

## Redox entropy of plastocyanin: Developing a microscopic view of mesoscopic polar solvation

David N. LeBard and Dmitry V. Matyushov<sup>a)</sup>

*Center for Biological Physics, Arizona State University, PO Box 871604, Tempe, Arizona 85287-1604, USA*

(Received 2 November 2007; accepted 12 March 2008; published online 18 April 2008)

We report applications of analytical formalisms and molecular dynamics (MD) simulations to the calculation of redox entropy of plastocyanin metalloprotein in aqueous solution. The goal of our analysis is to establish critical components of the theory required to describe polar solvation at the mesoscopic scale. The analytical techniques include a microscopic formalism based on structure factors of the solvent dipolar orientations and density and continuum dielectric theories. The microscopic theory employs the atomistic structure of the protein with force-field atomic charges and solvent structure factors obtained from separate MD simulations of the homogeneous solvent. The MD simulations provide linear response solvation free energies and reorganization energies of electron transfer in the temperature range of 280–310 K. We found that continuum models universally underestimate solvation entropies, and a more favorable agreement is reported between the microscopic calculations and MD simulations. The analysis of simulations also suggests that difficulties of extending standard formalisms to protein solvation are related to the inhomogeneous structure of the solvation shell at the protein-water interface combining islands of highly structured water around ionized residues along with partial dewetting of hydrophobic patches. Quantitative theories of electrostatic protein hydration need to incorporate realistic density profile of water at the protein-water interface. © 2008 American Institute of Physics. [DOI: 10.1063/1.2904879]

### I. INTRODUCTION

Calculations of the thermodynamics of hydrated biopolymers present a significant challenge to theoretical algorithms. In many cases the problem can be treated by numerical simulations with force fields assigned to the biomolecule (solute) and water (solvent).<sup>1</sup> The obvious difficulty is the large computational load and still existing uncertainties in the treatment of the long-range electrostatics. The problem, however, becomes more nontrivial when derivatives of thermodynamic potentials, e.g., redox entropy, need to be computed or when the solvation thermodynamics changes on the time and length scale unattainable to standard molecular dynamics (MD) protocols, e.g., in problems related to protein folding.<sup>2,3</sup> In all such cases, coarse graining of the system is required and that can be done on various length scales.<sup>4–6</sup> Dielectric continuum algorithms, solving the boundary Poisson problem, are computationally very efficient. In this approximation, all length scales below the largest distance of microscopic correlations (density and/or polarization) are averaged out into a continuum surrounding the solute. These approaches are normally represented by either direct solution of the Poisson–Boltzmann equation on the real-space grid<sup>7</sup> or even more approximate formalisms under the umbrella of the generalized Born approximation.<sup>8</sup>

When the cavity cut out by the solute from the continuum dielectric is properly parametrized, equations of continuum electrostatics provide a reasonable estimate of the solvation Gibbs energy.<sup>9–11</sup> The fundamental problem of this

approach is that the local structure of the solvent around the solute, averaged out in the continuum limit, is what effectively forms the dielectric cavity. While this structure can be parametrized by choosing proper van der Waals (vdW) radii,<sup>12</sup> this parametrization needs to be redone every time the thermodynamic state of the solvent changes. This difficulty makes continuum formalisms unreliable for the calculation of derivatives of the Gibbs energy, for instance, the entropy of solvation.<sup>13–16</sup> In addition, the surface of a protein is highly heterogeneous combining hydrophobic patches exposing nonpolar residues and hydrophilic patches made of ionized/polar residues. While the water structure is rigid around ionized residues, probably resembling the well-studied case of solvation around simple ions, water is much less structured at hydrophobic patches with the potential for dewetting<sup>17</sup> or/and oscillations of the water occupation.<sup>18,19</sup> It is clear that simplistic continuum does not represent this complex reality,<sup>3</sup> and one needs to incorporate the ability of the solvent to fluctuate into the solvation model.

The goal of this paper is to extend the microscopic view of solvation in polar solvents, which we have been developing in the past in application to small- and medium-size solutes,<sup>14,15,20</sup> to solvation of solutes of mesoscopic dimension, biopolymers in the first place. The length scale of this problem presents an obvious obstacle to numerical simulation techniques. On the other hand, the same length scale allows one to hope that some of the short-range features of the solvent structure around the solute, making solvation of small molecules so specific, might average out on a larger scale. If this averaging is realized for solvation of biopolymers, it would allow coarse-grained models to efficiently op-

<sup>a)</sup>Electronic mail: dmitrym@asu.edu.

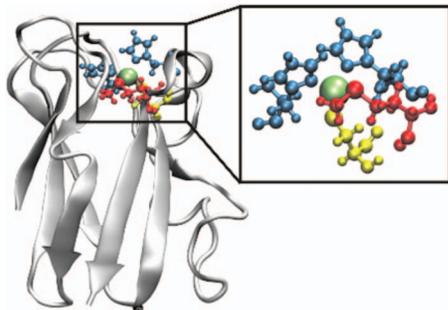


FIG. 1. (Color) Structure of plastocyanin: the active site includes copper ion (green), 2 histidines (blue), methionine (red), and cysteine (orange) residues.

erate in this field complementing direct numerical simulations. Our approach to the problem is to coarse grain the solvent response into a number of solvent correlation functions (structure factors) representing the nuclear modes affecting electrostatic solvation. The microscopic nature of the solvent response is then incorporated into the wave-vector dependence of these structure factors efficiently filtering out the length scales insignificant for solvation.

This study poses the central question for the future development of such techniques: What are the solvent modes which play the central role in the thermodynamics of mesoscopic polar solvation and what are the theory ingredients critical for capturing the basic physics of large-scale solvation? We study this problem here by carrying out extensive MD simulations of solvation of plastocyanin (PC) in TIP3P water in the temperature range of 280–310 K. This fully atomistic approach is compared to continuum electrostatics and to our microscopic algorithm, operating with  $\mathbf{k}$ -space correlation functions, which was designed to scale efficiently on the mesoscopic length scale.

Plastocyanin from spinach is a single polypeptide chain of 99 residues forming a  $\beta$ -sandwich, with a single copper ion coordinated by two sulfurs from cysteine and methionine and two nitrogens from histidine residues (Fig. 1). The presence of the copper ion, which can change redox state, allows PC to function as a mobile electron carrier in the photosynthetic apparatus of plants and bacteria. It accepts an electron from ferredoxin *f* and diffusively carries it to another docking location where the electron is donated to the oxidized form of photosystem I.<sup>21</sup>

Because of high solvent accessibility of the redox site, PC presents a convenient model to study the hydration effect on the redox properties of metalloproteins, a situation also encountered for ferredoxins.<sup>22,23</sup> The redox thermodynamics of PC has been characterized experimentally<sup>24–26</sup> and combined quantum/simulation calculations have been done as well.<sup>27,28</sup> The early focus of the theoretical studies had been on unusually high redox potentials of copper proteins, which was assigned to the nontraditional distorted tetrahedral coordination on the copper ion.<sup>29–31</sup> In particular, the Cu–S bond to methionine is unusually long and is actually broken in the reduced state of PC at  $\text{pH} < 3.8$ .<sup>32</sup> The protein is also highly charged at  $\text{pH} \approx 7$  (–9.0 in reduced state and –8.0 in oxidized states). The charge is made by 15 negatively charged deprotonated residues (nine glutamic and six aspartic acids)

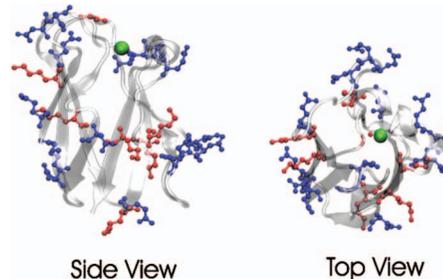


FIG. 2. (Color) Distribution of the positive and negative charge on the surface of the protein. The positively and negatively charged residues are shown, respectively, in red and blue. The copper ion is shown in green.

and six positively charged lysine residues with amino groups protonated (Fig. 2). The asymmetric charge distribution located on the protein surface creates the dipole moment of 2200 D in the oxidized state (Ox) and of 2470 D in the reduced state (Red), both numbers are calculated relative to the center of partial charges.

The redox potential of the protein includes a component from the local ligand field of the active site and the Gibbs energy of solvation. The computation of the former requires quantum mechanics, making the problem of calculating the overall redox potential a very nontrivial exercise.<sup>23,28,33,34</sup> Calculations of solvation thermodynamics can be reasonably accomplished using partial atomic charges parametrized from quantum calculations in the vacuum. The experimental input comes from measurements of redox entropy<sup>24,25,35</sup> since the temperature-independent ligand-field component is expected to vanish in the temperature derivative.

## II. MICROSCOPIC SOLVATION MODEL

The principal idea of the microscopic solvation model is to reduce the problem of solvation of an arbitrary solute in a polar solvent to a formalism combining two major blocks: Electrostatics of an isolated solute and nonlocal correlation functions of the pure solvent. The idea of assembling separate solute and solvent properties in a solvation model is obviously not new going back to Born<sup>36</sup> and Onsager<sup>37</sup> and all the subsequent development of continuum electrostatics in application to solvation.<sup>7–10,38</sup> The advantage of our approach is in avoiding the necessity to know the microscopic solute-solvent structure, which is the main complexity of microscopic solvation models and is also their main advantage when the problem is successfully resolved by either solving integral equations<sup>39</sup> or by applying time-dependent<sup>40,41</sup> or equilibrium<sup>42–44</sup> density functional methods. Inserting a solute into a dense liquid creates a significant distortion of its structure, and the incomplete account of the coupling between the short-range density profile around the solute with the long-range polarization field is perhaps the weakest part of our formulation when applied to small solutes.<sup>45</sup> [This deficiency is almost completely offset by averaging of the density profile around a nanoscale solute (see below).] On the other hand, a strong side of our formalism, its ability to treat solvation of large solutes of irregular shape and arbitrary charge distribution,<sup>15,46</sup> becomes particularly useful in application to protein solvation.

The reduction of the many-body solvation problem to an irreducible representation in terms of a few basic correlation functions depends on the symmetry of the solute-solvent interaction potential. The number of correlation functions is known to grow with increasing the rank of the solvent multipole included in the interaction potential.<sup>47</sup> Solvent dipoles are for the most part sufficient for solvation in polar liquids,<sup>48</sup> in which case the solute-solvent interaction potential  $V_{0s}$  (“0” and “s” are used for the solute and solvent, respectively) is a sum of pairwise interactions of the solute electric field  $\mathbf{E}_0(\mathbf{r})$  with the solvent dipoles,

$$V_{0s} = - \sum_{j=1}^N \mathbf{m}'_j \cdot \mathbf{E}_0(\mathbf{r}_j). \quad (1)$$

Here,  $\mathbf{m}'_j$  is the dipole moment characteristic of the bulk state of the solvent;  $m'$  is usually higher than the vacuum dipole  $m$  because of the collective field of the induced solvent dipoles.<sup>49</sup> For instance, the dipole moment of water in the liquid state, 2.4–2.6 D,<sup>50</sup> is higher than the gas-phase dipole of 1.83 D.

We will focus on the electrostatic component of the chemical potential of solvation  $\mu_{0s}$  which contains all the information relevant to electrostatic solvation. Linear response approximation (LRA) significantly simplifies the problem and provides several equivalent routes to  $\mu_{0s}$ . One can consider the full interaction between the atomic charges of the solute and the solvent and determine  $\mu_{0s}$  as half of the average solute-solvent interaction energy,<sup>51</sup>  $\mu_{0s} = \langle V_{0s} \rangle / 2$ . Alternatively, one can use the second cumulants,  $\langle (\delta V_{0s})^2 \rangle$  or  $\langle (\delta V_{0s})^2 \rangle_0$ .<sup>14</sup> In the first expression, the angular brackets  $\langle \dots \rangle$  refer to an ensemble average over the fluctuations  $\delta V_{0s}$  in the solvent in equilibrium with the full charge distribution of the solute. For the second expression,  $\langle \dots \rangle_0$  implies that all the charges of the solute have been set to zero, and fluctuations of the solvent in the solute vicinity are regulated only by short-range solute-solvent interactions, molecular repulsions in the first place. In the LRA, the two cumulants are equal,<sup>52</sup> which physically means that the solute electrostatic forces do not significantly change the solvent structure around the solute established by the prevalence of short-range repulsions.<sup>53</sup> Computer simulations for the most part support this picture<sup>54–56</sup> with a few exceptions of very strong solute-solvent electrostatic coupling found for small solutes.<sup>16,52,57</sup>

This observation opens up a significant simplification of the calculation algorithms. Instead of solving the inhomogeneous problem of restructuring the solvent in an external field of the solute, it appears to be sufficient to look at the statistics of solvent fluctuations around the repulsive core of the solute. This strategy is used here and we will base our calculations on the relation

$$- \mu_{0s} = (\beta/2) \langle (\delta V_{0s})^2 \rangle_0, \quad (2)$$

where  $\delta V_{0s} = V_{0s} - \langle V_{0s} \rangle_0$  and  $\beta = 1/(k_B T)$ .

By using the interaction potential according to Eq. (1), one can rewrite Eq. (2) in the form typical for Gaussian (LRA) models of solvation,<sup>58,59</sup>

$$- \mu_{0s} = \frac{1}{2} \tilde{\mathbf{E}}_0(\mathbf{k}_1) * \chi(\mathbf{k}_1, \mathbf{k}_2) * \tilde{\mathbf{E}}_0(\mathbf{k}_2). \quad (3)$$

Here, the two-rank tensor  $\chi(\mathbf{k}_1, \mathbf{k}_2)$  is the response function<sup>60</sup> of the system composed of a dipolar solvent and a solute to a weak field of the solute. The inhomogeneous character of the problem is reflected by the fact that  $\chi(\mathbf{k}_1, \mathbf{k}_2)$  depends on two wave vectors,  $\mathbf{k}_1$  and  $\mathbf{k}_2$ , separately and not on  $\mathbf{k}_1 - \mathbf{k}_2$ , as is the case with response functions of homogeneous solvents. The asterisk in Eq. (3) refers to both the tensor contraction and integration in inverted  $\mathbf{k}$  space. In addition,  $\tilde{\mathbf{E}}_0(\mathbf{k})$  is the Fourier transform of the electric field of the solute defined by the integral limited to the solvent volume  $\Omega$

$$\tilde{\mathbf{E}}_0(\mathbf{k}) = \int_{\Omega} \mathbf{E}_0(\mathbf{r}) e^{i\mathbf{k} \cdot \mathbf{r}} d\mathbf{r}. \quad (4)$$

The shape of the solute thus enters both the response function  $\chi(\mathbf{k}_1, \mathbf{k}_2)$  and the field Fourier transform  $\tilde{\mathbf{E}}_0(\mathbf{k})$ . The charge distribution of the solute, which determines the electric field  $\tilde{\mathbf{E}}_0(\mathbf{k})$ , is given by its electronic density and is commonly represented by partial atomic charges.

The main challenge of this formalism, as well as of other Gaussian solvation theories,<sup>61</sup> is how to connect the inhomogeneous response function  $\chi(\mathbf{k}_1, \mathbf{k}_2)$  to the shape of the solute repulsive core and the self-correlation functions of the solvent modes affecting solvation. Two modes naturally appear in most theories: Dipolar (orientational) polarization and density fluctuations.<sup>20,43,62,63</sup> For the former, the combination of axial symmetry introduced by the wave vector  $\mathbf{k}$  with the vector character of the dipolar polarization  $\mathbf{P}(\mathbf{k})$  allows one to split the two-rank tensor  $\chi_s(\mathbf{k}) = (\beta/\Omega) \times \langle |\delta \mathbf{P}(\mathbf{k})|^2 \rangle$  into the longitudinal and transverse dyads,<sup>64</sup>

$$\chi_s(\mathbf{k}) = \frac{3y}{4\pi} [\mathbf{J}^L S^L(k) + \mathbf{J}^T S^T(k)], \quad (5)$$

where  $\mathbf{J}^L = \hat{\mathbf{k}}\hat{\mathbf{k}}$ ,  $\mathbf{J}^T = 1 - \hat{\mathbf{k}}\hat{\mathbf{k}}$ . In Eq. (5),  $y$  is the effective density of both permanent and induced dipoles in the liquid which commonly appears in theories of dielectrics,<sup>49</sup>

$$y = (4\pi/9)\beta m'^2 \rho + (4\pi/3)\rho \alpha. \quad (6)$$

In Eq. (6),  $\alpha$  is the dipolar polarizability of the solvent particle.

The scalar functions  $S^L(k)$  and  $S^T(k)$  in Eq. (4) are, correspondingly, the longitudinal and transverse structure factors of dipolar fluctuations of the homogeneous solvent (see below). The  $k=0$  values of these structure factors are related to the dielectric constant  $\epsilon_s$  by the following equations:

$$S^L(0) = (\epsilon_s - 1)/(3y\epsilon_s), \quad (7)$$

$$S^T(0) = (\epsilon_s - 1)/(3y).$$

Also, the trace of  $\chi_s(0)$  over the Cartesian projections gives the Kirkwood  $g$  factor,<sup>65</sup>

$$g_K = \frac{1}{3} [S^L(0) + 2S^T(0)]. \quad (8)$$

The expansion of the solvation chemical potential in the Mayer functions corresponding to the solute-solvent interaction potential leads to the following form for the response function:<sup>14,20</sup>

$$\chi(\mathbf{k}_1, \mathbf{k}_2) = \chi_p(\mathbf{k}_1, \mathbf{k}_2) + \chi_d(\mathbf{k}_1, \mathbf{k}_2), \quad (9)$$

where

$$\chi_p(\mathbf{k}_1, \mathbf{k}_2) = \chi_s(\mathbf{k}_1) \delta_{\mathbf{k}_1, \mathbf{k}_2}. \quad (10)$$

In Eq. (10),  $\Omega \delta_{\mathbf{k}_1, \mathbf{k}_2} = (2\pi)^3 \delta(\mathbf{k}_1 - \mathbf{k}_2)$  is the Kronecker symbol and  $\chi_d(\mathbf{k}_1, \mathbf{k}_2)$  in Eq. (9) is the component of the response originating from the local fluctuations of the solvent density around the solute,<sup>20</sup>

$$\chi_d(\mathbf{k}_1, \mathbf{k}_2) = (3y/8\pi)(1 - S(k_1))\theta_0(\mathbf{k}_1 - \mathbf{k}_2). \quad (11)$$

Here,  $S(k) = N^{-1}(|\delta\rho(\mathbf{k})|^2)$  is the density-density structure factor of the homogeneous solvent and  $N$  is the number of solvent molecules. In addition,  $\theta_0(\mathbf{k})$  is the Fourier transform of the step function  $\theta_0(\mathbf{r})$  defining the solute shape. It is equal to unity for  $\mathbf{r}$  inside the solute and is zero otherwise.

The problem with the direct perturbation result in Eq. (9) is that it contains the transverse polarization response function  $\propto S^T(k)$  diverging in its continuum,  $k \rightarrow 0$ , limit as the solvent dielectric constant goes to infinity [Eq. (7)]. The problem is really caused by the nonspherical shape of the solute. The electric field of the solute charges is longitudinal. However, when the symmetry of the solute is different from the symmetry of the charge distribution in a sense that the cavity boundary does not coincide with the equipotential surface, the Fourier integral in Eq. (4) generates a transverse component of  $\tilde{\mathbf{E}}_0(\mathbf{k})$ . Notice that this is always the case when electron transfer reactions are considered.<sup>66</sup> A transverse component in  $\tilde{\mathbf{E}}_0(\mathbf{k})$  then results in a “transverse catastrophe” for solvents of high polarity. The problem was well recognized in early studies<sup>20,62,63</sup> which suggested to use only the longitudinal component of the field  $\tilde{\mathbf{E}}_0(\mathbf{k})$ . As a matter of fact, the problem lies in the response function of the dipolar polarization field which needs to be renormalized with the account of the solute repulsive core, a procedure similar to applying boundary conditions to the Poisson equation of continuum electrostatics.

The Li-Kardar-Chandler<sup>59,67</sup> Gaussian model allows one to achieve a correct renormalization of the inhomogeneous polarization response function  $\chi_p(\mathbf{k}_1, \mathbf{k}_2)$  eliminating the “transverse catastrophe.”<sup>68</sup> This approach introduces another simplification by replacing all the short-range solute-solvent interactions by hard-core repulsions. This simplification, however, leads to an exact solution for the  $\mathbf{k}$ -space response functions with the result<sup>68</sup>

$$\chi_p(\mathbf{k}_1, \mathbf{k}_2) = \chi_s(\mathbf{k}_1) \delta_{\mathbf{k}_1, \mathbf{k}_2} - \chi''(\mathbf{k}_1) \theta_0(\mathbf{k}_1 - \mathbf{k}_2) \chi_s(\mathbf{k}_2). \quad (12)$$

The second summand in Eq. (12) is the correction of the response function of the homogeneous solvent, appearing in Eq. (10), by the repulsive core of the solute. The response function  $\chi''(\mathbf{k}_1)$  then incorporates both  $\chi_s$  and the information about the solute shape.<sup>14,68</sup>

A direct substitution of Eq. (12) into Eq. (3) results in a six-dimensional (6D) integral convolution in  $\mathbf{k}$ -space which

is not numerically tractable. In order to arrive at a computationally efficient procedure, a mean-field approximation was introduced,<sup>14</sup> which replaces the inhomogeneous electric field of the solvent inside the solute by a mean cavity field,

$$\mathbf{F}_0 = \frac{f}{8\pi} \int_{\Omega} \mathbf{E}_0 \cdot \mathbf{D}_r \frac{d\mathbf{r}}{r^3}, \quad (13)$$

where

$$f = \frac{2(\epsilon_s - 1)}{2\epsilon_s + 1}. \quad (14)$$

Here,  $\mathbf{D}_r = 3\hat{\mathbf{r}}\hat{\mathbf{r}} - 1$  is the two-rank dipolar tensor with  $\hat{\mathbf{r}} = \mathbf{r}/r$ .  $\mathbf{F}_0$  becomes the Onsager reaction field<sup>37</sup> for a spherical solute with point dipole located at the center.

The mean-field approximation reduces the problem of calculating the solvation thermodynamics to a numerically tractable three-dimensional integral in  $\mathbf{k}$  space. The chemical potential of solvation then becomes a sum of two components arising from the longitudinal ( $L$ ) and transverse ( $T$ ) polarization fluctuations,  $\mu_{0s}^{L,T}$ , and a third component arising from the density fluctuations,  $\mu_{0s}^d$ ,

$$\mu_{0s} = \mu_{0s}^L + \mu_{0s}^T + \mu_{0s}^d. \quad (15)$$

The transverse component  $\mu_{0s}^T$  is defined by the  $\mathbf{k}$  integral of the transverse projection of the solute field,  $\tilde{\mathbf{E}}_0^T(\mathbf{k})$ , with the transverse polarization structure factor,

$$-\mu_{0s}^T = g_K^{-1} S^L(0) \frac{3y}{8\pi} \int \frac{d\mathbf{k}}{(2\pi)^3} |\tilde{\mathbf{E}}_0^T(\mathbf{k})|^2 S^T(k). \quad (16)$$

The transverse field component is defined by subtracting the longitudinal projection,

$$\tilde{\mathbf{E}}_0^L(\mathbf{k}) = \hat{\mathbf{k}}(\hat{\mathbf{k}} \cdot \tilde{\mathbf{E}}_0), \quad (17)$$

from the total inverted-space field  $\tilde{\mathbf{E}}_0$ ,

$$\tilde{\mathbf{E}}_0^T(\mathbf{k}) = \tilde{\mathbf{E}}_0(\mathbf{k}) - \tilde{\mathbf{E}}_0^L(\mathbf{k}). \quad (18)$$

Equation (16) is the main result of the application of the Gaussian model<sup>59</sup> to polar solvation. It replaces  $S^T(k)$  of the direct perturbation expansion in Eqs. (9) and (10) with the renormalized function,

$$3yS^T(k) \rightarrow (3yS^L(0)/g_K)S^T(k). \quad (19)$$

The  $k=0$  limit of the transverse response changes from  $3yS^T(0) = \epsilon_s - 1$  to  $3(\epsilon_s - 1)/(2\epsilon_s + 1)$  thus eliminating the “transverse catastrophe” of direct perturbation expansions.

The component  $\mu_{0s}^L$  of the solvation chemical potential in Eq. (15) is obtained by inverted-space integration with the longitudinal polarization structure factor:

$$-\mu_{0s}^L = \frac{3y}{8\pi} \int \frac{d\mathbf{k}}{(2\pi)^3} S^L(k) \left[ |\tilde{\mathbf{E}}_0^L|^2 - |\tilde{\mathbf{E}}_0^T|^2 f \frac{\mathbf{F}_0 \cdot \tilde{\mathbf{E}}_0^L}{\mathbf{F}_0 \cdot \tilde{\mathbf{E}}_0^T} \right]. \quad (20)$$

There is a significant physics behind the appearance of the transverse field in the brackets of Eq. (20). Longitudinal dipolar polarization is short ranged and thus does not propagate over macroscopic distances. On the contrary, transverse polarization is long ranged. Therefore, inducing transverse

polarization modifies the electric field acting on the solvent dipoles resulting in the second term in the brackets in Eq. (20).

The density component in Eq. (15) can formally be obtained by multiplying the response function  $\chi_d(\mathbf{k}_1, \mathbf{k}_2)$  [Eq. (11)] with the Fourier transforms of the electric field and integrating over  $\mathbf{k}_1$  and  $\mathbf{k}_2$ . This, however, results in a 6D convolution integral to be avoided in numerical applications. An alternative approach is to use direct-space integration when the density component becomes

$$-\mu_{0s}^d = 3yF(\mathbf{r}) * \mathcal{F}^{-1}[(1 - S(k))\theta_0(\mathbf{k})]. \quad (21)$$

Here,  $F(\mathbf{r}) = (8\pi)^{-1}E_0^2(\mathbf{r})$  is the density of the electrostatic field energy and the asterisk indicates integration in real space over the volume  $\Omega$  occupied by the solvent. In addition,  $\mathcal{F}^{-1}$  is the inverse Fourier transform of the function indicated in the brackets.

Solvation by the overall dipolar polarization, including nuclear and electronic components, was considered in the formalism outlined above. For solvation problems relevant to spectroscopy and charge-transfer reactions nuclear component of polarization needs to be extracted. This is achieved by replacing the density  $y$  [Eq. (6)] of all, permanent and induced, dipoles in the equations above with the density of permanent dipoles only,  $y \rightarrow y_p = (4\pi/9)\beta m'^2 \rho$ . In addition, the  $k=0$  values of the structure factors need to be modified to account for screening of the dipolar interactions by the high-frequency dielectric constant  $\epsilon_\infty$ . The  $k=0$  values for these nuclear structure factors,  $S_n^{L,T}(k)$ , now become<sup>15</sup>

$$S_n^L(0) = (\epsilon_\infty^{-1} - \epsilon_s^{-1})/(3y_p), \quad (22)$$

$$S_n^T(0) = (\epsilon_s - \epsilon_\infty)/(3y_p).$$

Once the  $k=0$  values for the structure factors are fixed by Eqs. (7) and (22), the scalar functions  $S^L(k)$  and  $S^T(k)$  can be calculated from our parametrization scheme, parametrized polarization structure factors (PPSF).<sup>14</sup> This analytical route to the polarization structure factors is tested here by comparing the results of solvation calculations employing the PPSF to the direct use of  $S^{L,T}(k)$  from MD simulations (see below).

### III. COMPUTATIONAL ALGORITHM

The computational algorithm is outlined in Fig. 3. The solute is parametrized by coordinates  $\mathbf{r}_j$ , van der Waals (vdW) radii  $a_j$ , and partial charges  $q_j$  of the atoms. The electric field of the solute is calculated at points  $\mathbf{r}_n$  of the  $N \times N \times N$  grid built on the  $L \times L \times L$  cube,

$$\mathbf{E}_0(\mathbf{r}_n) = \sum_{j=1}^{N_q} \frac{q_j(\mathbf{r}_n - \mathbf{r}_j)}{|\mathbf{r}_n - \mathbf{r}_j|^3}, \quad (23)$$

where  $N_q$  is the number of solute charges. The array  $\mathbf{E}_0(\mathbf{r}_n)$  is converted to inverted space by using fast Fourier transform technique.<sup>69</sup> The field  $\tilde{\mathbf{E}}_0(\mathbf{k})$  is split into longitudinal and transverse components and used in  $\mathbf{k}$  integration in Eqs. (16) and (20) with the corresponding structure factors of the dipolar polarization. As is illustrated in Fig. 3, the calculation input is subdivided into two separate components related to

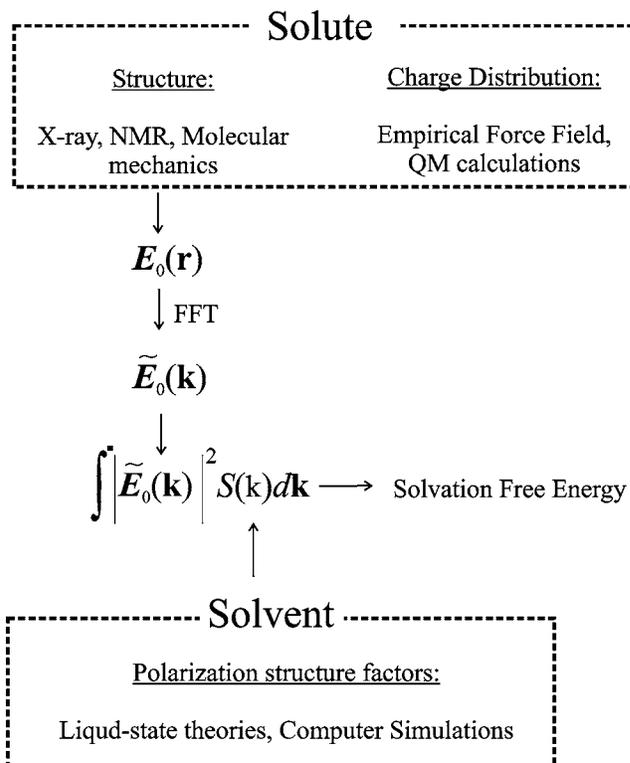


FIG. 3. Diagram of the computational algorithm.

the solute and solvent properties. Details of the calculations for each of these are given below.

#### A. Solute

The Fourier transform of the Coulomb field is conditionally convergent. Therefore, in order to avoid numerical divergence in the Fourier integral, real space is divided into three regions: Hard core of the solute (region 1), region outside a sphere of radius  $R$  (region 3), and the region between the solute surface and the sphere (region 2, Fig. 4). The Fourier integral then becomes

$$\tilde{\mathbf{E}}_0(\mathbf{k}) = \int_{r < R} \mathbf{E}_0(\mathbf{r}) e^{-i\mathbf{k} \cdot \mathbf{r}} d\mathbf{r} + \tilde{\mathbf{E}}_R, \quad (24)$$

where the first integral is taken over region 2 and the second integral is over region 3:

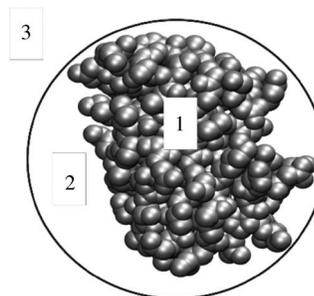


FIG. 4. Separation of real space into regions for the calculation of the Fourier transform of the solute electric field [Eq. (24)]. The Fourier transform is calculated numerically in region 2 and analytically [Eq. (27)] in region 3. The field is set equal to zero within the hard repulsive core of the solute (region 1).

TABLE I. Atomic partial charges for copper and its four ligands in the reduced (Red) and oxidized (Ox) states of PC.

Set	Red					Ox				
	Cu	$N_\delta^a$	$N_\delta^b$	$S_\gamma^c$	$S_\delta^d$	Cu	$N_\delta^a$	$N_\delta^b$	$S_\gamma^c$	$S_\delta^d$
I	1.0	-0.7	-0.7	-1.23	-0.09	2.0	-0.7	-0.7	-1.23	-0.09
II	-0.49	-0.445	-0.495	-0.369	-0.24	0.35	-0.42	-0.47	-0.26	-0.24

<sup>a</sup>His87.<sup>b</sup>His37.<sup>c</sup>Cys84.<sup>d</sup>Met92.

$$\tilde{\mathbf{E}}_R(\mathbf{k}) = \int_{r>R} \mathbf{E}_0(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}} d\mathbf{r}. \quad (25)$$

The center of the sphere is taken at the center of the charge distribution defined by the relation

$$\mathbf{r}_q = \frac{\sum_{j=1}^{N_q} q_j \mathbf{r}_j}{\sum_{j=1}^{N_q} q_j}. \quad (26)$$

The radius of the cutoff sphere,  $R$ , is chosen to minimize the part of the grid which is used in numerical calculation of the Fourier transform. In our calculations, the radius  $R$  is chosen by adding the solvent diameter  $\sigma$  to the largest distance from  $\mathbf{r}_q$  to the solvent-accessible surface (SAS) of the solute (vdW radii of the surface atoms plus the radius of the solvent molecule).

The Fourier transform outside the sphere can be evaluated analytically. For the location of charges relative to the center of charge given as  $\mathbf{s}_j = \mathbf{r}_j - \mathbf{r}_q$ , the solution for  $\tilde{\mathbf{E}}_R$  can be obtained by expanding  $\mathbf{E}_0(\mathbf{r})$  in  $s_j/R < 1$ ,

$$\begin{aligned} \tilde{\mathbf{E}}_R(\mathbf{k}) = & -4\pi e \sum_j q_j \sum_{n=1}^{\infty} \left( \frac{s_j}{R} \right)^{n-1} \\ & \times \frac{j_{n-1}(kR)}{k} [\hat{\mathbf{s}}_j P'_{n-1}(\cos \theta_j) - \hat{\mathbf{k}} P'_n(\cos \theta_j)]. \end{aligned} \quad (27)$$

Here,  $\cos \theta_j = \hat{\mathbf{s}}_j \cdot \hat{\mathbf{k}}$ ,  $j_n(x)$  is the spherical Bessel function, and  $P_n(\cos \theta_j)$  is the Legendre polynomial;  $P'_n(x)$  denotes the derivative of  $P_n(x)$ .

## B. Charging scheme

The formal charge of the copper ion is +2 and that of the cysteine sulfur is -1 in the oxidized state of PC. The charge is, however, delocalized among the ligands and the metal center. The main factor in this delocalization is a strong covalency of a copper-sulfur (Cys) bond. Calculations by Solomon and Lowery<sup>70</sup> assign 40% of spin density of an unpaired electron to copper and 36% to cysteine's sulfur in the ground state of oxidized protein. The extent of delocalization varies significantly depending on the level of quantum mechanical calculations used.<sup>71-73</sup> The electron-nuclear double resonance experiments,<sup>74</sup> which require additional calibration on quantum calculations, result in the following net charges on the residues coordinating copper:<sup>75</sup> -0.25 (His), -0.51 (Cys), and -0.04 (Met). The more recent map-

ping of the electron spin density to NMR relaxation<sup>29</sup> gives a much lower extent of delocalization: -0.11 (Cys), -0.025 (His), and 0 (Met).

The uncertainties in the extent of electron delocalization pose the question of their impact on the calculation of the redox thermodynamics. In order to study this question, we have performed calculations of the solvation part of the redox potential and the corresponding entropy using different charge sets. Set I is chemically fake assuming charge +2 on copper in Ox state and the net charge of -1 on cysteine. The negative charge is placed on cysteine's sulfur in addition to -0.23 from CHARMM22 protein parametrization. The rest of the protein charges are from the standard CHARMM parametrization. The reduced state for set I is obtained by changing the metal charge to +1. The charges for copper and its four ligands are summarized in Table I. Set II is based on the charging scheme listed by Ullmann *et al.*<sup>72</sup> for the oxidized state of PC. The reduced form is obtained by placing an extra negative charge on copper and its three ligands,  $N_\delta$  (His87),  $N_\delta$  (His37),  $S_\gamma$  (Cys84), and  $S_\gamma$  (Met92) in proportion extracted from NMR experiments (Table I).<sup>29</sup> Finally, a third charge distribution is completely parametrized at the density functional theory level for the charges and force constants of the copper and ligand atoms and consistent with the Amber force field.<sup>76</sup> In addition, Amber FF03 parametrization<sup>77</sup> was applied to all nonligand residues (set II). There were various numbers of TIP3P water molecules for each of the charge distributions: 5 874 (set I), 5 886 (set II), and 4 628 (set III).

We ran separate simulations (about 5 ns) for each charging scheme to find that the results are not strongly affected by the choice of atomic charges (Table II). This was also noticed in some other recent simulations.<sup>78,79</sup> We have therefore implemented charge scheme II in all simulations of metalloproteins reported here since it presents a reasonable balance between being simple and realistic.

## C. Solvent

The polarization structure factors entering the equations for the solvation chemical potential are characteristics of the homogeneous solvent. They can be obtained numerically by averaging the projections of dipole moments  $\hat{\mathbf{e}}_j$  on an arbitrary chosen direction of the  $\mathbf{k}$  vector,  $\hat{\mathbf{k}} = \mathbf{k}/k$ ,

TABLE II. Temperature-dependent  $\langle V_{0s} \rangle_{Ox}$  (eV) for PC(Ox). The results are obtained from MD simulations, DELPHI calculations (with vdW and SS cavities), and from NRFT calculations. The calculations were done with three charge distributions of the active site (I–III, see text for description).

$T$ (K)	MD <sup>a</sup>			DelPhi(vdW)			DelPhi(SS)			NRFT <sup>b</sup>
	I	II	III	I	II	III	I	II	III	II
280	-68.29	-69.46	-68.37	-104.86	-102.07	-99.76	-46.90	-47.47	-46.37	-104.6(-37.2)
285		-69.69		-104.81	-102.03	-99.71	-46.88	-47.45	-46.36	-103.3(-36.5)
290	-70.73	-66.01		-104.76	-101.98	-99.67	-46.87	-47.43	-46.34	-102.1(-35.8)
295		-66.97		-104.71	-101.93	-99.62	-46.84	-47.41	-46.32	-100.7(-35.1)
300	-67.06	-66.09	-65.84	-104.65	-101.88	-99.57	-46.81	-47.39	-46.30	-99.3(-34.5)
305		-68.58		-104.59	-101.82	-99.52	-46.80	-47.37	-46.28	-98.3(-33.9)
310		-67.51	-66.68	-104.53	-101.77	-99.46	-46.79	-47.35	-46.26	-97.1(-33.3)

<sup>a</sup>Average interaction energy of PC(Ox) with the water molecules in the simulation cell. The energy of the protein becomes more negative by  $\approx 25$  eV when the interactions of the protein atoms with the sodium ions neutralizing the system are included (see also Fig. 7). The simulation protocol does not allow to produce the Debye–Hückel screening of the Coulomb interactions and this shift might not represent the real situation with proteins dissolved in buffered solutions.

<sup>b</sup>The numbers in the parentheses indicate the density component of the equilibrium interaction energy,  $\langle V_{0s} \rangle_{Ox}^d$ .

$$S^L(k) = \frac{3}{N} \left\langle \sum_{ij} (\hat{\mathbf{e}}_j \cdot \hat{\mathbf{k}}) (\hat{\mathbf{k}} \cdot \hat{\mathbf{e}}_i) e^{i\mathbf{k} \cdot \mathbf{r}_{ij}} \right\rangle, \quad (28)$$

$$S^T(k) = \frac{3}{2N} \left\langle \sum_{ij} [(\hat{\mathbf{e}}_j \cdot \hat{\mathbf{e}}_i) - (\hat{\mathbf{e}}_j \cdot \hat{\mathbf{k}}) (\hat{\mathbf{k}} \cdot \hat{\mathbf{e}}_i)] e^{i\mathbf{k} \cdot \mathbf{r}_{ij}} \right\rangle,$$

where  $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$  and  $N$  is the number of liquid dipoles.

Unfortunately, experiment does not provide spatially resolved correlators of dipoles in polar liquids and one has to resort to using either computer simulations or liquid-state theories. Parametrizing homogeneous structure factors by computer simulations is a very attractive avenue for studying hydration because of continuously improving empirical potentials for water<sup>80</sup> on one hand and the reliance of biological applications on aqueous solvation on the other. For the problems related to derivatives of thermodynamic potentials, e.g., entropy and volume of solvation, the structure factors need to be tabulated at different temperatures and/or pressures. In this paper, polarization structure factors  $S^{L,T}(k)$  were obtained from  $\approx 25$  ns MD trajectories of TIP3P water<sup>81</sup> at different temperatures (Fig. 5). Neumann's algorithm<sup>82</sup> of finite-size corrections for the cutoff of long-range electrostatic interactions was employed in the analysis of the simulation results. The results for the dielectric constant of TIP3P water obtained from this analysis are consistent with other literature sources (see also Ref. 15).

In parallel to simulations, we have used dipolar structure factors from our PPSF parametrization scheme.<sup>14</sup> This approach is based on the analytical solution of the mean-spherical approximation for the fluid of dipolar hard spheres<sup>83</sup> which is parametrized to give  $k=0$  values from Eqs. (7) and (22).<sup>14,15</sup> The PPSF parametrization gives solvation free energies essentially identical to those obtained with the structure factors from simulations, as was also found for a smaller polypeptide solute in our previous publication.<sup>15</sup> It appears therefore that the local tetrahedral order of water, which is, of course, not captured by dipolar hard spheres, is not significant for the energetics of polar solvation dominated by orientational correlations ruled by dipole-dipole forces. The density fluctuations are, on the other hand, dominated by repulsions. We used, therefore, the

density structure factor from the Percus–Yevick solution<sup>60</sup> for a hard-sphere fluid of the same density as water in calculations of the density response in Eq. (21).

We need to emphasize here that using dielectric constants of model dipolar fluids in the structure factors would give us wrong results. The dielectric constant of a molecular liquid is affected by short-range molecular correlations through the Kirkwood factor; hydrogen bonds and molecular quadrupoles are among significant factors.<sup>49</sup> We account for all these effects in the PPSF scheme by using the experimental, either from computer or laboratory data, dielectric constants in Eqs. (7) and (22). Once the  $k=0$  limit is set up by the experimental input, the behavior of  $S^{L,T}(k)$  in the range of  $k < 2\pi/\sigma$  ( $\sigma$  is the solvent diameter) is well reproduced by solutions obtained for model dipolar fluids. These solutions will fail at  $k \geq 2\pi/a$ , where  $a$  is the characteristic distance between partial charges within the solvent molecule. This part of the spectrum of polarization fluctuations is correctly captured by models based on interaction-site integral

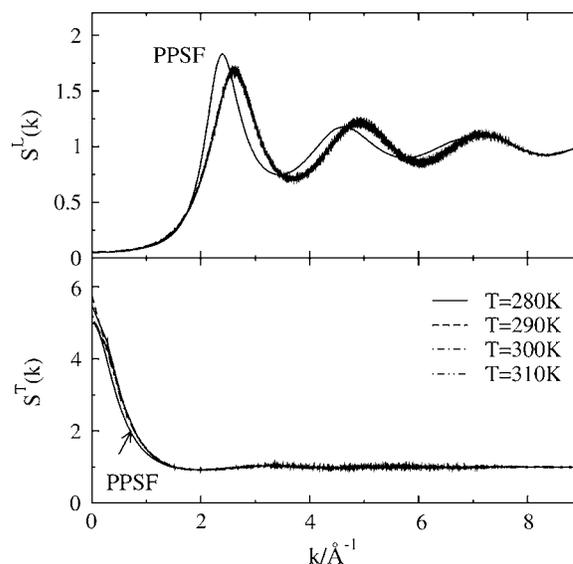


FIG. 5. Longitudinal ( $L$ ) and transverse ( $T$ ) polarization structure factors of TIP3P water [Eq. (28)] calculated at different temperatures from  $\approx 25$  ns MD trajectories. Also shown is the PPSF calculation at  $T=300$  K.

equations,<sup>39</sup> but that range of wave vectors normally does not contribute to the solvation energy. In fact, the range of  $k$  values relevant for the solvation problem is limited by  $k < 2\pi/R$ , where  $R$  is the characteristic dimension of the solute. For large solutes, only the long-wavelength part of the polarization structure factors is really needed for the solvation energy calculations. As is shown in Fig. 5, there is a mismatch between the PPSF longitudinal structure factor and MD simulations. However, this difference makes no effect on the calculated solvation energies.

We can summarize our results on parametrizing the solvent properties by stating that the model fluid of dipolar hard spheres can serve as a reliable reference system for calculating polar solvation given the macroscopic properties, the density of dipoles  $\gamma$  and the dielectric constant  $\epsilon_s$ , have been taken from experiment (either laboratory or computer). The theory thus adds an additional parameter  $\gamma$  to the dielectric constant used in electrostatic solvation theories to produce a fully microscopic solvent response. In practical applications of the theory (e.g., in case of solvation in ambient water presented below), the parameter  $\gamma$  needs to be calculated from the molecular properties of the solvent. We use the 1-R Wertheim theory<sup>84</sup> to calculate the effective dipole moment of the solvent (see Ref. 85 for comparison to simulations). The solvent input is thus made by five parameters:  $\{\sigma, \rho, m, \alpha, \epsilon_s\}$ . One needs, in addition, the high-frequency dielectric constant  $\epsilon_\infty$  for the reorganization energy calculations and the temperature slopes of two dielectric constants, as well as the isobaric expansivity of the solvent, for the solvation entropy calculations. The big advantage of the PPSF scheme is that all these parameters have been tabulated for many solvents commonly used in solution chemistry making our method broadly applicable to solvation calculations in polar molecular solvents.

Despite the fact that the dielectric constant is sensitive to local correlations, the polarization structure factors in the long-wavelength limit are fully determined by dipolar correlations general for all polar liquids and not much sensitive to details of the local structure which is, of course, very different in water than in a hard-sphere dipolar fluid. There are several advantages to using dipolar hard spheres as the reference system. First, all thermodynamic and structural properties are controlled by only two parameters, the reduced density  $\rho\sigma^3$  and the dipolar density  $\gamma$ . Second, this system has been well characterized both analytically and numerically. It has served many times as a starting point for developing theories of polar liquids,<sup>60</sup> similarly to the role played by the fluid of hard spheres in theories of nonpolar liquids.<sup>53</sup> Once that stated, we, however, want to stress that the theory itself is based on the structure factors of an arbitrary polar medium with the Gaussian fluctuation spectrum and is not limited to a choice of any particular reference system.

## IV. SIMULATIONS PROTOCOL

AMBER 8.0 (Ref. 86) was used for all MD simulations. The initial configuration of PC was created using a protonated version of the x-ray crystal structure at 1.7 Å resolution (PDB: 1ag6, Ref. 87). This initial configuration of the

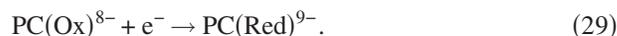
protein was first minimized in vacuum by the conjugate gradient method for 10 000 steps to allow the protein to remove any bad initial contacts. Then the system was solvated in a rectilinear box with several thousand TIP3P molecules,<sup>81</sup> providing at least two to three solvation shells around the protein. To neutralize the charge, a number of sodium ions equal to the total charge of the protein were added. The protein was then relaxed for a few thousand steps while water and sodium were positionally constrained. Finally, the entire system containing solvent, counterions, and protein was energy minimized in 100 000 steps.

Next, the system was heated in a *NVT* ensemble for 30 ps from 0 K to the desired temperature followed by volume expansion in a 1 ns *NPT* run. *NVT* production runs, following density equilibration, lasted from 6 to 18 ns. The last 5–10 ns at the end of each trajectory were used to calculate the averages. The timestep for all MD simulations was 2 fs, and SHAKE was employed to constrain bonds to hydrogen atoms. Constant pressure and temperature simulations employed Berendsen barostat and thermostat, respectively.<sup>88</sup> The long-range electrostatic interactions were handled using a smooth particle mesh Ewald summation with a 9 Å limit in the direct-space sum. The total charge for the protein was  $-9.0$  for the reduced state and  $-8.0$  for the oxidized state.

The MD trajectories were produced in parallel runs at ASU's HPC facility and required 4.0–4.5 years of CPU time for collecting the simulation data followed by 1.8–2.0 years of the analysis. The analysis was also done in parallel on the Opteron cluster using a parallel code developed for this project which directly reads binary AMBER files.

## V. RESULTS

The calculations presented here are focused on two properties: The solvent portion of the redox chemical potential,  $\Delta\mu_s$ , and the solvent reorganization energy  $\lambda_s$ , both corresponding to the half reaction



The former can in principle be calculated as the difference of solvation chemical potentials in the Red and Ox states. However, this approach involves calculating the difference in two large numbers, which is computationally unreliable. Instead, we use the linear response approximation to calculate  $\Delta\mu_s$  according to the equation

$$\Delta\mu_s = \mu_{0s}^{\text{Red}} - \mu_{0s}^{\text{Ox}} = -\Delta\tilde{\mathbf{E}}_0 * \chi * \bar{\mathbf{E}}_0. \quad (30)$$

Here,  $\bar{\mathbf{E}}_0 = (\tilde{\mathbf{E}}_0^{\text{Ox}} + \tilde{\mathbf{E}}_0^{\text{Red}})/2$  and  $\Delta\tilde{\mathbf{E}}_0 = \tilde{\mathbf{E}}_0^{\text{Red}} - \tilde{\mathbf{E}}_0^{\text{Ox}}$  are the mean and the difference of the electric fields in the Red and Ox states. Similarly, the solvent reorganization energy is calculated from

$$\lambda_s = \frac{1}{2}\Delta\tilde{\mathbf{E}}_0 * \chi * \Delta\tilde{\mathbf{E}}_0. \quad (31)$$

Equation (31) applies to the reorganization energy of nonpolarizable solvents employed in computer simulations. For laboratory data, nuclear polarization should be separated from the overall solvent polarization and the response

function  $\chi$  is replaced by the nuclear response function  $\chi_n$  as explained above and in more detail in Ref. 15.

### A. Redox thermodynamics

The solvation thermodynamics calculated here can be related to experimental redox entropies reported by measuring the temperature dependence of the standard or midpoint electrode potentials.<sup>24,25,35</sup> An electrochemical experiment corresponds to bringing a solution containing given numbers of oxidized and reduced reagents, which are not necessarily in equilibrium (the ratio of their numbers is not a Boltzmann factor), in contact with a metal electrode. The equilibrium is established between the electronic subsystem of the redox pair and the electrode in such a way that the electrode is charged and its electrochemical potential  $\mu$  is shifted from the vacuum Fermi energy  $\epsilon_F$  by the electrostatic potential  $\phi$ :  $\mu = \epsilon_F - e\phi$ .

The numbers of the oxidized and reduced forms of the redox pair,  $N_{\text{Ox}}$  and  $N_{\text{Red}}$ , are assumed to be large enough so that they are not affected by charging the electrode. The electrochemical potential of the electrode then becomes equal to the absolute electrochemical potential of the redox couple in the solution.<sup>89</sup> The latter can be found from simple statistical arguments. The grand-canonical free energy of two fermionic subsystems of  $N_{\text{Ox}}$  and  $N_{\text{Red}}$  electronic levels is<sup>90</sup>

$$\beta\Omega = -N_{\text{Ox}} \ln(1 + e^{\beta(\mu - \epsilon_{\text{Ox}})}) - N_{\text{Red}} \ln(1 + e^{\beta(\mu - \epsilon_{\text{Red}})}), \quad (32)$$

where  $\epsilon_{\text{Ox}}$  and  $\epsilon_{\text{Red}}$  are the average energies of the electronic levels in the corresponding redox states. The chemical potential is then found by requiring that the derivative  $-(\partial\Omega/\partial\mu)_T$  is equal to the total number of electrons  $N_{\text{Red}}$ . For the energy gap between Ox and Red states greater than  $k_B T$ , this requirement results in the Nernst equation,<sup>91</sup>

$$\mu = \frac{\epsilon_{\text{Ox}} + \epsilon_{\text{Red}}}{2} - k_B T \ln(N_{\text{Ox}}/N_{\text{Red}}), \quad (33)$$

in which the standard potential is given by the mean of the average electronic energies

$$\phi^0 = -\frac{\epsilon_{\text{Ox}} + \epsilon_{\text{Red}}}{2e}. \quad (34)$$

The same result follows from the use of the stationary condition (zero electrode current) for the rates of reduction and oxidation<sup>92</sup>

$$k_{\text{Ox}} c_{\text{Ox}} = k_{\text{Red}} c_{\text{Red}}, \quad (35)$$

where  $c_{\text{Ox/Red}}$  are the surface concentrations. By using the Marcus equation for the reaction rate,<sup>93</sup>

$$k_{\text{Ox/Red}} \propto \exp\left[-\beta \frac{(\epsilon_{\text{Ox/Red}} - \mu)^2}{4\lambda_s}\right], \quad (36)$$

and neglecting the logarithmic correction including the ratio of two surface concentrations, one gets equal rates at  $\mu \approx (\epsilon_{\text{Ox}} + \epsilon_{\text{Red}})/2$ . The double-well Marcus free energy surface for the electrode electron transfer is then symmetrical as illustrated in Fig. 6. This picture bears a clear similarity with

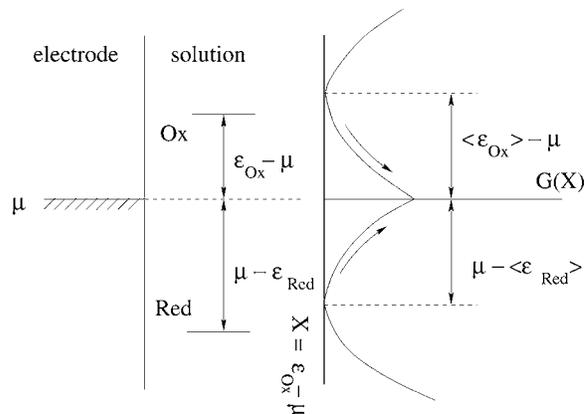


FIG. 6. Contact of a redox pair with the metal electrode.  $\epsilon_{\text{Ox}} - \mu$  and  $\epsilon_{\text{Red}} - \mu$  show the fluctuating energy gaps for reduction and oxidation electron transfer, respectively. The equilibrium electrochemical potential of the electrode is established when the equilibrium energy gaps are equal for the reduction and oxidation reactions [Eq. (33)]. The Marcus electron transfer parabolas, shown by the dependence of free energy  $F(X)$  on the energy gap coordinate  $X = \epsilon_{\text{Ox}} - \mu$ , are symmetric in this case producing equal oxidation and reduction currents [Eq. (35)].

the formation of the Fermi level in the forbidden band of a semiconductor, as was noticed by Reiss.<sup>94</sup>

The electronic energies are given by the sums of their vacuum components,  $\epsilon_{\text{Ox/Red}}^0$ , and the interaction of the electric field of the electron  $\mathbf{E}_e$  with the polarization of the solvent in equilibrium with the total electric field of the molecule in the solution,<sup>95</sup>

$$\epsilon_{\text{Ox/Red}} = \epsilon_{\text{Ox/Red}}^0 - \mathbf{E}_e * \chi * \mathbf{E}_0^{\text{Ox/Red}}. \quad (37)$$

Taking into account that

$$\mathbf{E}_e = \mathbf{E}_0^{\text{Red}} - \mathbf{E}_0^{\text{Ox}} = \Delta\mathbf{E}_0, \quad (38)$$

one gets from Eqs. (34), (37), and (38) the commonly used connection between the standard electrode potential and the solvation part of the redox free energy,

$$\phi^0 = -\frac{\epsilon_{\text{Red}}^0 + \epsilon_{\text{Ox}}^0}{2e} - \frac{\Delta\mu_s}{e}, \quad (39)$$

where  $\Delta\mu_s$  is given by Eq. (30). The first term in this equation disappears in the temperature derivative reported experimentally,<sup>24-26</sup>

$$e \left( \frac{\partial\phi^0}{\partial T} \right)_P = \Delta s_s = s_s^{\text{Red}} - s_s^{\text{Ox}} = - \left( \frac{\partial\Delta\mu_s}{\partial T} \right)_P. \quad (40)$$

We need to stress here that redox entropies in polar solutions are sensitive to the presence of electrolyte.<sup>22,96</sup> One therefore can expect only a qualitative agreement between experiments done in buffered protein solutions<sup>25,26</sup> and our calculations at zero ionic strength.

From Eqs. (34), (37), and (38), one can directly derive the equation for the solvation redox free energy,

$$\Delta\mu_s = (\langle\Delta V_{0s}\rangle_{\text{Ox}} + \langle\Delta V_{0s}\rangle_{\text{Red}})/2, \quad (41)$$

where  $\Delta V_{0s}$  is the difference in the solute-solvent interaction energies in the Red and Ox states and the averages are taken over the corresponding ensembles. The same average vertical gaps can be used to calculate the reorganization energy as

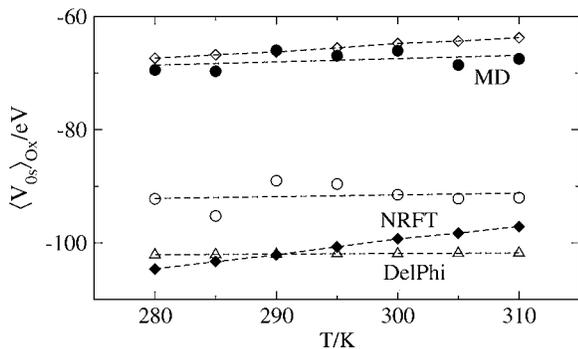


FIG. 7. Average solute-solvent interaction energy  $\langle V_{0s} \rangle_{Ox}$  obtained from MD simulations (closed circles), NRFT (diamonds), and DELPHI continuum calculations (vdW cavity, triangles). The closed diamonds refer to the total average energy including the polarization and density components, while the open diamonds denote the polarization component only. The dashed lines represent linear regressions through the points.

$$\lambda_s = (\langle \Delta V_{0s} \rangle_{Ox} - \langle \Delta V_{0s} \rangle_{Red})/2. \quad (42)$$

We need to caution here that, while Eq. (34) is a statistical-mechanical result, Eqs. (37) and (42) are based on the LRA for the solute-solvent interaction energy and might be affected by deviations from this approximation.

## B. Solute-solvent average energy

In addition to our NRFT formalism, we have used the dielectric continuum approximation implemented in the DELPHI program suite<sup>7</sup> in the solvation calculations. Dielectric constant of ambient water was used for the solvent continuum and  $\epsilon_s=1$  for the protein. This latter choice was driven by our desire to compare continuum and microscopic calculations of solvation thermodynamics since the latter does not assume any polarization of the protein. Three different charging schemes have been used and compared to MD simulations (Table I and Sec. IV). In the following, we will discuss the results relevant to charging scheme II only, which are also visualized in Fig. 7. Table II lists the results of NRFT and DELPHI calculations of the average energy  $\langle V_{0s} \rangle_{Ox}$  of PC in the Ox state, where  $V_{0s}$  refers to the interaction of the protein with the water molecules in the simulation box. Sodium ions used to neutralize the simulation box shift the interaction potential down by about  $-25$  eV as is shown by open circles in Fig. 7. It is not clear how realistic this number might be since Debye–Hückel screening is not accounted for by the simulations. In addition, we found that the variance of the Coulomb interactions with the counterions is negligible thus resulting in a very small contribution to the reorganization energy, which is known to be the case from analytical models which do include the Debye–Hückel screening.<sup>97</sup>

The NRFT calculations listed in Table II and shown in Fig. 7 have been done by using Eqs. (3) and (9) in which the electric field of PC in Ox state was used for  $\tilde{E}_0(\mathbf{k})$ . The close diamonds in Fig. 7 refer to the total solvent response, while open diamonds represent the polarization response only [ $\chi_p$  in Eq. (9)]. Two interesting observations result from examining Fig. 7: (i) A close proximity of the NRFT result to the standard (vdW) continuum calculation and (ii) a good agreement between the polarization portion of the NRFT cal-

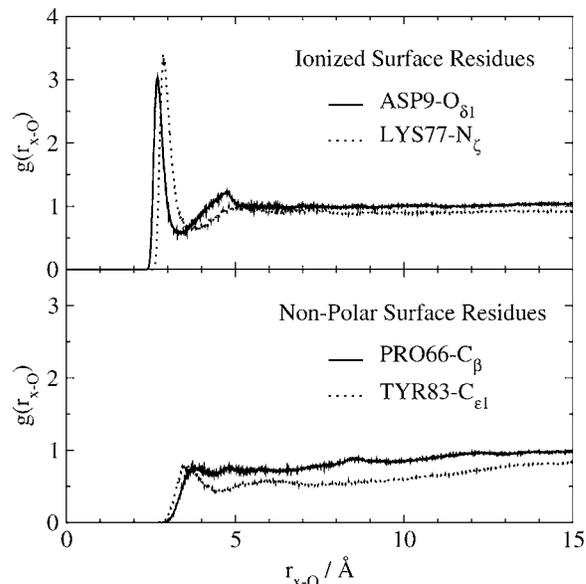


FIG. 8. Radial distribution functions between surface residues of PC and oxygens of water. The upper panel shows ionized residues and the lower panel refers to nonpolar residues. The legends in the figure list: aspartic acid (ASP), the probe atom is the oxygen at the first  $\delta$  position; lysine (LYS), the probe atom is nitrogen at the  $\zeta$  position; proline (PRO), the probe atom is the  $\beta$  carbon; tyrosine (TYR), with the first  $\epsilon$  carbon as the probe atom.

culations and MD results. The continuum electrostatics does not reproduce the slope of the average energy as we also discuss below in relation to the redox entropy.

In order to understand the origin of the close agreement between MD and the polarization component of the solvent response, one needs to recall what comes to the calculation of the polarization and density components of the solvation free energy. The polarization response is calculated by assuming that the only influence of the solute on the polarization field is to exclude it from the solute volume represented by the step function  $\theta_0(\mathbf{r})$  equal to one inside the solute and zero otherwise [Eq. (12)]. The density component corrects this result by taking into account the inhomogeneous density profile formed at the surface of a hard-wall solute. In dense liquids, such a profile is characterized by a sharp peak of the radial distribution function in the first solvation shell of the solute. Correspondingly, reflecting the belief that the short-range structure of liquids is primarily determined by repulsions,<sup>53</sup> the density structure factor  $S(k)$  in Eq. (21) was taken in our calculations from the Percus–Yevick solution for hard spheres.<sup>60</sup> The close proximity of the full NRFT calculation to the standard DELPHI/vdW algorithm (Fig. 7) illustrates the fact that the common parametrization of the atomic radii is based on the experience learned for hydration of small ions with tightly bound first solvation shell. In the present algorithm, this physics is accommodated by the density component of the solvation free energy.

The structure of water at the protein surface is quite different from what is normally obtained by inserting a small solute in a molecular solvent. The structure is inhomogeneous, including islands of highly structured water around polar and ionized residues and a much softer density profile at the hydrophobic patches. This reality is illustrated in Fig. 8 which shows pair distribution functions between ionized and

TABLE III. Redox thermodynamics of PC (eV).

T (K)	$\Delta\mu_s(\text{TIP3P})$				$\lambda_s(\text{TIP3P})$				$\lambda_s(\text{H}_2\text{O})$		
	MD	NRFT	vdW <sup>a</sup>	SAS <sup>b</sup>	MD	NRFT	vdW <sup>c</sup>	SAS	NRFT <sup>d</sup>	vdW	SAS
280	-3.05	-6.21	-9.63	-7.15	0.845	1.37(0.59)	3.645	0.690	0.82	2.728	0.417
285	-2.73	-6.15	-9.62	-7.14	0.814	1.35(0.58)	3.641	0.685	0.80	2.725	0.417
290	-2.71	-6.10	-9.62	-7.14	0.646	1.33(0.57)	3.636	0.685	0.79	2.722	0.417
295	-3.14	-6.04	-9.61	-7.14	0.565	1.32(0.56)	3.630	0.685	0.78	2.718	0.417
300	-2.82	-6.01	-9.60	-7.14	0.435	1.30(0.55)	3.623	0.684	0.77	2.714	0.417
305	-2.94	-5.94	-9.59	-7.13	0.426	1.29(0.54)	3.620	0.684	0.76	2.710	0.417
310	-2.53	-5.89	-9.59	-7.13	0.543	1.27(0.53)	3.618	0.684	0.74	2.710	0.416

<sup>a</sup>Poisson–Boltzmann calculations with vdW radii assigned to the protein atoms (standard vdW cavity).

<sup>b</sup>Poisson–Boltzmann calculation with solvent radius added to the radii of the protein atoms exposed to the solvent (solvent-accessible cavity, SAS).

<sup>c</sup>The continuum calculations are done with the dielectric constant of TIP3P water  $\epsilon_s=95$  and  $\epsilon_\infty=1.0$ . The dielectric constant of the protein interior was put equal to unity in order to be consistent with the microscopic calculations. The temperature variation of the dielectric constant of  $d\epsilon_s/dT=-0.654\text{ K}^{-1}$  (Ref. 15) was adopted for the entropy calculations listed in Table IV.

<sup>d</sup>Calculations in ambient water using  $\epsilon_\infty=1.78$ ,  $\epsilon_s=78$ ,  $d\epsilon_s/dT=-0.398\text{ K}^{-1}$ , and  $d\epsilon_\infty/dT=-2.75\times 10^{-4}\text{ K}^{-1}$ . In addition, temperature expansion was included with the constant pressure expansivity coefficient  $\alpha_p=2.6\times 10^{-4}\text{ K}^{-1}$ .

non-polar residues and water's oxygens. While the distribution functions of ionized residues are reminiscent of the structures typically observed around small solutes in dense solvents, the water structure around nonpolar residues is quite different: There is no first shell peak and water interface is shifted by  $\approx 1\text{ \AA}$ , in accord with simulations of nanoscale hydrophobic solutes.<sup>98</sup>

The stronger attraction of the surface water molecules to the bulk than to a nonpolar hydrophobic patch of the protein (cavity expulsion potential<sup>99,100</sup>) results in a weak dewetting of the surface<sup>19</sup> with the density at the interface lower than in the bulk (Fig. 8). Since there are only a few charged residues on the protein surface (see the Discussion section), the average surface structure is closer to a stepwise cutoff introduced in the polarization component of the response function than to a structured liquid at the surface of a small polar/ionic solute.<sup>101</sup> This observation explains a good agreement between MD and polarization calculations of the solvation thermodynamics in this paper as well as an equally impressive agreement with the simulations obtained in our previous calculations of charge transfer across a polypeptide bridge.<sup>15</sup>

### C. Solvent Gibbs and reorganization free energies

The results of calculations of the redox solvation energy and solvent reorganization energy [Eqs. (41) and (42)] using

different levels of the theory are listed in Table III. Redox and reorganization entropies are given in Table IV. In addition, the temperature dependence of  $\Delta\mu_s$  and  $\lambda_s$  are visualized in Figs. 9 and 10. The differences in the theoretical results arise from the different level of structural solvent information incorporated in each formalism. For completeness, we have also listed in Table III the NRFT calculations with the solvent parameters of ambient water.

The continuum electrostatics gain access to the solvation entropy through the temperature variation of the solvent dielectric constant. Therefore, since the dielectric constant commonly decreases with heating, the solvent becomes effectively less polar and the solvation free energy of a charge distribution increases, i.e., becomes less negative. If one considers redox species positively charged in both redox states, the oxidized state carries a larger charge and, hence, the difference of Red and Ox solvation energies has a positive value decreasing with increasing temperature. The redox entropy in Eq. (40) is then positive as is typically observed for simple inorganic ions.<sup>35,102</sup> By the same arguments, the species carrying negative charge in both redox states should have a negative redox entropy, which is the case for the negatively charged PC in our calculations.

Despite the right sign of the redox entropy, the magnitudes of both the Gibbs solvation energy and the entropy are

TABLE IV. Redox solvation free energy  $\Delta\mu_s$  and redox entropy  $\Delta s_s$  in for the Red/Ox states of PC. Also listed are the reorganization energy and reorganization entropy,  $s_\lambda=-\partial\lambda/\partial T$ . All energies are in eV and entropies are in meV/K,  $T=300\text{ K}$ .

Method	$\Delta\mu_s$	$\Delta s_s$	$\lambda_s$	$s_\lambda$
DELPHI with vdW cavity	-9.92	-1.25	3.62	0.97
DELPHI with solvent-accessible cavity	-7.12	-0.45	0.64	0.005
Nonlocal polarization response functions <sup>a</sup>	-6.01(-1.33)	-10.5(4.8)	1.30(0.55)	3.2(2.0)
Molecular dynamics <sup>b</sup>	-2.81	-7.40	0.54	10.7
Experiment		-0.4 <sup>c</sup>		
		-1.4 <sup>d</sup>		

<sup>a</sup>The numbers in the parentheses indicate contributions to the solvation free energy and entropy from density fluctuations.

<sup>b</sup>MD results refer to linear fits through the simulation points.

<sup>c</sup>Reference 25.

<sup>d</sup>Reference 24.

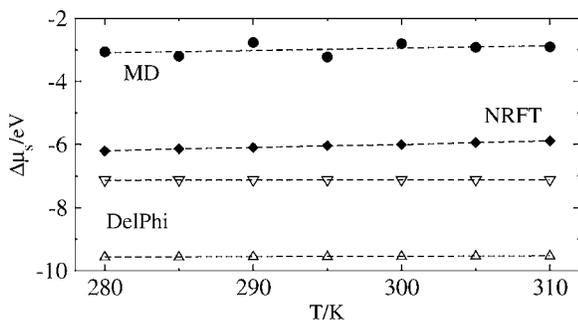


FIG. 9. Solvation Gibbs energy from MD simulations (closed circles), NRFT calculations (closed diamonds), continuum DELPHI calculation with the standard cavity definition (up triangles), and the cavity surface augmented by the solvent radius  $\sigma/2$  (down triangles). The dashed lines are linear regressions through the points.

markedly different in continuum and microscopic/simulation approaches:  $|\Delta\mu_s|$  is higher in the standard implementation of DELPHI (vdW cavity) than the NRFT value by a factor of 1.5 while the redox entropy is lower by a factor of ten. The use of the solvent-accessible cavity brings the value of  $\Delta\mu_s$  in a closer proximity to the NRFT, but the redox entropy is lowered even more (Table IV). The magnitude of  $\Delta\mu_s$  from NRFT is significantly higher than from MD even if the density component is subtracted from the total response. This initially comes a bit of surprise given a good agreement between the polarization-NRFT and MD values of  $\langle V_{0s} \rangle_{\text{Ox}}$  reported in Table II and Fig. 7. We do not currently have a good explanation of this disagreement (see Discussion below).

Concerning the reorganization energy calculations, the standard DELPHI/vdW algorithm gives  $\lambda_s$  three times larger than NRFT and almost an order of magnitude larger than the MD simulations (Table IV). The origin of large  $\lambda_s$  in the standard (vdW) implementation of the continuum model is in placing highly polar dielectric into the small pocket near copper which water molecules do not visit in MD simulations. Figure 11 shows the pair distribution function between the Cu ion and oxygen of water testifying to the fact that water never comes to Cu closer than 6 Å and the maximum of the first solvation shell appears at 6.7 Å. The addition of the solvent radius to the cavity corrects this error reducing the reorganization energy to the level of 0.64 eV consistent

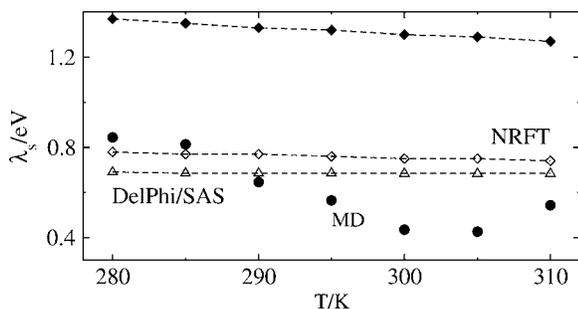


FIG. 10. Reorganization energy of PC vs temperature calculated from MD simulations (closed circles), from NRFT (diamonds), and from dielectric continuum using solvent-accessible cavity definition (triangles). The dashed lines are linear regressions through the points. The filled diamonds refer to the full NRFT calculation and the open diamonds denote the polarization response only.

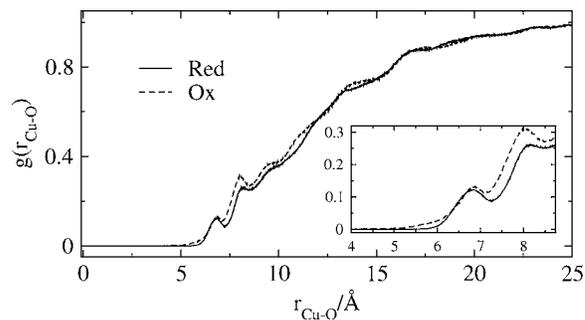


FIG. 11. Pair distribution function between oxygen of water and Cu of PC in the reduced (Red) and oxidized (Ox) states.

with the polarization component of the NRFT (0.75 eV, Table IV).

The combination of the absolute values of the reorganization energy with the reorganization entropies clearly indicates that re-scaling of the dielectric cavity, often employed in various continuum formulations, does not solve the solvation problem. Entropies calculated by including the effect of density fluctuations on the solvent response are generally in better agreement with MD simulations than the results obtained from polarization response only. All solvation free energies obtained from such calculations, however, significantly exceed the simulation results. The problem lies in the assumed dense structure of the liquid around the solute, which is obviously not realized for a hydrated protein. Capturing density fluctuations is essential for the correct calculations of the entropies, but the effective density should be adjusted to that of the structurally loose hydration layer at the protein surface.

One might argue that the failure to reproduce the results of the simulations by the NRFT calculations can be traced back to the deficiencies of the model and not to the specific structure created by the protein in water. We believe that this objection cannot explain the differences between our calculations and MD simulations. In order to illustrate this point, we show in Fig. 12 the comparison of the NRFT calculations with our recent MD results for solvation of a small charge-transfer molecule *p*-nitroaniline in SPC/E water.<sup>16</sup> The partial charges for this molecule, used in the calculations were

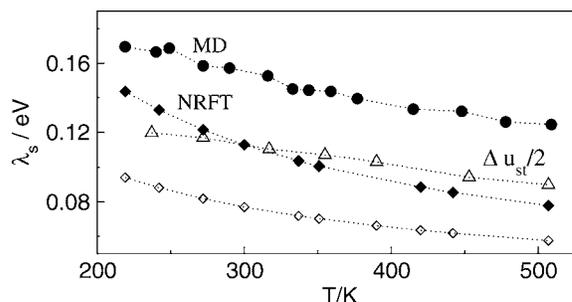


FIG. 12. Reorganization energy of charge-transfer transition in *p*-nitroaniline dissolved in SPC/E water. The results are obtained by MD simulations (Ref. 16) (closed circles) and NRFT calculations (diamonds). Closed and open diamonds refer to the full and polarization response, respectively. Triangles denote half of the Stokes shift from MD simulations; in the LRA,  $\Delta u_{st}/2 = \lambda_s$ . The dielectric constants of SPC/E water at different temperatures, required for the NRFT input, were taken from MD simulations (Ref. 16).

tabulated in Ref. 103. The comparison with the NRFT method is somewhat complicated in this case by nonlinear solvation effects related to specific binding of water to the oxygen sites and seen in the deviation of the half of the Stokes shift from the reorganization energy. Nevertheless, the NRFT calculation with the density component included is in a reasonable agreement with the simulation data for both the reorganization energy and entropy, while the polarization response alone clearly underestimates the reorganization energy.

## VI. DISCUSSION

This paper has addressed the thermodynamics of electrostatic interactions of the protein with water. The free energy of electrostatic protein/water coupling is a part of the overall free energy of protein solvation. This component is experimentally probed by changing the charge of the redox site in redox reactions. What has been left out from our analysis is the electrostatic interaction of the active site with the protein charges. This component is probably less significant for the redox entropy (see below), but needs to be included into a complete analysis of the redox thermodynamics. The extension of the present algorithm to this case is currently delayed by the absence of reliable models of microscopic polarization correlation functions of the protein. Once these are available, the formalism can be extended to include protein electrostatics. Our hope for this development is based on the observation, from our current simulations, that the cross correlation of interaction potentials of the redox site with water and protein is small (approximately 10%) (Ref. 104) in the overall cumulant  $\langle(\delta V_{0s})^2\rangle$  allowing one to treat, in the first approximation, the protein and water cumulants separately.

The formalism developed here is based on the LRA suggesting that knowing the variance of electrostatic fluctuations around a neutral repulsive solute is sufficient to calculate the solvation chemical potential [Eq. (2)]. One can arrive at Eq. (2) from the following simple considerations. The chemical potential of solvation can be written as the integral over the magnitude of the solute-solvent electrostatic interaction  $\epsilon = V_{0s}$  as follows:

$$e^{-\beta\mu_{0s}} = \int d\epsilon P(\epsilon, \beta) e^{-\beta\epsilon}. \quad (43)$$

Here, the probability density  $P(\epsilon, \beta)$  of reaching the value  $\epsilon$  is obtained by taking the statistical average over the reference Hamiltonian  $H_0$  which includes all the interaction potentials in the system except the solute-solvent electrostatic potential  $V_{0s}$ ,

$$P(\epsilon, \beta) = Q_0^{-1} \int \delta(\epsilon - V_{0s}) e^{-\beta H_0} d\Gamma, \quad (44)$$

where  $Q_0 = \int \exp[-\beta H_0] d\Gamma$  and  $\Gamma$  denotes the phase space volume.<sup>90</sup> Since the solute-solvent component of  $H_0$  includes mostly isotropic short-range interactions, one can put  $\langle V_{0s} \rangle_0 = 0$ . In addition, the Gaussian approximation can be applied to  $P(\epsilon, \beta)$ ,

$$P(\epsilon, \beta) \propto \exp\left[-\frac{\epsilon^2}{2\sigma_\epsilon^2}\right], \quad (45)$$

with

$$\sigma_\epsilon^2 = \langle(\delta V_{0s})^2\rangle_0. \quad (46)$$

Combining Eqs. (43) and (45), one immediately arrives at Eq. (2).

One of the important lessons of Eqs. (43) and (44) is that all the thermodynamic information required for the solvation problem is contained in the distribution of electrostatic potential fluctuations around a “nonpolar” solute in which all the electrostatic interactions with the solvent have been switched off. For large solutes, the spectrum of electrostatic fluctuations around the repulsive solute core will ultimately determine the thermodynamics of solvation. The Gaussianity of electrostatic fluctuations created by polar solvents inside nonpolar solutes [Eq. (45)] still needs testing, in particular at large scale, since it was questioned for small solutes dissolved in water.<sup>105</sup> However, it is important to stress that Eqs. (43) and (44) are exact and not limited to linear solvation.

The reasoning outlined above works well for small- and medium-size solutes, as is exemplified in Fig. 12. However, the application of the same procedure to protein solvation studied in this paper has encountered some difficulties suggesting that methods developed over several decades and successfully applied to solvation of small solutes in dense polar solvents are probably not directly transferable without significant modifications to mesoscopic hydration of proteins. The application of our algorithm to the calculation of the average solute-solvent interaction energy turned out to be quite successful when the density profile of water around the protein is approximated by a step function (Fig. 7). The calculations agree with MD within simulation uncertainties ( $\leq 5\%$ ). As mentioned above, this comes as a result of averaging between tight and loose water structures at the protein surface. The application of the same procedure to solvation of the difference charges of the active site (reorganization energy, Fig. 10) was less successful, but the agreement is probably still acceptable, in particular, at lowest temperatures.

Where the calculations and simulations come in significant disagreement is for the free energy of solvation obtained in simulations as the mean of two vertical transition energies [Eq. (41)]. Some recent simulations of cavities in force-field water models<sup>106,107</sup> and of uncharged protein<sup>108</sup> have suggested a possible origin of this effect. It was found that water structured at the protein surface creates a positive potential within an uncharged cavity/protein. In terms of Eq. (45), this implies a constant shift of the solute-solvent energy  $\epsilon \rightarrow \epsilon - \epsilon_0$ . Since we have assumed random orientations of water around an uncharged solute, as is indeed the case for small solutes,<sup>16</sup> our calculations do not include the effect of a positive background potential and include only changes of the potential in response to protein’s charges. In the case of a constant background potential  $\Phi$ , Eq. (41) modifies to

$$\Delta\mu_s = (\langle\Delta V_{0s}\rangle_{\text{Ox}} + \langle\Delta V_{0s}\rangle_{\text{Red}})/2 - \Delta q\Phi, \quad (47)$$

where  $\Delta q = q_{\text{Red}} - q_{\text{Ox}} = -1$ . Ashbaugh<sup>106</sup> reported an average positive potential of about  $e\Phi \approx 9$  kcal/mol for cavities in SPC water comparable in size with PC. In terms of our calculations, it amounts a positive shift of the simulation data by about 0.4 eV which will increase the current gap of about 3.2 eV (300 K) between the MD and NRFT.

The origin of the difference in solvation free energies between calculations and MD simulations might be related to the weakly dewetted water density profile near the active site. This would imply that some of the properties of water structure around large cavities expelled by the protein from its volume are quite different from the common experience gained with small solutes. This qualitative difference between small-size and large-size solvation has recently gained appreciation for hydrophobic solvation<sup>101</sup> as we discuss next.

The Gaussian model of hydrophobic solvation goes back to the Pratt–Chandler theory of hydrophobicity<sup>109</sup> recently extended by Pratt and co-workers.<sup>110,111</sup> The formulation of the theory of hydrophobic solvation follows a path similar to the one outlined in Eqs. (43)–(46) asking what is the free energy  $\mu_\Omega$  needed to solvate a solute of volume  $\Omega_0$ . It is given by the Gaussian probability,<sup>101,110</sup>

$$\beta\mu_\Omega \approx \frac{\rho^2\Omega_0^2}{2\chi_\Omega}, \quad (48)$$

with the fluctuation of the number of solvent particles in volume  $\Omega_0$ ,

$$\chi_\Omega = \langle(\Delta N)^2\rangle_\Omega = \rho\Omega_0 + \rho^2 \int_{\Omega_0} d\mathbf{r}d\mathbf{r}' h_{ss}(|\mathbf{r} - \mathbf{r}'|). \quad (49)$$

In Eq. (49),  $h_{ss}(r)$  is the pair correlation function of the homogeneous solvent.

The Gaussian probability of electron transfer carries a close similarity with Eq. (48) giving the activation free energy as

$$\beta\mu^{\text{act}} = \frac{X_0^2}{2\sigma_\epsilon^2}, \quad (50)$$

where  $X_0$  is the average vertical donor-acceptor energy gap and  $\sigma_\epsilon^2$  is the variance of the solute-solvent interaction potential when  $V_{0s} = \Delta V_{0s}$  is used in Eqs. (44) and (45). Not surprisingly, the structure of the equation for  $\sigma_\epsilon^2$  resembles Eq. (49):

$$\begin{aligned} \sigma_\epsilon^2 = & \rho \int V_{0s}^2(1) d\Gamma_1 + \rho^2 \int V_{0s}(1)V_{0s}(2)h_{ss}(1,2) d\Gamma_1 d\Gamma_2 \\ & + \rho^2 \int V_{0s}^2(1)h_{ss}(1,2)\theta_0(2) d\Gamma_1 d\Gamma_2. \end{aligned} \quad (51)$$

The last summand in this equation represents the density component of the response since it transforms into the  $\mathbf{k}$ -space integral with the density structure factor of the liquid [Eq. (21)].

Equation (51) carries a close resemblance with the Pratt–Chandler theory of hydrophobicity.<sup>109</sup> The sum of the

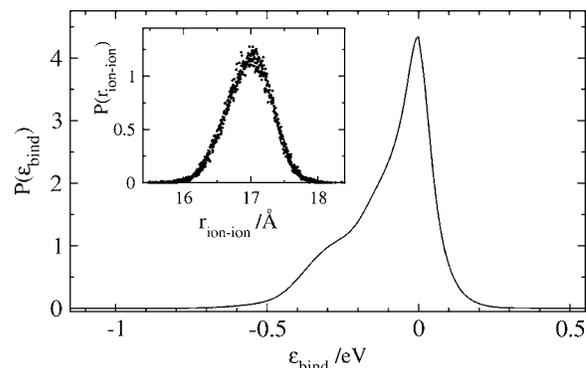


FIG. 13. The distribution of binding energies of water molecules in the first solvation shell of PC(Ox) defined by adding the water diameter 2.87 Å to all protein atoms exposed to solution. The binding energy is defined as the total interaction energy of a given water molecule with all atoms of the protein. The inset shows the distribution of distances of charged residues at the protein surface.

first and the third terms is the average of the squared solute-solvent potential  $V_{0s}^2(1)$  over the solute-solvent density profile,<sup>112</sup>

$$\rho(\mathbf{r}) = \rho(1 + c_{0s}(\mathbf{r})) + \rho^2 \int c_{0s}(\mathbf{r}')h_{ss}(|\mathbf{r} - \mathbf{r}'|)d\mathbf{r}', \quad (52)$$

in which the solute-solvent direct correlation function  $c_{0s}(\mathbf{r})$  is replaced by its lowest density expansion,<sup>60</sup>  $c_{0s}(\mathbf{r}) \approx -\theta_0(\mathbf{r})$  ( $\theta_0(\mathbf{r})$  is one inside the solute and zero otherwise). The observation that the density profile around solutes of size  $\geq 1$  nm can be approximated by a step function (Fig. 11, also see Fig. 3 in Ref. 101) amounts to neglecting the second term in Eq. (52) and, correspondingly, the density components  $\chi_d$  in the solvent response function [Eq. (9)].

The fact that the spectrum of electrostatic potential fluctuations around a nonpolar solute gives complete information about polar solvation brings this latter problem in close relation to the problem of hydrophobic solvation. It was realized in recent years that hydrophobic solvation of small and large solutes are qualitatively different.<sup>19,101,112</sup> Solvation character changes at the critical size of  $\approx 1$  nm from entropy-dominated solvation of small solutes (Gaussian statistics<sup>110</sup>) to enthalpy-dominated solvation of large solutes driven by the creation of the solute-solvent interface (non-Gaussian statistics<sup>113</sup>). The solvent interface around large solutes involves partial dewetting, strongly sensitive to the strength of solute-solvent attractions<sup>17</sup> and the appearance of large-size interfacial density fluctuations.<sup>101</sup> In case of protein solvation, surface water creates a nonzero potential around uncharged proteins discussed above, while density fluctuations lead to complex protein-solvent dynamics<sup>18</sup> and are probably connected to “slaving” of the protein dynamics by the solvent.<sup>114</sup>

The inhomogeneous nature of the protein/water interface is illustrated in Fig. 13. The inset in the figure shows the distribution of distances between charged residues at the protein surface with the average of about 17 Å. This distance is sufficient for a buildup of partially dewetted water structures at the hydrophobic surface between the charges, with the critical distance estimated to be about 1 nm.<sup>101</sup> Figure 13

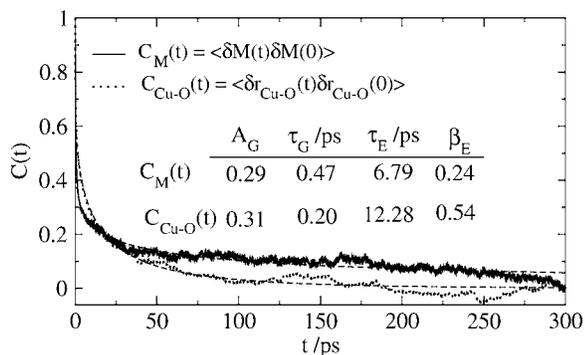


FIG. 14. Normalized time self-correlation functions for Cu-water minimum distance,  $C_{Cu-O}(t)$  (dotted line), and for the number of particles  $M(t)$  in the first solvation shell,  $C_M(t)$  (solid line). The dashed lines refer to the fits of correlation functions obtained from MD simulations to Eq. (53). The parameters of the fits are listed in the plot.

also shows the distribution of binding energies of water molecules residing in the first solvation shell of the protein. The distribution is very wide ( $\approx 1.6$  eV) and clearly bimodal with the lower-energy shoulder developing at the average energy of binding to charged residues. The main portion of  $\langle M \rangle \approx 505$  water molecules residing in the first solvation shell represents weakly bound molecules at hydrophobic patches. The distribution tail stretching into the positive side of the energy scale also indicates that some of these surface waters find their configuration energetically unfavorable. The average binding energy at the interface is low,  $-0.11$  eV. When multiplied with the average number of first shell molecules, it gives the solvation energy of  $-55.6$  eV, which accounts for 81% of the total solvation energy of PC(Ox) listed in Table II.

Dynamical information extracted from MD trajectories also supports the picture of heterogeneous, roughly bimodal, statistics of water motions at the protein/water interface. Figure 14 shows the normalized time self-correlation functions of minimum distance between PC's copper and the water oxygen,  $C_{Cu-O}(t)$ . It is compared to the self-correlation function of the number of molecules  $M(t)$  in the first solvation shell,  $C_M(t)$ . Both correlation functions, when normalized, are fitted well by the combination of a fast initial Gaussian decay followed by the slow stretched-exponential tail,

$$C(t) = A_G e^{-(t/\tau_G)^2} + (1 - A_G) e^{-(t/\tau_E)^{\beta_E}}, \quad (53)$$

where  $\tau_G$  and  $\tau_E$  are corresponding relaxation times and  $\beta_E$  is the stretching exponent. As is seen from the fitting parameters listed in Fig. 14, based on the exponential relaxation time, the first shell waters generally relax about twice as fast as those closest to the Cu ion.

The inhomogeneous structure of the protein/water interface certainly affects the thermodynamics of polar solvation. It is currently not clear how the crossover from the Gaussian regime of small solutes to the interface-dominated regime of large solutes will translate to the problem of electrostatic solvation. The calculations and simulations performed in this study gave some insights into the kind of problems which the theory needs to address. It is clear that the inhomogeneous nature of the protein surface requires algorithms for the local

density profile<sup>61</sup> or the solvent-accessible surface<sup>115</sup> to be a part of a quantitative solvation theory. The current formulation can be improved by using a local density approximation as, for instance, applied by Ramirez and Borgis.<sup>44</sup> In this approach, the bulk dipolar density  $y$  [Eq. (6)] is replaced by the local dipolar density  $y(\mathbf{r}) = (4\pi/9)\beta m'^2 \rho(\mathbf{r}) + (4\pi/3)\alpha \rho(\mathbf{r})$  with the local density profile calculated from, for instance, Lum–Chandler–Weeks theory.<sup>112</sup> Our computational algorithm will then require the Fourier transform of the field  $\mathbf{E}_0(\mathbf{r})\rho(\mathbf{r})/\rho$  instead of  $\mathbf{E}_0(\mathbf{r})(1-\theta_0(\mathbf{r}))$  in the current implementation. In addition, density fluctuations of the interfacial region need to be addressed since the mean position of the interface has no physical significance in the presence of large-amplitude fluctuations.<sup>101</sup> What is also clear is that the density fluctuations of the interfacial region present a nuclear mode, largely diminished for solvation of small solutes, which grows in significance for solvation of biopolymers. Future theoretical development needs to address this physical reality.

Both NRFT and MD simulations predict a substantial temperature dependence of the redox potential (approximately  $\approx 5-7$  meV/K). This magnitude of the temperature variation is prohibitively high since a temperature change of  $\approx 15$  K would shift the redox potential by about 100 mV potentially terminating many enzymatic redox reactions. The theoretically predicted redox entropies refer to zero ionic strength and the solvent contribution to the redox potential. The contribution from the protein to the overall redox potential is available from our simulations, but it depends weakly on temperature contributing only  $\approx -0.9$  meV/K to the redox entropy. The combined solvent/protein entropy is then significantly higher than redox entropies experimentally observed in buffered solutions<sup>24,25</sup> (Table IV). This difference raises the question of the role of the ionic atmosphere in stabilizing the redox potential of metalloproteins. The existing experimental evidence for small redox molecules<sup>116,117</sup> and simulations of metalloproteins<sup>22,118</sup> all indicate a substantial effect of the ionic atmosphere on the redox entropy, which might compensate for the large entropy due to water and protein. In order to get a crude estimate of this effect, we have plotted in Fig. 7 the total interaction energy of the protein with water and ions used to neutralize the system. The slope of the overall interaction energy as a function of temperature decreases suggesting a shift of the entropy in the desired direction. We need to warn, however, that the simulation setup does not include the Debey–Hückel screening of the ionic interactions and the results are qualitative at best. Since our present simulations and most of other simulation data available in the literature have little to say about the ionic strength effects, laboratory measurements of redox entropies at different buffer concentrations are required to shed more light on this problem.

## ACKNOWLEDGMENTS

This research was supported by the National Science Foundation (CHE-0616646). The code used for the NRFT calculations is available at <http://theochemlab.asu.edu/>

codes.html. We are grateful to Marco Sola for introducing us to the problem of protein redox entropy.

- <sup>1</sup>C. L. Brooks, M. Karplus, and B. M. Pettitt, *Adv. Chem. Phys.* **71**, 1 (1988).
- <sup>2</sup>D. Eisenberg and A. D. McLachlan, *Nature (London)* **319**, 199 (1986).
- <sup>3</sup>R. Zhou, G. Krilov, and B. J. Berne, *J. Phys. Chem. B* **108**, 7528 (2004).
- <sup>4</sup>K. A. Dill, T. M. Truskett, V. Vlachy, and B. Hribar-Lee, *Annu. Rev. Biophys. Biomol. Struct.* **34**, 173 (2005).
- <sup>5</sup>M. Mueller, K. Katsov, and M. Schick, *Phys. Rep.* **434**, 113 (2006).
- <sup>6</sup>H. Gohlke and M. F. Thorpe, *Biophys. J.* **91**, 2115 (2006).
- <sup>7</sup>W. Rocchia, S. Sridharan, A. Nicholls, E. Alexov, A. Chiabrera, and B. Honig, *J. Comput. Chem.* **23**, 128 (2002).
- <sup>8</sup>M. Schaefer and M. Karplus, *J. Phys. Chem.* **100**, 1578 (1996).
- <sup>9</sup>M. R. Gunner and B. Honig, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 9151 (1991).
- <sup>10</sup>K. A. Sharp, *Biophys. J.* **73**, 1241 (1998).
- <sup>11</sup>K. Siriwong, A. A. Voityuk, M. D. Newton, and N. Röscher, *J. Phys. Chem. B* **107**, 2595 (2003).
- <sup>12</sup>D. Qiu, P. Shenkin, F. Hollinger, and W. Still, *J. Phys. Chem. A* **101**, 3005 (1997).
- <sup>13</sup>P. Vath, M. B. Zimmt, D. V. Matyushov, and G. A. Voth, *J. Phys. Chem. B* **103**, 9130 (1999).
- <sup>14</sup>D. V. Matyushov, *J. Chem. Phys.* **120**, 7532 (2004).
- <sup>15</sup>A. A. Milischuk, D. V. Matyushov, and M. D. Newton, *Chem. Phys.* **324**, 172 (2006).
- <sup>16</sup>P. K. Ghorai and D. V. Matyushov, *J. Phys. Chem. A* **110**, 8857 (2006).
- <sup>17</sup>R. Zhou, X. Huang, C. J. Margulis, and B. J. Berne, *Science* **305**, 1605 (2004).
- <sup>18</sup>M. Tarek and D. J. Tobias, *Phys. Rev. Lett.* **88**, 138101 (2002).
- <sup>19</sup>N. Choudhury and B. Pettitt, *J. Am. Chem. Soc.* **129**, 4847 (2007).
- <sup>20</sup>D. V. Matyushov, *Chem. Phys.* **174**, 199 (1993).
- <sup>21</sup>M. Ubbink, M. Ejdebäck, B. G. Karlsson, and D. S. Bendall, *Structure (London)* **6**, 323 (1998).
- <sup>22</sup>P. D. Swartz and T. Ichiye, *Biochemistry* **35**, 13772 (1996).
- <sup>23</sup>M. Sulpiuzi, S. Raugei, J. VandeVondele, P. Carloni, and M. Sprik, *J. Phys. Chem. B* **111**, 3969 (2007).
- <sup>24</sup>N. Sailasuta, F. C. Anson, and H. B. Gray, *J. Am. Chem. Soc.* **101**, 455 (1979).
- <sup>25</sup>G. Battistuzzi, M. Borsari, L. Loschi, F. Righi, and M. Sola, *J. Am. Chem. Soc.* **121**, 501 (1999).
- <sup>26</sup>G. Battistuzzi, M. Bellei, M. Borsari, G. W. Canters, E. de Waal, L. J. C. Jeuken, A. Ranieri, and M. Sola, *Biochemistry* **42**, 9214 (2003).
- <sup>27</sup>D. M. Lockwood, Y.-K. Cheng, and P. J. Rossky, *Chem. Phys. Lett.* **345**, 159 (2001).
- <sup>28</sup>S. N. Datta, J. Sudhamsu, and A. Pandey, *J. Phys. Chem. B* **108**, 8007 (2004).
- <sup>29</sup>D. F. Hansen and J. J. Led, *J. Am. Chem. Soc.* **126**, 1247 (2004).
- <sup>30</sup>E. Solomon, R. Szilagy, S. DeBeerGeorge, and L. Basumallick, *Chem. Rev. (Washington, D.C.)* **104**, 419 (2004).
- <sup>31</sup>E. I. Solomon, *Inorg. Chem.* **45**, 8012 (2006).
- <sup>32</sup>J. M. Guss, P. R. Harrowell, M. Murata, V. A. Norris, and H. C. Freeman, *J. Mol. Biol.* **192**, 361 (1986).
- <sup>33</sup>P. J. Stephens, D. R. Jollie, and A. Warshel, *Chem. Rev. (Washington, D.C.)* **96**, 2491 (1996).
- <sup>34</sup>M. Olsson, G. Hong, and A. Warshel, *J. Am. Chem. Soc.* **125**, 5025 (2003).
- <sup>35</sup>E. L. Yee, R. J. Cave, K. L. Guyer, P. D. Tyma, and M. J. Weaver, *J. Am. Chem. Soc.* **101**, 1131 (1979).
- <sup>36</sup>M. Born, *Z. Phys.* **1**, 45 (1920).
- <sup>37</sup>L. Onsager, *J. Am. Chem. Soc.* **58**, 1486 (1936).
- <sup>38</sup>J. G. Kirkwood, *J. Chem. Phys.* **2**, 351 (1934).
- <sup>39</sup>F. O. Raineri and H. L. Friedman, *Adv. Chem. Phys.* **107**, 81 (1999).
- <sup>40</sup>A. Yoshimori, *J. Theor. Comput. Chem.* **3**, 117 (2004).
- <sup>41</sup>I. Burghardt and B. Bagchi, *Chem. Phys.* **329**, 343 (2006).
- <sup>42</sup>T. Biben, J. P. Hansen, and Y. Rosenfeld, *Phys. Rev. E* **57**, R3727 (1998).
- <sup>43</sup>R. Ramirez, R. Gebauer, M. Mareschal, and D. Borgis, *Phys. Rev. E* **66**, 031206 (2002).
- <sup>44</sup>R. Ramirez and D. Borgis, *J. Phys. Chem. B* **109**, 6754 (2005).
- <sup>45</sup>A. A. Milischuk and D. V. Matyushov, *J. Chem. Phys.* **124**, 204502 (2006).
- <sup>46</sup>D. N. LeBard, M. Lilichenko, D. V. Matyushov, Y. A. Berlin, and M. A. Ratner, *J. Phys. Chem. B* **107**, 14509 (2003).
- <sup>47</sup>A. A. Milischuk and D. V. Matyushov, *J. Chem. Phys.* **123**, 044501 (2005).
- <sup>48</sup>D. V. Matyushov and M. D. Newton, *J. Phys. Chem. A* **105**, 8516 (2001).
- <sup>49</sup>G. Stell, G. N. Patey, and J. S. Høye, *Adv. Chem. Phys.* **18**, 183 (1981).
- <sup>50</sup>G. G. Ferenczy and C. A. Reynolds, *J. Phys. Chem. A* **105**, 11470 (2001).
- <sup>51</sup>D. V. Matyushov and B. M. Ladanyi, *J. Chem. Phys.* **110**, 994 (1999).
- <sup>52</sup>E. A. Carter and J. T. Hynes, *J. Chem. Phys.* **94**, 5961 (1991).
- <sup>53</sup>H. C. Andersen, D. Chandler, and J. D. Weeks, *Adv. Chem. Phys.* **34**, 105 (1976).
- <sup>54</sup>J.-K. Hwang and A. Warshel, *J. Am. Chem. Soc.* **109**, 715 (1987).
- <sup>55</sup>R. A. Kuharski, J. S. Bader, D. Chandler, M. Sprik, M. L. Klein, and R. W. Impey, *J. Chem. Phys.* **89**, 3248 (1988).
- <sup>56</sup>J. Blumberger and M. Sprik, *J. Phys. Chem. B* **109**, 6793 (2005).
- <sup>57</sup>T. Fonseca and B. M. Ladanyi, *J. Mol. Liq.* **60**, 1 (1994).
- <sup>58</sup>S. Lee and J. T. Hynes, *J. Chem. Phys.* **88**, 6853 (1988).
- <sup>59</sup>D. Chandler, *Phys. Rev. E* **48**, 2898 (1993).
- <sup>60</sup>J. P. Hansen and I. R. McDonald, *Theory of Simple Liquids* (Academic, Amsterdam, 2003).
- <sup>61</sup>J. D. Weeks, *Annu. Rev. Phys. Chem.* **53**, 533 (2002).
- <sup>62</sup>L. E. Fried and S. Mukamel, *J. Chem. Phys.* **93**, 932 (1990).
- <sup>63</sup>B. Bagchi and A. Chandra, *Adv. Chem. Phys.* **80**, 1 (1991).
- <sup>64</sup>P. Madden and D. Kivelson, *Adv. Chem. Phys.* **56**, 467 (1984).
- <sup>65</sup>C. J. F. Böttcher, *Theory of Electric Polarization* (Elsevier, Amsterdam, 1973), Vol. 1.
- <sup>66</sup>Y. I. Kharkats, A. A. Kornyshev, and M. A. Vorotyntsev, *J. Chem. Soc., Faraday Trans. 2* **72**, 361 (1976).
- <sup>67</sup>H. Li and M. Kardar, *Phys. Rev. A* **46**, 6490 (1992).
- <sup>68</sup>D. V. Matyushov, *J. Chem. Phys.* **120**, 1375 (2004).
- <sup>69</sup>W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran 77: The Art of Scientific Computing* (Cambridge University Press, Cambridge, 1996).
- <sup>70</sup>E. I. Solomon and M. D. Lowery, *Science* **259**, 1575 (1993).
- <sup>71</sup>L. W. Ungar, N. F. Scherer, and G. A. Voth, *Biophys. J.* **72**, 5 (1997).
- <sup>72</sup>G. M. Ullmann, E.-W. Knapp, and N. M. Kostić, *J. Am. Chem. Soc.* **119**, 42 (1997).
- <sup>73</sup>P. Comba and R. Remenyi, *J. Comput. Chem.* **23**, 697 (2002).
- <sup>74</sup>M. M. Werst, C. E. Davoust, and B. M. Hoffman, *J. Am. Chem. Soc.* **113**, 1533 (1991).
- <sup>75</sup>C. A. P. Libeu, M. Kikimoto, M. Nishiyama, S. Horinouchi, and E. T. Adman, *Biochemistry* **36**, 13160 (1997).
- <sup>76</sup>J. O. D. Kerpel and U. Ryde, *Proteins: Struct., Funct., Genet.* **36**, 157 (1999).
- <sup>77</sup>Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, and P. Kollman, *J. Comput. Chem.* **24**, 1999 (2003).
- <sup>78</sup>M. Cascella, A. Magistrato, I. Tavernelli, P. Carloni, and U. Rothlisberger, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 19641 (2006).
- <sup>79</sup>J. Blumberger and M. L. Klein, *J. Am. Chem. Soc.* **128**, 13854 (2006).
- <sup>80</sup>B. Guillot, *J. Mol. Liq.* **101**, 219 (2002).
- <sup>81</sup>W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, *J. Chem. Phys.* **79**, 926 (1983).
- <sup>82</sup>M. Neumann, *Mol. Phys.* **57**, 97 (1986).
- <sup>83</sup>M. S. Wertheim, *J. Chem. Phys.* **55**, 4291 (1971).
- <sup>84</sup>M. S. Wertheim, *Mol. Phys.* **37**, 83 (1979).
- <sup>85</sup>S. Gupta and D. V. Matyushov, *J. Phys. Chem. A* **108**, 2087 (2004).
- <sup>86</sup>D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, Jr., A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods, *J. Comput. Chem.* **26**, 1668 (2005).
- <sup>87</sup>Y. Xue, M. Okvist, O. Hansson, and S. Young, *Protein Sci.* **7**, 2099 (1998).
- <sup>88</sup>H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, *J. Chem. Phys.* **81**, 3684 (1984).
- <sup>89</sup>H. Reiss, *J. Electrochem. Soc.* **135**, 247C (1988).
- <sup>90</sup>L. D. Landau and E. M. Lifshits, *Statistical Physics* (Pergamon, New York, 1980).
- <sup>91</sup>J. O. Bockris, *Modern Electrochemistry* (McDonald, London, 1970).
- <sup>92</sup>W. Schmickler, *Interfacial Electrochemistry* (Oxford University Press, New York, 1996).
- <sup>93</sup>R. A. Marcus, *J. Chem. Phys.* **43**, 679 (1965).
- <sup>94</sup>H. Reiss, *J. Phys. Chem.* **89**, 3783 (1985).
- <sup>95</sup>A. V. Gorodyskii, A. I. Karasevskii, and D. V. Matyushov, *J. Electroanal. Chem.* **315**, 9 (1991).

- <sup>96</sup>G. R. Moore, G. W. Pettigrew, and N. K. Rogers, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 4998 (1986).
- <sup>97</sup>E. D. German and A. M. Kuznetsov, *Elektrokhimiya* **28**, 294 (1992).
- <sup>98</sup>T. Hotta, A. Kimura, and M. Sasai, *J. Phys. Chem. B* **109**, 18600 (2005).
- <sup>99</sup>F. H. Stillinger, *J. Solution Chem.* **2**, 141 (1973).
- <sup>100</sup>G. Hummer and S. Garde, *Phys. Rev. Lett.* **80**, 4193 (1998).
- <sup>101</sup>D. Chandler, *Nature (London)* **437**, 640 (2005).
- <sup>102</sup>J. Florian and A. Warshel, *J. Phys. Chem. B* **103**, 10282 (1999).
- <sup>103</sup>P. K. Ghorai and D. V. Matyushov, *J. Phys. Chem. B* **110**, 1866 (2006).
- <sup>104</sup>D. N. LeBard and D. V. Matyushov, *J. Phys. Chem. B* **112**, 5218 (2008).
- <sup>105</sup>C. Hartnig and M. T. M. Koper, *J. Chem. Phys.* **115**, 8540 (2001).
- <sup>106</sup>H. S. Ashbaugh, *J. Phys. Chem. B* **104**, 7235 (2000).
- <sup>107</sup>S. Rajamani, T. Ghosh, and S. Garde, *J. Chem. Phys.* **120**, 4457 (2004).
- <sup>108</sup>D. S. Cerutti, N. A. Baker, and J. A. McCammon, *J. Chem. Phys.* **127**, 155101 (2007).
- <sup>109</sup>L. R. Pratt and D. Chandler, *J. Chem. Phys.* **67**, 3683 (1977).
- <sup>110</sup>S. Garde, G. Hummer, A. E. García, M. E. Paulaitis, and L. R. Pratt, *Phys. Rev. Lett.* **77**, 4966 (1996).
- <sup>111</sup>H. S. Ashbaugh and L. R. Pratt, *Rev. Mod. Phys.* **78**, 159 (2006).
- <sup>112</sup>K. Lum, D. Chandler, and J. Weeks, *J. Phys. Chem. B* **103**, 4570 (1999).
- <sup>113</sup>D. M. Huang and D. Chandler, *Phys. Rev. E* **61**, 1501 (2000).
- <sup>114</sup>P. W. Fenimore, H. Frauenfelder, B. H. McMahon, and R. D. Young, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 14408 (2004).
- <sup>115</sup>J. Dzubiella, J. M. J. Swanson, and J. A. McCammon, *Phys. Rev. Lett.* **96**, 087802 (2006).
- <sup>116</sup>J. T. Hupp, Y. Dong, R. L. Blackbourn, and H. Lu, *J. Phys. Chem.* **97**, 3278 (1993).
- <sup>117</sup>D. M. D'Alessandro, P. C. Junk, and F. R. Keene, *Supramol. Chem.* **17**, 529 (2005).
- <sup>118</sup>14 Na<sup>+</sup> and 6 Cl<sup>-</sup> ions dissolved in 1402 water molecules around rubredoxin protein were used in the simulations in Ref. 22. Given that electrostatic interactions were truncated and the number of water molecules was probably insufficient to build the complete second solvation shell, it is not clear whether this simulation protocol provided an adequate modeling of the ionic strength effects.