration of the protein-water system; summation over all residues $i = 1, \ldots, N$ and all waters $j = 1, \ldots, N_s$ is taken in Eq. (1).

Several approximations need to be made in the transition from the general Hamiltonian in Eq. (1) to a solvated elastic network. The first approximation is the assumption that an equilibrium configuration of the protein atomic coordinates is available from the structural data, and quadratic expansion in atomic displacements can be done around it. The model thus deals with one conformational state of the protein only and has nothing to say about transitions between distinct protein conformations.

The first step in coarse-graining the model is to replace the collection of protein atomic coordinates with a collection of beads. We will follow here the standard approach[13–15] of representing each residue with a single bead, with its position given by the coordinates of the C^α atom. The quadratic expansion of $E(\mathbf{R})$ in small displacements $\delta r_i^\alpha = r_i^\alpha - r_{0i}^\alpha$ of individual beads relative to equilibrium positions $\mathbf{r}_{0i}$ leads to the relation

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

in which $H_{ij}^{\alpha\beta}$ is a $3N \times 3N$ Hessian matrix and C is the scaling force constant; α, β indicate the Cartesian projections, and summation over repeated Greek indices is assumed.

The electrostatic component of the problem is represented by the standard formulation of atomic force fields. This implies that each atom of the protein carries the charge q_{ik}, where $i = 1, \ldots, N$ numbers the residues and k represents an atom within residue i. The linear, in the displacements $\delta \mathbf{r}_i$, expansion of the protein-water interaction term in Eq. (1) results in the following equation:

$$\delta \mathbf{E}_i \cdot \mathbf{m}_j = \sum_k q_{ik} \delta \mathbf{r}_i \cdot \mathbf{T}_{ij} \cdot \mathbf{m}_j, \tag{3}$$

where $\mathbf{T}_{ij} = -\nabla_i \nabla_j |\mathbf{r}_{0i} - \mathbf{r}_j|^{-1}$ is the dipolar tensor connecting the C^α of residue i with the dipole of water j.

In the elastic network constructed here, all atoms of a given residue experience a uniform displacement $\delta \mathbf{r}_i$ from their equilibrium positions. The librations of the residues are therefore neglected. This approximation leads to a significant simplification in Eq. (3) since only charged residues with $\sum_k q_{ik} = q_i \neq 0$ contribute to the sum. Clearly, uniform displacements of only charged residues contribute to the creation of the dipole moment fluctuation $\delta \boldsymbol{\mu}_i = q_i \delta \mathbf{r}_i$. We therefore get for the energy of the protein-water system

$$H = (C/2) \sum_{i,j} H_{ij}^{\alpha\beta} \delta r_i^\alpha \delta r_j^\beta - \sum_i q_i \delta r_i^\alpha T_{ij}^{\alpha\beta} m_j^\beta. \tag{4}$$

We now proceed to calculating the free energy of the hydrated protein by tracing out the fluctuations of the dipole moments of water. Adiabatic approximation is assumed at this step, that is we assume that water is a fast subsystem, equilibrating to each instantaneous configuration of the network of beads. Since only low-frequency dissipative motions of the protein are modeled by the elastic network (translations and rotations of the protein as a whole are excluded), this approximation is expected to be accurate.

Averaging over the configurations of the dipole moments of water produces partial free energy, i.e., free energy depending on the manifold of instantaneous displacements $\delta \mathbf{r}_i$. If the fluctuations of the dipolar polarization field of water are Gaussian, as is mostly the case for polar liquids,[45–47] this free energy is given by the following equation:

$$F = (C/2) \sum_{ij} \tilde{H}_{ij}^{\alpha\beta} \delta r_i^\alpha \delta r_j^\beta, \tag{5}$$

where the Hessian matrix, renormalized by solvation, becomes

$$\tilde{H}_{ij}^{\alpha\beta} = H_{ij}^{\alpha\beta} - C^{-1} \kappa_{ij}^{\alpha\beta} q_i q_j. \tag{6}$$

In Eq. (6), $\boldsymbol{\kappa}_{ij}$ is the rank-2 tensor representing the dipolar response of the solvent to a dipole $\delta \boldsymbol{\mu}_j$ created at the residue j. The dipolar polarization field induced in the solvent in response to this perturbation in turn propagates to induce the dipole $\delta \boldsymbol{\mu}_i$ at residue i. The corresponding free energy cost contributes to the renormalization of the Hookean force constants for the residues involved, as represented by the second term in Eq. (6).

The physical meaning of the solvation terms in Eqs. (5) and (6) is quite clear. At $i = j$, the second term in Eq. (6) represents the solvation free energy of the fluctuation dipole $\delta \boldsymbol{\mu}_i$. Correspondingly, $i \neq j$ terms are the water-mediated couplings of the dipolar fluctuations through the solvent polarization. Note that direct Coulomb interactions between $\delta \boldsymbol{\mu}_i$ and $\delta \boldsymbol{\mu}_j$ are a part of $E(\mathbf{R})$ in Eq. (1) and are not included in $\boldsymbol{\kappa}_{ij}$.

The dipolar, off-diagonal interactions in $\boldsymbol{\kappa}_{ij}$ propagate the elastic perturbation of the protein through the solvent by

excitation hopping between the ionized residues. This mechanism occurs in addition to direct propagation of elastic forces through elastic contacts between the neighbors in the network (Fig. 2). We note that the standard coarse-grained models of protein electrostatics, such as generalized Born models,[48] do not include off-diagonal terms in their solvation free energy. These terms are however sufficiently long-ranged, scaling as r_{ij}^{-3} with the distance between the residues, and they can potentially modify the response of the hydrated protein to either mechanical or electrostatic perturbations.

A. Polar response of hydration water

The dipolar susceptibility $\boldsymbol{\kappa}_{ij}$ in Eq. (6) generally requires either liquid-state models of solvation or electrostatic continuum approaches for its calculation. Here, we start with the former to introduce a sequence of steps to reduce the full complexity of polar response to a clear physical picture and a computationally efficient algorithm.

The susceptibility $\boldsymbol{\kappa}_{ij}$ is given by the convolution of the dipolar tensors $\mathbf{T}_i = \mathbf{T}(\mathbf{r}_{0i} - \mathbf{r})$ and $\mathbf{T}_j = \mathbf{T}(\mathbf{r}_{0j} - \mathbf{r})$, representing the electric fields of the dipoles $\delta\boldsymbol{\mu}_i$ and $\delta\boldsymbol{\mu}_j$ at the point $\mathbf{r}$ in water, with the spacial correlation function of the dipolar fluctuations of water interfacing the protein. The polar response, quadratic in both the electric field and the solvent dipolar polarization, can be conveniently represented by the convolution of reciprocal space $\mathbf{k}$-integrals[27,28,49]

$$\boldsymbol{\kappa}_{ij} = \tilde{\mathbf{T}}_i(\mathbf{k}) * \boldsymbol{\chi}(\mathbf{k}, \mathbf{k}') * \mathbf{T}_j(\mathbf{k}'). \tag{7}$$

Here, the asterisks between the tensors refer to tensor contraction over common indices and integration over common $\mathbf{k}$-variables. Further, the response function $\boldsymbol{\chi}(\mathbf{k}, \mathbf{k}')$ depends on two wave-vectors to reflect the inhomogeneous nature of the problem caused by the presence of the protein in solution. Finally, $\tilde{\mathbf{T}}_i(\mathbf{k})$ is the Fourier transform of the residue-water dipolar tensor taken over the volume Ω occupied by water

$$\tilde{\mathbf{T}}_i(\mathbf{k}) = \int_\Omega \mathbf{T}(\mathbf{r} - \mathbf{r}_{0i}) e^{i\mathbf{k}\cdot\mathbf{r}} d\mathbf{r}. \tag{8}$$

As we have shown elsewhere,[27] the nonlocal part of $\boldsymbol{\chi}(\mathbf{k}, \mathbf{k}')$ is mostly due to transverse polarization fluctuations, given by the component of the dipolar polarization perpendicular to the unit vector $\hat{\mathbf{k}} = \mathbf{k}/k$.[27,50,51] This transverse response is in fact fairly small for most solvation problems[52] (e.g., the Born solvation energy is entirely longitudinal) and will be neglected here. This approximation eliminates the dependence on the second wave-vector with the result[27]

$$\boldsymbol{\chi}(\mathbf{k}, \mathbf{k}') = \hat{\mathbf{k}}\hat{\mathbf{k}} \chi_s^L(k) (2\pi)^3 \delta(\mathbf{k} - \mathbf{k}'). \tag{9}$$

Here, $\chi_s^L(k)$ is the longitudinal dipolar susceptibility of the homogeneous liquid depending on the scalar magnitude k only. It is typically given as a product of the density of dipoles in the liquid y and the longitudinal structure factor $S^L(k)$:[51,53] $\chi_s^L(k) = (3y/4\pi) S^L(k)$. The dipolar density parameter $y = (4\pi/9)\beta\rho m^2$ is defined by the liquid number density ρ and molecular dipole m; $\beta = 1/(k_B T)$ is the inverse temperature.

With the form of the response function given by Eq. (9), the convolution in Eq. (7) is reduced to a 3D integral. While this problem is numerically tractable,[28,49] the number of integrals to be evaluated is $\sim N_i^2/2$, where N_i is the number of ionized residues. This is still a numerically intense computation, and simplifications are desired.

We will further simplify the problem by modeling the calculation of the dipolar tensor of a given residue in Eq. (8). The full calculation of the Fourier transform requires numerical integration over the volume outside a typically complex shape of the protein.[28] To avoid this computationally extensive step, the concept of the relative accessible surface area[54] will be employed here. Specifically, the volume integral in Eq. (8) will be replaced with the integral outside the sphere of radius s, representing the average distance of the closest approach of the water molecules to the residue, and scaled with the fraction of the surface area α_i exposed to the solvent

$$\tilde{\mathbf{T}}_i(\mathbf{k}) = -4\pi \mathbf{D} \alpha_i \frac{j_1(ks)}{ks} e^{i\mathbf{k}\cdot\mathbf{r}_{0i}}. \tag{10}$$

In this equation, $\mathbf{D} = 3\hat{\mathbf{k}}\hat{\mathbf{k}} - \mathbf{1}$, $j_n(x)$ is the spherical Bessel function, and α_i is the ratio of the solvent exposed area a_i to the overall surface area of the residue

$$\alpha_i = a_i/(4\pi s^2). \tag{11}$$

The reduction of Eq. (10) yields an analytical solution for the solvent response function

$$\kappa_{ij}^{\alpha\beta} = \frac{4y\alpha_i\alpha_j}{3}\left[\frac{a(s,r)}{s^3}\delta_{\alpha\beta}\delta_{ij} - (1-\delta_{ij}) b(s,r) T_{ij}^{\alpha\beta}\right]. \tag{12}$$

Here, indices are dropped for brevity in $r = r_{ij}$ and $T_{ij}^{\alpha\beta}$ is the direct-space dipolar tensor connecting beads i and j. The first summand in the brackets in Eq. (12) represents the solvation free energy of the dipole at a charged bead and the second term represents the dipole-dipole interaction between the two charged beads.

The dipole moment $\delta\boldsymbol{\mu}_i = q_i \delta\mathbf{r}_i$ is produced by translation of a charged bead with the overall charge q_i by the distance $\delta\mathbf{r}_i$. Because κ_{ij} represent polar response of water to these residue displacements, both diagonal and off-diagonal terms in Eq. (12) are of dipolar character.

The coefficients $a(s, r)$ and $b(s, r)$ in Eq. (12) are obtained as one-dimensional integrals including the longitudinal structure factor of the liquid $S^L(k)$ to account for non-local correlations between the solvent dipoles. These integrals are listed and calculated in the Appendix. We show there that the dependence on s and r can be lifted in these functions and they in fact are well represented by constants, $a(s, r) = S^L(0)A$, $b(s, r) = S^L(0)$. Here, $S^L(0) = (3y)^{-1}(1 - \epsilon_s^{-1})$ represents the longitudinal dielectric response of a homogeneous polar liquid with the dielectric constant ϵ_s. We finally get for the solvation tensor

$$\kappa_{ij}^{\alpha\beta} = \frac{4\alpha_i\alpha_j}{9}\left(1 - \frac{1}{\epsilon_s}\right)\left[\frac{A}{s^3}\delta_{\alpha\beta}\delta_{ij} - (1-\delta_{ij}) T_{ij}^{\alpha\beta}\right]. \tag{13}$$

The constant $A = 3.54$ scaling the solvation term in Eq. (13) is calculated in the Appendix assuming $s = 4.4$ Å

for the closest-approach distance between the center of a surface amino acid and the oxygen of water. The result is not strongly affected by the choice of s. Note, however, that A absorbs the thermodynamic state of the solvent into it and will depend on thermodynamic variables (temperature, pressure, etc.) through the corresponding polarization structure factor. All calculations and molecular dynamics (MD) simulations presented here refer to the temperature of 300 K and ambient pressure.

B. Dissipative elastic network

The free energy of the hydrated protein in Eq. (5) is a quadratic form in residues' displacements $\delta\mathbf{r}_i$. It can be used to calculate the response to external perturbations or equilibrium variances once the network Hessian $H_{ij}^{\alpha\beta}$ has been specified. We will use here the Hookean springs Hamiltonian suggested by Tirion.[13] This potential, $E(\mathbf{R}) = (C/2)\sum_{ij} D_{ij}(r_{ij} - r_{0,ij})^2$, 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 ($r_{0,ij} = |\mathbf{r}_{0i} - \mathbf{r}_{0j}|$). In this potential, D_{ij} is the connectivity matrix. Its value is set to unity when r_{ij} is within the cutoff distance r_c and is set to zero otherwise. In addition, $D_{ij} = \varepsilon > 1$ for covalently bound neighbors. This scaling accounts for a stronger bonding of residues along the backbone, and is known to better model the vibrational density of states of the protein.[55] Finally, the renormalization of the network by solvation of and water-mediated interactions between ionized residues will follow Eqs. (5) and (6).

The equations of motion for the network beads need to be specified in order to calculate the time-dependent response functions.[56] The elastic network obviously lacks dissipative dynamics typical for soft condense phases. Alternatives to purely mechanical equations of motions can be sought in terms of Langevin dynamics of individual beads.[22,57] Introducing dissipation at the level of individual beads is not necessarily an obvious choice.[9,58] Attempts to derive Langevin equations of motions for individual beads inevitably result in coupled equations of motion.[59] Normal modes of vibrations are decoupled in the static limit, and that decoupling propagates dynamically in the first-order perturbation.[60] Based on these considerations, we have previously opted to introduce dissipation to normal modes $\mathbf{q}_m$ diagonalizing the network Hessian.[61] The equation of motion for such overdamped dynamics is[56]

$$\int_0^t \zeta(t-t')\dot{\mathbf{q}}_m(t')dt' + \lambda_m\mathbf{q}_m = \mathbf{F}(t) + \mathbf{R}(t), \tag{14}$$

where $\zeta(t - t')$ is a memory function, $\mathbf{F}(t) = \mathbf{F}_\omega e^{i\omega t}$ is an external oscillating force, and $\mathbf{R}(t)$ is a randomly fluctuating force. The latter satisfies the generalized fluctuation-dissipation relations[62,63]

$$\langle\mathbf{R}(t)\rangle = 0, \quad \langle\mathbf{R}(t)\cdot\mathbf{R}(0)\rangle = k_B T\zeta(t). \tag{15}$$

The eigenvalues λ_m of normal modes q_m in Eq. (14) are obtained by diagonalizing the Hessian in Eq. (6) with unitary matrix $\mathbf{U}$ according to the standard rules of linear algebra. The eigenvalues are therefore affected by solvation-induced softening of the interface. We indeed observe a shift of the vibrational density of states to softer modes when the network is solvated.

Applying Laplace-Fourier transform[56] to Eq. (14) results in the displacement response function for the collective mode $\mathbf{q}_m$. It is given as a scalar function connecting the average displacement to the external force,[62] $\langle\mathbf{q}_m(\omega)\rangle = \chi_m(\omega)\mathbf{F}_\omega$, where $\langle\mathbf{q}_m(t)\rangle = \langle\mathbf{q}_m(\omega)\rangle e^{i\omega t}$. From Eq. (14), one gets

$$\chi_m(\omega) = [i\omega\tilde{\zeta}(\omega) + \lambda_m]^{-1}, \tag{16}$$

where $\tilde{\zeta}(\omega)$ is the Laplace-Fourier transform of the friction kernel $\zeta(t)$. The entire set of $3N$ eigenvalues λ_m is produced by diagonalizing the Hessian with the unitary matrix $\mathbf{U}$. The inclusion of all eigenvalues of the Hessian results in the response function of the bead displacements

$$\chi_{ij}^{\alpha\beta}(\omega) = C^{-1}\sum_m U_{mi}^{\gamma\alpha}\chi_m(\omega)U_{mj}^{\gamma\beta}. \tag{17}$$

The response function in Eq. (17) can be used to determine the dynamics of correlated motions within the protein scaffold.[64] These dynamics can potentially be complex, nonexponential, and cover many relaxation times.[65] It is currently not clear how to construct memory kernels of the network dynamics to satisfy the experimental constraints. We therefore have adopted a simplified approximation, which allows us to fit correlation functions from MD simulations 100–150 ns in length.[61]

The distance- and self-correlation functions of protein residues calculated from MD trajectories 100–150 ns long typically show two characteristic relaxation times of overdamped motion and, correspondingly, two Debye peaks in their loss spectra. We have previously found[61] that this phenomenology can be reproduced by assigning a two-Debye form to $\chi_m(\omega)$, with two characteristic friction coefficients, ζ_l and ζ_h. We therefore use the following friction kernel in our calculations

$$\chi_m(\omega) = \frac{a}{i\omega\zeta_h + \lambda_m} + \frac{1-a}{i\omega\zeta_l + \lambda_m}, \tag{18}$$

where the amplitude a specifies the relative weight of each relaxation component. This is a phenomenological function, open to future improvements.

The physical reality of two characteristic relaxation times in the dynamics of atomic displacements was modeled differently by Moritsugu and Smith.[58] They separated each principal-component mode calculated from MD trajectories into fast and slow components. The former is described by a damped oscillator with a simple Markovian friction, and the latter, linearly coupled to the fast one, by diffusion. The relative weight of the fast component (a in Eq. (18)) was shown to decay with increasing temperature, with a characteristic kink at the dynamical transition of the protein.[58]

C. Electrostatic response functions

The network response function $\boldsymbol{\chi}_{ij}(\omega)$ in Eq. (17) describes the displacement of residue i induced by a weak oscillating force applied to residue j. Since the residue displacement uniformly moves all of its atomic charges, this linear susceptibility can be used to build electrostatic response

functions of either electrostatic potential or electric field at a given location within the protein.[61]

Assume that an oscillatory probe charge $q_0(t) = q_\omega e^{i\omega t}$ is placed at some location $\mathbf{r}_0$ within the protein. This charge will act on residue i with the force $-q_0(t)\mathbf{E}_{0i}$, where $\mathbf{E}_{0i}$ is the electric field produced at $\mathbf{r}_0$ by all charges of residue i in their equilibrium positions

$$\mathbf{E}_{0i} = \sum_k \frac{q_{ik}(\mathbf{r}_0 - \mathbf{r}_{ik})}{|\mathbf{r}_0 - \mathbf{r}_{ik}|^3}. \tag{19}$$

Here, q_{ik} are the atomic charges of residue i with the equilibrium coordinates $\mathbf{r}_{ik}$.

The force acting on residue i will propagate through the elastic network to residue j according to the response function $\boldsymbol{\chi}_{ij}(\omega)$. The displacement of that residue will in turn alter the electrostatic potential produced by the protein at $\mathbf{r}_0$. After summing over all residues in the network, one arrives at the frequency-dependent susceptibility of the electrostatic potential

$$\chi_\phi(\omega) = -\sum_{i,j} E^\alpha_{0j}\chi^{\alpha\beta}_{ij}(\omega)E^\beta_{0i}. \tag{20}$$

This susceptibility determines the alteration of the electrostatic potential produced by the charges of the protein matrix $\delta\phi_0(\omega)$ at the position $\mathbf{r}_0$ of the probe charge q_ω: $\delta\phi_0(\omega) = \chi_\phi(\omega)q_\omega$.

Similarly, one can define the electric field alteration $\delta\mathbf{E}_0$ produced by the protein matrix at $\mathbf{r}_0$ in response to placing an oscillating probe dipole $\boldsymbol{\mu}_\omega$ at that point. This frequency-dependent susceptibility is based on convoluting $\boldsymbol{\chi}_{ij}(\omega)$ with the dipolar tensors $\mathbf{T}_{ik} = \mathbf{T}(\mathbf{r}_0 - \mathbf{r}_{ik})$ connecting the residue charge q_{ik}, located at $\mathbf{r}_{ik}$, to the position of the probe dipole at $\mathbf{r}_0$. The result is

$$\chi_E^{\alpha\beta}(\omega) = \sum_{i,j,k,l} q_{ik}T^{\alpha\gamma}_{ik}\chi^{\gamma\delta}_{ij}(\omega)T^{\delta\beta}_{jl}q_{jl}. \tag{21}$$

Here, as above, summation runs over the repeated Greek indices denoting Cartesian projections of the corresponding tensors. The difference in signs in Eqs. (20) and (21) comes from the fact that the free energy invested into the creation of the potential alteration is $(1/2)q_\omega\delta\phi_0(\omega)$, while for the dipole one has $-(1/2)\boldsymbol{\mu}_\omega \cdot \delta\boldsymbol{E}_0(\omega)$.

D. Dielectric response

When a uniform oscillatory external field $\mathbf{E}_0(t) = \mathbf{E}_\omega e^{i\omega t}$ is applied to a protein, it induces the dipole moment $\delta\mathbf{M}(\omega) = \sum_j \delta\boldsymbol{\mu}_j(\omega)$. Since only movements of the charged residues produce non-zero dipoles, $\delta\mathbf{M}(\omega) = \sum_j q_j\delta\mathbf{r}_j(\omega)$. Substituting the network displacements susceptibility, one arrives at the relation

$$\delta M^\alpha(\omega) = \sum_{i,j} q_i q_j \chi^{\alpha\beta}_{ij}(\omega)E^\beta_\omega. \tag{22}$$

Assuming that the external field is along the z-axis of the laboratory frame, one gets for the dipolar susceptibility

$$\chi_M(\omega) = \sum_{ij} q_i q_j \chi^{zz}_{ij}(\omega) = (1/3)\sum_{ij} q_i q_j \chi^{\alpha\alpha}_{ij}(\omega). \tag{23}$$

This dipolar susceptibility refers to the dipole moment induced at a single protein molecule. It can be used to calculate the complex-valued dielectric constant $\epsilon_p(\omega)$ of a protein sample (either powder or polycrystal) by applying the standard derivation of the theory of dielectrics.[66] The result is

$$\frac{(\epsilon_p(\omega) - 1)(2\epsilon(\omega) + 1)}{9\epsilon_p(\omega)} = \frac{4\pi}{3}\rho_p\chi_M(\omega) + y_e, \tag{24}$$

where ρ_p is the number density of the protein molecules in the material. The density of induced dipoles y_e in this equation describes electronic dipolar polarizability of the protein. It can be estimated using the Clausius-Mossotti equation if the refractive index of the protein powder is known.[66]

E. Allosteric response

As an example of the application of the formalism of response functions to the allosteric response of a protein, we will consider the displacement $\delta r_i^\alpha(\omega)$ of residue i in response to binding a charge $q(t) = q_\omega e^{i\omega t}$ at position $\mathbf{r}_0$. The response is therefore effectively of piezoelectric type,[67] creating deformation at a distant site in response to electric stimulus.

Binding of an ion causes both global and local perturbations of the protein. From the global perspective, it changes the free energy of the entire protein by the free energy of binding ΔF and, in addition, exerts Coulomb forces on all charges of the protein. The global perspective can be studied, as is typically done in electrostatics,[67–70] by asking what is the free energy cost $F_i(q)$ of transferring a small probe charge q to the binding site, where $i = 1, 2$ labels the two stable conformations of the protein, ion-free and ion-bound. The electrostatic potential susceptibility [Eq. (20)] addresses this question. The free energy cost is obviously

$$F_i(q) = \frac{1}{2}\chi^{(i)}_\phi(0)(q - q_{i0})^2 + F_{i0}, \tag{25}$$

where $q_{10} = 0$, $F_{10} = 0$, and $q_{20} = q_i$ is the charge of the binding ion and $F_{20} = \Delta F$ is the binding free energy. We have also added the dependence of the response function on the protein state since susceptibility $\chi^{(i)}_\phi(0)$ can be sensitive to structural changes of the protein.

The amount of transferred charge q can be viewed as the reaction coordinate for the global free energy cost of binding the partial charge q. The continuous charge variable represents the fraction of the ensemble that has changed its binding state, without flipping in a new conformation, as a result of a spontaneous thermal fluctuation requiring the input of free energy $F_i(q) - F_i(q_{0i})$. The crossing of the free energy surfaces, $F_1(q^\dagger) = F_2(q^\dagger)$, will define the transition state and the corresponding free energy barrier. This picture allows for a "Marcus inverted region behavior,"[71] i.e., there is an optimal binding free energy minimizing the free energy barrier. The activation barrier starts to grow when ΔF falls below the optimal value (inverted region). In addition, because the curvatures of the two parabolas may differ, there is a scenario in which no crossing in the inverted region occurs, i.e., the activation barrier becomes infinite and the reaction is not allowed. Note, however, that the curvatures of two surfaces calculated for the NtrC protein studied below are nearly identical

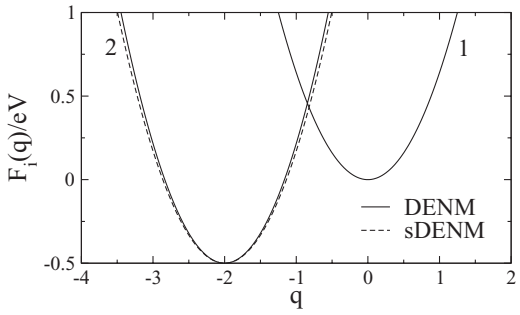

FIG. 3. Free energy surfaces representing the free energy penalty (reversible work) of changing the charge at a binding site of an allosteric protein. The calculations are done for attaching carbamoylphosphate ($q_{02} = -2$) to the bacterial enhancer-binding protein NtrC (Fig. 8). The curves refer to DENM in the unphosphorylated state ($q_{01} = 0$) and to DENM and sDENM in the phosphorylated state ($q_{02} = -2$). The elastic network is defined with $k_BT/C = 0.75$ Å^2 and $\epsilon = 125$; the cutoff radius is 15 Å. The free energy of binding ΔF is unknown and was set at -0.5 eV for the purpose of illustration.

(Fig. 3). There is also little sensitivity of the overall free energy functions to solvation of the surface residues (compare DENM and sDENM calculations). This lack of global sensitivity is in stark contrast with a strong effect of solvation on individual residue displacements, as we show below.

Most of the interest in the field is driven not by the global thermodynamics of binding, but by the need to understand biological function caused by it.[4,36] A typical problem is to calculate the displacement of a distant residue in response to binding. We will approach this question, as above, by considering an oscillatory charge placed at $\mathbf{r}_0$. This charge will interact with each residue j through the electric field $-\mathbf{E}_{0j}$ given by Eq. (19). This interaction induces the force acting on each bead in the network, $\delta F_j^\alpha(\omega) = -E_{0j}^\alpha q_\omega$.

In the linear response approximation,[56] the average displacement of residue i is given by summing up the forces produced by the charge q_ω at all residues of the network with their response function $\chi_{ij}^{\alpha\beta}(\omega)$ propagating the force at j into a displacement at i

$$\langle \delta r_i^\alpha(\omega)\rangle = -\sum_j \chi_{ij}^{\alpha\beta}(\omega) E_{0j}^\beta q_\omega. \tag{26}$$

In the calculations below we will present the scalar displacement

$$\delta r_i(\omega) = \left[\langle \delta r_i'^\alpha(\omega)\rangle \langle \delta r_i'^\alpha(\omega)\rangle\right]^{1/2}, \tag{27}$$

where $\delta r_i'^\alpha(\omega)$ is the real part of the complex-valued displacement and summation over repeated Greek indices is performed.

The frequency ω of the oscillatory charge might effectively represent the time scale of charge binding, such as the frequency of binding/unbinding events, typically occurring on the nanosecond time scale for small electrolyte ions.[6] The limit $\omega = 0$ in this formalism will represent stationary, i.e., adiabatically slow, binding.

III. RESULTS

Before presenting the results of specific calculations, we start with some crude estimates of the effect of solvation of charged residues on the properties of the elastic network. Assuming the elastic spring constant of $k_BT/C \simeq 1$ Å^2, consistent with other estimates in the literature,[13,14] one can estimate the effect of solvation on the network Hessian. We consider the diagonal element in Eq. (6),

$$\tilde{H}_{ii}^{\alpha\alpha} = 2 - 4A\alpha_i^2 q^2/(3Cs^3)\left(1 - \epsilon_s^{-1}\right). \tag{28}$$

With $q = e$, $s = 4.4$ Å, $\epsilon_s = 78$, and $A = 3.54$, the second term becomes $30\alpha_i^2$. This estimate suggests that any singly-charged residue exposed to water to more than a quarter of its surface will have a negative elastic constant with its non-covalent neighbors and will be held in the equilibrium position only by covalent bonds within the network. Such ionized residue would lose mechanical stability and dissolve in water if not held in place by its covalent neighbors. It is clear that solvation makes a major effect on the elastic response of charged residues.

Sections III A–III C will present illustrative calculations using the sDENM model. These calculations are not designed to fully describe the systems considered, but instead offer case studies to demonstrate the model capabilities and physical problems the model can potentially address. We have therefore chosen to keep the set of network parameters constant in all these calculations. We fix the cutoff radius at $r_c = 15$ Å and the enhancement of the binding constant between the covalent neighbors at $\epsilon = 125$. Changing either of these parameters does not affect the qualitative outcome of the calculations.[61] The friction coefficients in the two-Debye friction kernel in Eq. (18) are set to reproduce the spectrum of the electrostatic potential fluctuations at the heme of cytochrome B562 shown below: $\zeta_l = 30$ ns, $\zeta_h = 0.006\zeta_l$, and $a = 0.35$ in Eq. (18). Finally, the force constant of springs connecting non-covalent neighbors in the network is set to $k_BT/C = 0.75$ Å^2, which is chosen to fit the variance of the electrostatic potential at the protein heme from MD simulations.[61] These force constants are softened by hydration of surface ionized residues according to Eq. (28).

A. Residue displacements and electrostatic response

The standard approach of experimental verification and parameterization of elastic protein networks is to compare the root-mean-square displacements (rmsd's) of residues with experiment or MD simulations. Crystallographic B-factors are often used,[72] but those are of limited value.[23] It was noted that reported B-factors are dominated by rigid-body motions of the proteins in the crystal.[73] In addition, there is a clear mismatch between the reported rmsd's of proteins in crystals and in their flexibility in solution, as is illustrated in Fig. 4 comparing rmsd's from B-factors of C^α's with their rmsd's found from MD. The MD simulations were done for hydrated cytochrome B562 (cytB, PDB entry 256B, Fig. 5) as described elsewhere.[61,74]

The mismatch between both the B-factors and the standard ENM as compared to MD is particularly notable for the flexible loop (residues 46 to 55) containing several ionized residues (Figs. 4 and 5). This region is clearly not restricted to a single configuration in solution and instead wanders through

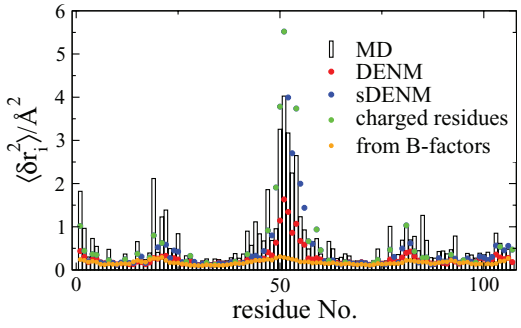

FIG. 4. Mean-square displacements of cytochrome B562 (cytB) C^{α}'s from MD simulations, DENM and sDENM calculations. The network parameters in the DENM/sDENM calculations are: $k_BT/C = 0.75$ Å^2, $\epsilon = 125$, and the cutoff radius is 15 Å. The solvent accessible surface for the loop residues labeled in Fig. 5 is scaled down to 45 Å^2.

a number of semi-stable conformations. The network, required to reside in a single conformation, is expected to lose stability because of this and similar segments. The standard ENM clearly avoids this instability by over-restricting the flexible residues. In contrast, when renormalization by solvation is introduced in sDENM, the network loses stability, as expected, due to solvation of the loop residues labeled in Fig. 5. Since calculations cannot be performed with an unstable network, we have artificially restricted the network by scaling down the solvent-accessible area of the residues labeled in Fig. 5 from the values calculated with visual molecular dynamics (VMD)[75] (in the range of 120–150 Å^2) to 45 Å^2. This rescaling prevents mechanical instability of the network, but preserves the physical reality of a flexible loop, as is seen from the corresponding rmsd's in Fig. 4.

Figure 6 shows the results of the calculations (cytB) for the electrostatic potential susceptibility $\chi_\phi(\omega)$ and the tensor contraction $\chi_E(\omega) = \chi_E^{\alpha\alpha}(\omega)$ for the electric field susceptibility. The effect of solvating surface residues is less pronounced for these susceptibilities, in particular for the more long-ranged electrostatic potential. A slow relaxation component, not resolved on the length of the MD trajectory, appears for the electric field susceptibility. This slow component arises from much slower motions of highly solvated residues in the network, also seen in the dielectric response of the protein.

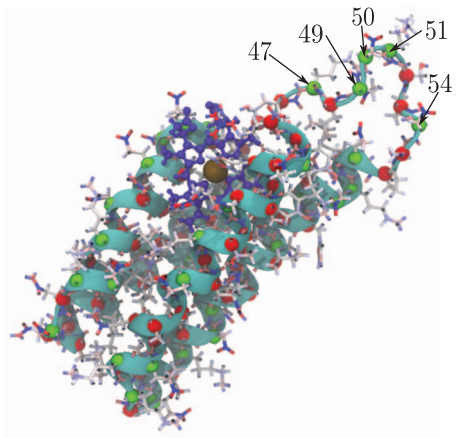

FIG. 5. Cartoon of cytochrome B562 (cytB) showing the positions of C^{α} (spheres) colored by residue charge: charged (green) and uncharged (red). The charged residues marked green (also green points in Fig. 4) are also required to have α_i greater than 0.16, used as a threshold number. The remaining C^{α}'s are marked red. The side chain atoms are colored by charge: negative (red), positive (blue), and neutral (white). The heme iron is colored brown while the remaining atoms of the heme are blue. Numbers label the unstable residues in the loop for which the water-exposed surface was scaled down to 45 Å^2 in order to maintain the network stability.

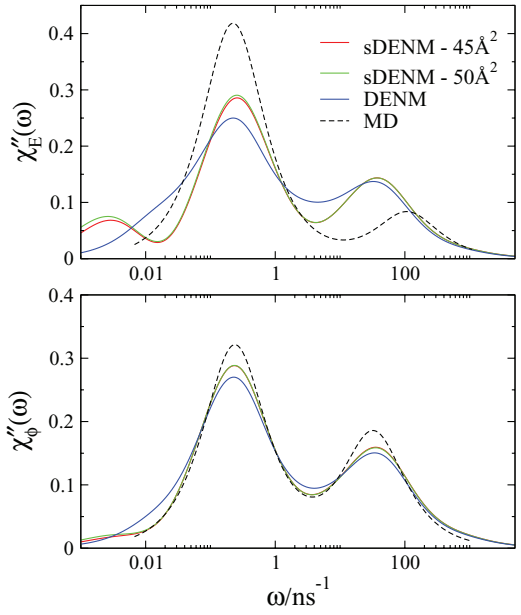

FIG. 6. Loss spectra $\chi_E''(\omega)/\chi_E'(0)$ and $\chi_\phi''(\omega)/\chi_\phi'(0)$ for cytB. Compared are MD, DENM, and sDENM calculations. To show the sensitivity of sDENM calculations to solvation of the loop residues in cytB, the results of choosing the solvent-accessible area of $a_i = 45$ Å^2 and of $a_i = 50$ Å^2 are shown. The two-Debye relaxation parameters are $\zeta_l = 30$ ns, $\zeta_h = 0.006\zeta_l$, and $a = 0.35$ [Eq. (18)]. The elastic network is defined with $k_BT/C = 0.75$ Å^2, $\epsilon = 125$, and the cutoff radius of 15 Å.

B. Dielectric susceptibility of the protein

The calculated imaginary part of the dielectric susceptibility (loss function) $\chi_M''(\omega)$ from Eq. (23) is shown in Fig. 7. Similarly to the case of $\chi_E''(\omega)$, it clearly shows the emergence of a slow peak, about two orders of magnitude slower than the main peak. The slow component comes from the hydrated residues with high exposure to water. This is

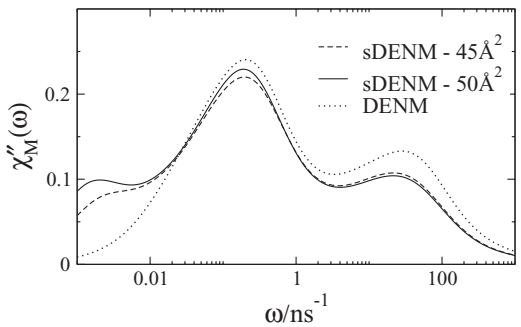

FIG. 7. $\chi_M''(\omega)/\chi_M'(0)$ for cytB. The parameters of the network are the same as in Fig. 6.

clearly seen from the sensitivity of the slow peak to the solvent-accessible surface assigned to the residues of the loop. We need to note that the network completely neglects librations of polar residues, focusing only on the polarization fluctuations produced by translational motions of the charged residues. An additional dielectric intensity might therefore come from the components missing from the model.[76]

Experimentally, partially hydrated protein powders show three relaxation processes at low temperatures, which merge into two processes at ambient temperature.[12] The fastest and the slowest processes disappear when the protein is dried. The relaxation time of the main peak from dielectric measurements matches well the relaxation time from neutron scattering experiments, in which the protein and hydration water signals can be separated by deuteration. The main peak is therefore assigned to low-frequency, dissipative protein motions.[12] In this regard, the main peak in Fig. 7 can be tentatively aligned with the main peak of dielectric measurements.

The slowest observed peak,[12] about two orders of magnitude slower than the main peak, has been hard to interpret by experimental means. Its strong dependence on the level of hydration, however, suggests that it should be linked to the protein. Indeed, our calculations give clear evidence that slow relaxation is related to overdamped motions of highly solvated ionized residues. The two orders of magnitude ratio of the slow and main relaxation times is in qualitative agreement with the dielectric measurements. Finally, the peak disappears when hydration of ionized residues is removed, which is analogous to drying the sample in experiment. The assignment of the slow peak should also emphasize the involvement of water in the relaxation process. Since water is fast and follows adiabatically the protein motions, ionized residues move by dragging hydration waters with them.

C. Allosteric response

The calculations of the allosteric response to ion binding have been done for a single-domain signaling protein NtrC. The structures of this protein have been resolved[77] both in unphosphorylated (denoted as NtrC) and in phosphorylated (denoted as P-NtrC) states. The latter state is short-lived, and it was maintained in solution at a large excess of the phosphordonor carbamoylphosphate carrying the charge of $q_{02} = -2$. The addition of this charge to Asp54 active site (Fig. 8) creates a Coulomb force acting on the neighboring atomic charges, such that each residue j experiences the force $-q_{02}\mathbf{E}_{0j}$. The perturbing force produced by ion binding is therefore fairly nonlocal, in contrast to a common assumption,[44] and the calculation of the response requires full account of this fact.

The frequency-dependent displacement of residue i in Eqs. (26) and (27) sums up all Coulomb forces acting on residues j from the active site labeled as "0." The unphosphorylated (NtrC) state of the protein is very mobile, with several loops continuously changing their conformation on the μs to ms time scale. These motions mostly disappear in a more compact phosphorylated state.[37] Not surprisingly, we have found that sDENM is rather unstable and only DENM calculations could be done on the NtrC state. Therefore, sDENM calculations were done only on the P-NtrC state. The results for $\delta r_{124}(\omega)$ ($i = 124$ is glutamine in NtrC) in both states are shown in Fig. 9. As expected, the inclusion of solvation in sDENM greatly enhances the displacement magnitude. The frequency dependence is also noteworthy. It implies the existence of μs motions of the protein responding to the charge perturbation.[37] If the frequency of binding/unbinding events exceeds this frequency, the protein does not have the ability to respond to the perturbation and this subset of motions dynamically freezes. As a result, the displacement diminishes.

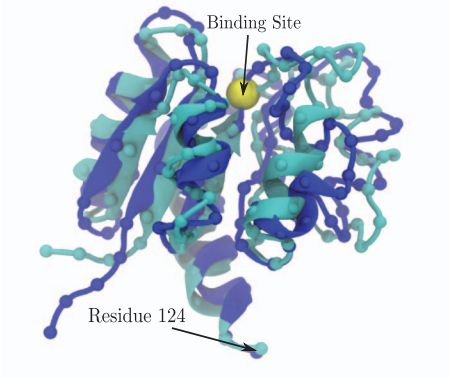

FIG. 8. Superimposed structures of the bacterial enhancer-binding protein NtrC in dephosphorylated (light blue, PDB entry 1DC7) and phosphorylated (dark blue, PDB entry 1DC8) states.[37,77] The nuclear magnetic resonance structure was determined[77] with carbamoylphosphate binding to Asp54. The displacement of Glu124 in response to a probe charge at Asp54 is shown in Fig. 9.

Figure 10 emphasizes a strong effect of solvation on ion binding allostery presented in Fig. 9 by showing zero-frequency displacements of all residues in the NtrC protein. The calculations have been done in DENM for NtrC and P-NtrC and in sDENM for P-NtrC. The results are compared with Δr_i displacements of C^α between the two structures (Fig. 8). It is clear that only by including solvation within sDENM does the displacements of residues in a given state reach the magnitudes comparable with the overall displacement amplitudes Δr_i. It appears that, while solvation does not strongly affect the global energetics of ion binding (Fig. 3),

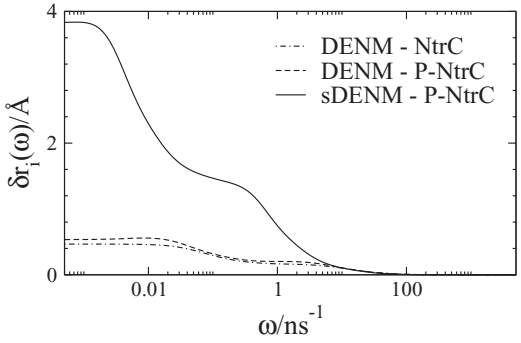

FIG. 9. Frequency-dependent allosteric displacement [Eq. (27)] for residue $i = 124$ (Glu) of NtrC in the unphosphorylated (NtrC) and phosphorylated (P-NtrC) states. The results of calculations within DENM and sDENM are compared as shown in the plot.

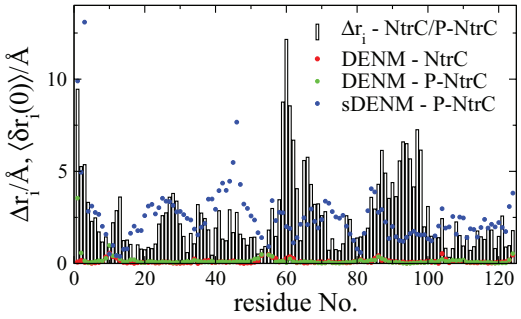

FIG. 10. Displacements Δr_i between two equilibrium structures (NtrC and P-NtrC) of the NtrC protein. $\delta r_i(0)$ shows the displacement of residue i in response to placing a unitary probe charge $q_\omega = 1$ at the position of C^α of the binding site (Asp54, Fig. 8).

it critically affects the allosteric amplification of ion binding through conformational transitions of individual residues.

IV. SUMMARY

Folded proteins have to maintain structural stability. At the same time many functions of enzymes and motor proteins involve large-scale domain movements in response to binding and release of ligands. This requirement makes one suggest that some stability needs to be sacrificed to allow amplification of a small perturbation into a large response. The question is what are the structural motifs that allow amplification without compromising the global stability. A related issue is the length of correlations, since long-ranged correlations are required for allosteric action.

The first obvious target to address the problem is elasticity. The protein is densely packed and any force perturbation propagates through its body as in a glass material. The elastic deformation spreads, however, through the elastic body and does not accommodate for a directed, specific action. Elasticity can capture motions of relatively rigid domains linked by flexible hinges,[44] but to a lesser extent the allostery of monomeric, single-domain systems.

One can alternatively turn attention to hydration water[78] as a possible medium for transferring signals. Water can store significant energy in the form of dipolar polarization and large entropy in its hydrogen-bond network, but it is also a highly non-specific medium. The required specificity might therefore reside at the protein-water interface, combining large energies stored in hydration with regulation achieved through identities of the surface residues.

The surface charges, and to some extent dipoles, carry large solvation free energies and are strongly correlated in their motions, with water-mediated correlations decaying as r^{-3}. Because of non-locality of correlations, an ensemble of ionized surface residues forms a strongly correlated net (or, in a sense, merged multiple pathways[36]) enveloping the entire protein. Given that multiple binding sites are typically involved in protein function, allostery might be designed not by building a specific site and attaching strings ("communication pathways"[6,79]) to it, but by pulling on the net wherever one finds a "knot." Some knots might be more important than the others from the perspective of biological function. The non-locality of this net excludes the possibility of well-defined pathways, they must be achieved by more specific interactions involving either nonpolar residues[80] or chains of hydrogen bonds.[79]

A network of ionized, hydrated surface residues is a general property of all hydrated proteins, affecting a number of observable properties. The formalism of solvated dissipative electro-elastic network captures this reality and projects it on a number of susceptibilities describing the response to a particular type of external perturbation of a given experiment. A number of observables, such as rmsd's, dielectric response, electrostatic susceptibilities, and allosteric response are affected by surface solvation. The general outcome is that elastic motions of the residues become significantly heterogeneous, with softening achieved at the sites carrying charges. The interfacial heterogeneity does not dramatically affect the global thermodynamics of the protein or the thermodynamics of ion binding, but is critical for local responses to external perturbations. While large-scale motions of the protein altering its shape occur on the nanosecond time scale, μs motions[37,81] are assigned to portions of the protein with highly hydrated ionized residues.

ACKNOWLEDGMENTS

This research was supported by the National Science Foundation (MCB-1157788). CPU time was provided by the National Science Foundation through TeraGrid resources (TG-MCB080116N).

APPENDIX: SOLVATION INTEGRALS

The solvation integrals in Eq. (12) are given by the following one-dimensional k-space integrals involving the longitudinal structure factor $S^L(k)$ of the homogeneous solvent:

$$a(s,r) = \delta_{r,0} + \frac{6s}{\pi}\int_0^\infty j_1(sk)^2 j_0(rk)(S^L(k)-1)dk,$$

$$b(s,r) = \theta(r-2s) + \frac{6s}{\pi}\left(\frac{r}{s}\right)^3 \int_0^\infty j_1(sk)^2 j_2(rk)(S^L(k)-1)dk. \quad \text{(A1)}$$

Here, $j_n(x)$ is a spherical Bessel function, s denotes the effective radius of a residue, and r is the distance between the centers of two beads in the network. Further, $r = 0$ corresponds to one bead and that configuration is represented by the Kronecker delta $\delta_{r,0}$, which is equal to unity when $r = 0$. Since dipolar structure factors satisfy the asymptote $S^L(k) \to 1$ at $k \to \infty$, this limit is separated from the numerical integral and is given by the first summand in each equation.

The dipolar structure factors of polar liquids have been intensively studied in the past.[53] Analytical models from liquid-state theories also exist.[56] Several studies reported structure factors of force-field water models.[82,83] These functions are accessible only from simulations since there is no known experimental technique giving access to them.

For the purpose of estimating the integrals in Eq. (A1) we have taken the longitudinal structure factor of TIP3P

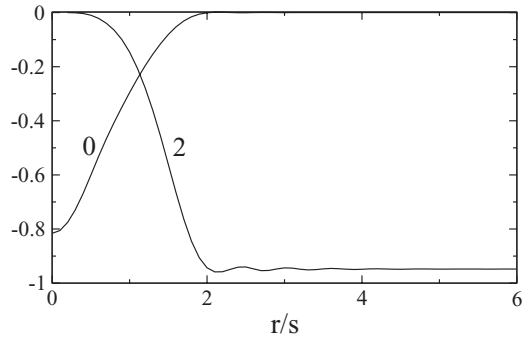

FIG. 11. Integrals entering Eq. (A1) calculated with the longitudinal structure factor of TIP3P water.[28] The labeling in the plot point to the integration with $j_0(rk)$ (first equation in Eq. (A1), labeled as "0") and with $j_2(rk)$ (second equation in Eq. (A1), labeled as "2"). The integrals are calculated as the function of r/s with fixed $s = 4.4$ Å.

water calculated from MD simulations.[28] A simple parameterization of this function is available[28,52] based on the solution of the mean-spherical closure for dipolar hard spheres.[56] The results of this integration are shown in Fig. 11. As is seen, the first integral involving $j_0(rk)$ quickly goes to zero when reaching the distance $r/s \simeq 2$. Since r is either zero, for one-bead solvation, or greater than $2s$, for different beads, only $r = 0$ needs to be considered for this function. We therefore put $a(s, r) = A(s)S^L(0)$, where $A(s = 4.4\ \text{Å}) = 3.54$ is numerically calculated. The value $s = 4.4$ Å is the sum of the average radius of 3 Å assigned to a residue and 1.4 Å for the radius of waters.

The situation is just the opposite for the second integral. It is zero at $r = 0$ and reaches the value $S^L(0) - 1$ at $r/s = 2$. This latter result implies that the continuum limit approximation $S^L(k) = S^L(0)$ applies in this case. We therefore put $b(s, r) = S^L(0)\theta(r - 2s)$.

The overall result of these calculations, incorporating solvent dipolar correlations through the longitudinal structure factor, is quite clear. The solvation energy, at $r = 0$, is renormalized by the factor A from the dielectric continuum limit $S^L(k) = S^L(0)$. This renormalization effectively reduces the cavity radius for dipolar solvation from the distance of the closest approach of water to the residue s to $s/A^{1/3}$. This is consistent with the common observation that the effective cavity radius should fall between s and the van der Waals radius of the solute $s - \sigma_s/2$ (σ_s is the water diameter). On the other hand, the dipolar water-mediated coupling between distant residues is well described by the continuum limit of the solvent dipolar response, and that fact is reflected in the constancy of $b(s, r)$.

[1] J. Monod, J. Wyman, and J.-P. Changeux, J. Mol. Biol. **12**, 88 (1965).
[2] D. Kern and E. R. P. Zuiderweg, Curr. Opin. Struct. Biol. **13**, 748 (2003).
[3] J.-P. Changeux and S. J. Edelstein, Science **308**, 1424 (2005).
[4] Q. Cui and M. Karplus, Protein Sci. **17**, 1295 (2008).
[5] D. E. Koshland, G. Némethy, and D. Filmer, Biochemistry **5**, 365 (1966).
[6] P. I. Zhuravlev and G. A. Papoian, Q. Rev. Biophys. **43**, 295 (2010).
[7] O. Miyashita, J. N. Onuchic, and P. G. Wolynes, Proc. Natl. Acad. Sci. U.S.A. **100**, 12570 (2003).
[8] P. W. Fenimore, H. Frauenfelder, B. H. McMahon, and R. D. Young, Proc. Natl. Acad. Sci. U.S.A. **101**, 14408 (2004).
[9] S. Hayward, A. Kitao, F. Hirata, and N. Go, J. Mol. Biol. **234**, 1207 (1993).
[10] D. Thirumalai, E. P. O'Brien, G. Morrison, and C. Hyeon, Annu. Rev. Biophys. **39**, 159 (2010).
[11] G. A. Papoian, J. Ulander, M. P. Eastwood, Z. Luthey-Schulten, and P. G. Wolynes, Proc. Natl. Acad. Sci. U.S.A. **101**, 3352 (2004).
[12] S. Khodadadi, S. Pawlus, and A. P. Sokolov, J. Phys. Chem. B **112**, 14273 (2008).
[13] M. M. Tirion, Phys. Rev. Lett. **77**, 1905 (1996).
[14] A. R. Atilgan, S. R. Durell, R. L. Jernigan, M. C. Demirel, O. Keskin, and I. Bahar, Biophys. J. **80**, 505 (2001).
[15] F. Tama and Y. H. Sanejouand, Protein Eng. **14**, 1 (2001).
[16] F. Tama and C. L. Brooks, J. Mol. Biol. **345**, 299 (2005).
[17] K. Moritsugu and J. C. Smith, Biophys. J. **93**, 3460 (2007).
[18] D. Riccardi, Q. Cui, and G. N. Phillips, Biophys. J. **96**, 464 (2009).
[19] T. D. Romo and A. Grossfield, Proteins: Struct., Funct., Bioinf. **79**, 23 (2011).
[20] E. Lyman, J. Pfaendtner, and G. A. Voth, Biophys. J. **95**, 4183 (2008).
[21] K. Hinsen and G. R. Kneller, J. Chem. Phys. **111**, 10766 (1999).
[22] B. T. Miller, W. Zheng, R. M. Venable, R. W. Pastor, and B. R. Brooks, J. Phys. Chem. B **112**, 6274 (2008).
[23] B. Halle, Proc. Natl. Acad. Sci. U.S.A. **99**, 1274 (2002).
[24] M. Lu and J. Ma, Biophys. J. **89**, 2395 (2005).
[25] F. Tama and C. L. Brooks, Annu. Rev. Biophys. Biomol. Struct. **35**, 115 (2006).
[26] M. Feig and C. L. Brooks III, Curr. Opin. Struct. Biol. **14**, 217 (2004).
[27] D. V. Matyushov, J. Chem. Phys. **120**, 7532 (2004).
[28] D. N. LeBard and D. V. Matyushov, J. Chem. Phys. **128**, 155106 (2008).
[29] H. B. Gray and J. R. Winkler, Proc. Natl. Acad. Sci. U.S.A. **102**, 3534 (2005).
[30] S. K. Pal and A. H. Zewail, Chem. Rev. **104**, 2099 (2004).
[31] M. Yang and J. L. Skinner, Phys. Chem. Chem. Phys. **12**, 982 (2010).
[32] A. T. Fafarman, P. A. Sigala, J. P. Schwans, T. D. Fenn, D. Herschlag, and S. G. Boxer, Proc. Natl. Acad. Sci. U.S.A. **109**, E299 (2012).
[33] D. V. Matyushov and M. D. Newton, J. Phys. Chem. A **105**, 8516 (2001).
[34] S. Takashima, *Electrical Properties of Biopolymers and Membranes* (Adam Hilger, Bristol, 1989).
[35] S. Khodadadi, J. H. Roh, A. Kisliuk, E. Mamontov, M. Tyagi, S. A. Woodson, R. M. Briber, and A. P. Sokolov, Biophys. J. **98**, 1321 (2010).
[36] A. del Sol, C.-J. Tsai, B. Ma, and R. Nussinov, Structure **17**, 1042 (2009).
[37] B. Volkman, D. Lipson, D. Wemmer, and D. Kern, Science **291**, 2429 (2001).
[38] L. Ma and Q. Cui, J. Am. Chem. Soc. **129**, 10261 (2007).
[39] P. Maragakis and M. Karplus, J. Mol. Biol. **352**, 807 (2005).
[40] W. Zheng, B. R. Brooks, and G. Hummer, Proteins **69**, 43 (2007).
[41] J.-W. Chu and G. A. Voth, Biophys. J. **93**, 3860 (2007).
[42] S. Tripathi and J. J. Portman, J. Chem. Phys. **135**, 075104 (2011).
[43] I. Bahar and A. J. Rader, Curr. Opin. Struct. Biol. **15**, 586 (2005).
[44] M. Ikeguchi, J. Ueno, M. Sato, and A. Kidera, Phys. Rev. Lett. **94**, 078102 (2005).
[45] R. A. Kuharski, J. S. Bader, D. Chandler, M. Sprik, M. L. Klein, and R. W. Impey, J. Chem. Phys. **89**, 3248 (1988).
[46] G. Hummer, L. R. Pratt, and A. E. Garcia, J. Phys. Chem. A **102**, 7885 (1998).
[47] J. Blumberger and M. Sprik, J. Phys. Chem. B **109**, 6793 (2005).
[48] W. C. Still, A. Tempczyk, R. C. Hawley, and T. Hendrickson, J. Am. Chem. Soc. **112**, 6127 (1990).
[49] A. A. Milischuk, D. V. Matyushov, and M. D. Newton, Chem. Phys. **324**, 172 (2006).
[50] P. Madden and D. Kivelson, Adv. Chem. Phys. **56**, 467 (1984).
[51] D. Kivelson and H. Friedman, J. Phys. Chem. **93**, 7026 (1989).
[52] D. V. Matyushov, J. Chem. Phys. **120**, 1375 (2004).
[53] M. S. Skaf and B. M. Ladanyi, J. Chem. Phys. **102**, 6542 (1995).
[54] W. Hasel, T. Hendrickson, and W. Still, Tetrahedron Comput. Methodol. **1**, 103 (1988).
[55] D. Ming and M. E. Wall, Phys. Rev. Lett. **95**, 198103 (2005).
[56] J. P. Hansen and I. R. McDonald, *Theory of Simple Liquids* (Academic, Amsterdam, 2003).
[57] S. G. Essiz and R. D. Coalson, J. Phys. Chem. B **113**, 10859 (2009).
[58] K. Moritsugu and J. C. Smith, J. Phys. Chem. B **110**, 5807 (2006).
[59] R. Soheilifard, D. E. Makarov, and G. J. Rodin, J. Chem. Phys. **135**, 054107 (2011).
[60] G. Lamm and A. Szabo, J. Chem. Phys. **85**, 7334 (1986).
[61] D. R. Martin, S. B. Ozkan, and D. V. Matyushov, Phys. Biol. **9**, 036004 (2012).
[62] R. Kubo, Rep. Prog. Phys. **29**, 255 (1966).
[63] R. Zwanzig, *Nonequlibrium Statistical Mechanics* (Oxford University Press, Oxford, 2001).

[64]S. Weiss, Nat. Struct. Biol. **7**, 724 (2000).
[65]W. Min, B. P. English, G. Luo, B. J. Cherayil, S. C. Kou, and X. S. Xie, Acc. Chem. Res. **38**, 923 (2005).
[66]B. K. P. Scaife, *Principles of Dielectrics* (Clarendon, Oxford, 1998).
[67]L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Pergamon, Oxford, 1984).
[68]M. Born, Z. Phys. **1**, 45 (1920).
[69]R. A. Marcus, J. Chem. Phys. **24**, 966 (1956).
[70]G. Hummer, L. R. Pratt, and A. E. Garcia, J. Phys. Chem. **100**, 1206 (1996).
[71]R. A. Marcus and N. Sutin, Biochim. Biophys. Acta **811**, 265 (1985).
[72]I. Bahar, T. R. Lezon, L.-W. Yang, and E. Eyal, Annu. Rev. Biophys. **39**, 23 (2010).
[73]R. Soheilifard, D. E. Makarov, and G. J. Rodin, Phys. Biol. **5**, 026008 (2008).
[74]D. V. Matyushov, J. Phys. Chem. B **115**, 10715 (2011).
[75]W. Humphrey, A. Dalke, and K. Schulten, J. Mol. Graphics **14**, 33 (1996).
[76]X. Song, J. Chem. Phys. **116**, 9359 (2002).
[77]D. Kern, B. Volkman, P. Luginbühl, M. J. Nohaile, S. Kustu, and D. E. Wemmer, Nature (London) **402**, 894 (1999).
[78]P. Ball, Chem. Rev. **108**, 74 (2008).
[79]D. Datta, J. M. Scheer, M. J. Romanowski, and J. A. Wells, J. Mol. Biol. **381**, 1157 (2008).
[80]S. Brüschweiler, P. Schanda, K. Kloiber, B. Brutscher, G. Kontaxis, R. Konrat, and M. Tollinger, J. Am. Chem. Soc. **131**, 3063 (2009).
[81]G. R. Kneller, K. Hinsen, and P. Calligari, J. Chem. Phys. **136**, 191101 (2012).
[82]P. A. Bopp, A. A. Kornyshev, and G. Sutmann, Phys. Rev. Lett. **76**, 1280 (1996).
[83]I. P. Omelyan, Mol. Phys. **93**, 123 (1998).