JMB

Calculation of the Dielectric Properties of a Protein and its Solvent: Theory and a Case Study

Gerald Löffler, Hellfried Schreiber and Othmar Steinhauser*

University of Vienna, Institute for Theoretical Chemistry Molecular Dynamics Group Währingerstraße 17/Parterre A-1090 Wien, Austria

This paper presents a rigorous derivation of a theory for the calculation of the frequency-dependent dielectric properties of each component of the system protein/water/ions with the aim of enabling comparison to experimentally determined dielectric properties. We apply this theory to a very long (13.1 ns) molecular dynamics simulation of an HIV1 zinc finger peptide, its co-ordinated zinc ion, and two chloride ions in a box of SPC/E water molecules. We find the dielectric relaxation of the water molecules restricted compared to pure water, giving rise to a static dielectric constant for the water-component of only 47. The peptide is found to have a complicated dielectric relaxation behaviour, with a static dielectric constant of 15. We also calculate the frequency-dependent conductivity of the ions in this system. We analyze all contributions to the calculation of these dielectric properties and find that the coupling between the dielectric relaxation of the peptide and that of the water-component is particularly important for correctly describing the dielectric constant of the peptide.

© 1997 Academic Press Limited

*Corresponding author

Introduction

It should be obvious that the calculation of such fundamental properties of matter as the frequency-dependent dielectric constant and the frequency-dependent conductivity from computer simulations is almost a goal in itself. Apart from that, a whole range of very popular algorithms that involve the calculation of the electrostatic field generated by a protein depend on the solution of the Poisson-Boltzmann (Gilson & Honig, 1987) equation. Solving this equation requires a value for the static dielectric constant of the protein and the surrounding dielectric medium. The latter is usually water and consequently does not pose a problem. However, the calculation of the former is still the subject of active research, and the goal of the work presented in this paper. The remaining part of this introduction will briefly review some important publications in the area of the calculation of the static or frequency-dependent dielectric properties of proteins or other biomolecules in aqueous solution. Along with a short description of our approach, this should serve the purpose of setting our work in the proper context. It should

be noted that the theory section of the paper should probably be read before the following literature overview, because many of the terms used in the overview are introduced in the theory section.

In 1988, when the extensive molecular dynamics (MD) simulations of proteins that are required for the calculation of dielectric properties were not yet feasible, Nakamura et al. (1988) calculated local static dielectric constants of bovine pancreatic trypsin inhibitor (BPTI) from normal mode analysis *in vacuo*. Electronic polarization of atoms and orientational polarization of local dipoles were considered and resulted in local dielectric constants ranging from 1 to 20 inside the protein.

King et al. (1991) evaluated the local static dielectric constant of different sites in trypsin and the static dielectric constant of pure water, based on MD simulations of 50 ps length. The MD simulation employed the SCAAS model which surrounds the solute with layers of solvent that range from unrestricted water molecules over restricted water molecules up to an electrostatic continuum Furthermore, they did not truncate electrostatic interactions. They calculated static dielectric constants from Kirkwood-Fröhlich theory (which was developed for a dielectric sphere surrounded by a continuum of uniform dielectric constant) and di-

---

Abbreviations used: MD, molecular dynamics; BPTI, bovine pancreatic trypsin inhibitor.

rectly *via* the average electric field and average polarization that result from the simulation. The resulting local static dielectric constants for sites in trypsin ranged from 3 to 20, where only the Kirkwood-Fröhlich approach gave dielectric constants above 10. They also gave the dipole auto-correlation functions for some protein sites, which varied among each other but commonly decayed to zero within the very short period of 14 ps.

Simonson *et al.* (1991) concentrated on the calculation of local static dielectric properties of proteins because they argued that the local variation of the dielectric properties is important for the biological function of a protein. To this end, they calculated the local generalized susceptibility out of the dipole-dipole correlation function of each site. They performed *in vacuo* MD simulations of deca-alanine (150 ps after equilibration) and cytochrome *c* (90 ps after equilibration) using the Charmm 19 force field. They estimated the static dielectric constant of deca-alanine to be 3.3 and of cytochrome *c* to be 3.5, based on an equation valid for a dielectric sphere surrounded by a continuum of uniform dielectric constant (Kirkwood-Fröhlich theory). They observed spatial variations of the susceptibility by a factor of 4 between different residues of the proteins.

Smith *et al.* (1993) performed a 1.4 ns and 1 ns MD simulation of BPTI and lysozyme in an explicit solvent environment (SPC/E or SPC), respectively, using the GROMOS force field and a twin range method based on the Coulomb potential for the calculation of the electrostatic interaction. They calculated the static and frequency-dependent dielectric constant of each protein under the assumption that the protein is spherical and that cross-correlation between the dipole moments of the protein and the water component can be neglected. In general, they applied the theory developed by Neumann *et al.* (1984) for one-component dielectrics to the protein in the three-component system protein/water/ions. Furthermore, since the simulation times of about 1 ns were not sufficiently long to sample the overall rotation of the proteins, they removed this motion (and translation as well, for that matter) during the analysis phase. Since the proteins were charged, and consequently the dipole moments of the proteins dependent on the origin of the coordinate system, they chose the centre of mass as the origin. The static dielectric constants were calculated to be 36 and 30 for BPTI and lysozyme, respectively. Neglecting the side-chains in calculating the dipole fluctuations led to static dielectric constants between 2 and 3, which is consistent with the picture of a low-dielectric protein core and a high-dielectric protein-water interface. The auto-correlation function of the dipole moment of each protein was fitted to a single-exponential function, giving rise to relaxation times of 1.8 ns and 3.6 ns, respectively. Based on these fits, the frequency-dependent dielectric constants were calculated and seen to vanish in the region above 10 MHz.

Antosiewicz *et al.* (1994) calculated pK_a values of ionizable groups in proteins using an algorithm that relies on the solution of the Poisson-Boltzmann equation for the calculation of electrostatic fields of proteins. These algorithms need the static dielectric constant of the protein as an input value (Gilson & Honig, 1987). Traditionally, a low dielectric constant of 2 to 4 was used, but Antosiewicz *et al.* report best agreement of their calculated pK_as with experiments when they chose a static dielectric constant of 20 for the protein. They argue that the high dielectric constant of the protein is needed to account for conformational relaxation effects that are not otherwise modelled in the algorithm.

In a considerable extension of their work from 1991, Simonson & Perahia (1995) applied the Kirkwood-Fröhlich theory to a 1 ns MD simulation of ferro- and ferricytochrome *c* in a spherical volume of water molecules, respectively. They found that the strong motion of the charged protein side-chains at the protein-solvent interface had a large effect on the protein dipole fluctuations. They argued that it is incorrect to consider these side-chains as being part of the protein, because if this is done, the protein can no longer be treated as a homogeneous dielectric material. If this is done nevertheless, the static dielectric constant was found to vary between 16 and 37. However, if the charged protein side-chains at the protein-solvent interface are viewed as part of the solvent, they found static dielectric constants of 4.7 and 3.4 for ferro- and ferricytochrome *c*, respectively. Again, they studied the spatial variation of the static dielectric properties of the protein, and found the static dielectric constant of the inner half of the proteins to be 1.5 to 2. They suggested that calculations involving the Poisson-Boltzmann equation should treat the inner part of a protein as a low-dielectric medium, whereas the charged surface groups should be seen as part of the solvent region.

Yang *et al.* (1995) performed a 1.155 ns MD simulation of the triple helical DNA strand d(CG·G)$_7$ in an ionic solution consisting of 837 water molecules, 37 sodium ions, and 16 chloride ions, using the Charmm22 force field and employing Ewald summation for the electrostatic interaction. They decomposed the system into the five components: base, sugar, phosphate, water, and ions. Since the phosphate and ion groups carried a net charge, their dipole moment was dependent on the origin of the coordinate system, which they chose to be the centre of mass of the DNA. In this way they also eliminated the contribution of the DNA to the conductivity of the system. They calculated the static dielectric constant of each of these components out of the fluctuations of the dipole moment of the respective component, thereby neglecting any cross-terms (which they evaluated and found to be small). Doing so they found the following static dielectric constants: 41.3 for water, 3.4 for the bases, 2.0 for the sugars, 33.0 for the phosphate groups,

and 15.5 for the DNA as a whole. They also calculated a static dielectric constant as opposed to a conductivity for the ion component. They chose the SPC/E water model and found a collective dipole relaxation time of 9.7 ps for this model in their system.

We see several problems with the above approaches to the calculation of the dielectric properties of solvated biomolecules in general and proteins in particular. (1) Dielectric properties are macroscopic properties. This is the reason why we consider proper boundary conditions a prerequisite for the ability of an MD simulation to correctly reproduce the dielectric properties of a system. We think that experience has shown that only periodic boundary conditions (either the usual toroidal boundary conditions which are based on a box-shaped simulation system, or simulations on a hypersphere) are the appropriate answer to this problem. In the simulation reported here, we used a box-shaped simulation system with periodic boundary conditions. (2) On the same line of argument, we believe that the rigorous treatment of electrostatic interactions is vital for a meaningful description of the dielectric properties of a system. This was clearly shown by Neumann & Steinhauser (1983a,b) and Neumann *et al.* (1984) for one-component systems. Unfortunately, employment of the Coulomb potential in conjunction with a cutoff (be it switched or not) is not enough. Similarly, in our opinion, full Coulomb interaction (without cutting of interactions) suffers from the problem that the long-range behaviour of the Coulomb potential leads to electrostatic boundary effects throughout the system, because no system can be large enough to not interfere with Coulomb interactions. What remains are techniques like lattice summation (Ewald summation being a special case (de Leeuw *et al.*, 1980a,b)), equivalent methods (Particle Mesh Ewald (Darden *et al.*, 1993)), and reaction field methods (Tironi *et al.*, 1995) for the treatment of the electrostatic interactions. Here, we used a classical implementation of Ewald summation. (3) The combination of Linear Response theory, the macroscopic definition of dielectric properties (phenomenological equations of matter), and a computer-adapted version of dielectric theory was very successful in the case of one-component systems (Neumann & Steinhauser, 1983a,b; Neumann *et al.*, 1984) and ionic solutions (Caillol, 1987; Caillol *et al.*, 1986, 1987, 1989a,b; Löffler *et al.*, 1997). We do not see a reason to depart from this approach, especially since it is questionable if the direct application of macroscopic formulas such as Kirkwood-Fröhlich theory to the results of MD simulations leads to correct answers. Consequently, we derived a theory of the frequency-dependent dielectric properties of all components of the three-component system protein/water/ions, which will be presented in the theory section of this paper, building on these experiences from one-component systems. (4) Furthermore, the results of Smith *et al.* (1993) clearly show that extremely long

MD simulations are required to sample all dielectric relaxation modes of a protein (including overall rotation). To meet this challenge we performed a 13.1 ns (after equilibration) MD simulation of a 18-residue peptide in an explicit water environment.

Theory

The dielectric properties of a system describe the reaction of this system to an externally applied electric field, which is usually taken to be spatially homogeneous. In the following we shall see that while the definition of the dielectric properties according to the phenomenological equations of matter involves the field acting inside the system (often termed the Maxwell field), linear response theory can give expressions for system properties only as a function of the external electric field. However, using fundamental electrostatic relations, it is possible to relate the Maxwell field to the external field.

Phenomenological equations of matter

For a macroscopic piece of matter, and for our simulation cell under periodic boundary conditions, if we assume it to be macroscopic, the following phenomenological equations (the so-called constitutive relations) define the dielectric properties of the components of the system:

$$\tilde{\mathbf{P}}_W(\omega) = \frac{\varepsilon_W(\omega) - 1}{4\pi} \tilde{\mathbf{E}}(\omega) \qquad (1)$$

is the definition of the frequency-dependent dielectric constant $\varepsilon_W(\omega)$ of the water-component of the system expressed in terms of the frequency-components $\tilde{\mathbf{E}}(\omega)$ of the Maxwell field (i.e. the field acting inside the system) and the frequency-components $\tilde{\mathbf{P}}_W(\omega)$ of the polarization of the water-component of the system. The polarization $\mathbf{P}_X(t)$ of a component X of the system is defined as the dipole-density of this component:

$$\mathbf{P}_X(t) = \frac{1}{V} \mathbf{M}_X(t) = \frac{1}{V} \sum_{i \in X} (q_i(t)\mathbf{r}_i(t)) \qquad (2)$$

Herein, q_i denotes the charge and $\mathbf{r}_i$ the position of particle i (being a member of the component X), and $\mathbf{M}_X$ is the dipole-moment of component X. The step from the time-domain to the frequency-domain is done *via* Fourier transformation, as usual.

Similarly:

$$\tilde{\mathbf{P}}_P(\omega) = \frac{\varepsilon_P(\omega) - 1}{4\pi} \tilde{\mathbf{E}}(\omega) \qquad (3)$$

defines the frequency-dependent dielectric constant $\varepsilon_P(\omega)$ of the protein expressed in terms of the Maxwell field and the frequency-components $\tilde{\mathbf{P}}_P(\omega)$ of the polarization of the protein.

Equations (1) and (3) define a linear relationship between the Maxwell field and the polarization of the components water and protein, respectively. For a justification of these simple relations, we refer to Appendix I. Finally:

$$\tilde{\mathbf{i}}_I(\omega) = \sigma_I(\omega)\tilde{\mathbf{E}}(\omega) \tag{4}$$

defines the frequency-dependent conductivity $\sigma_I(\omega)$ of the ion-component of the system expressed in terms of the Maxwell field and the frequency-components $\mathbf{i}_I(\omega)$ of the current-density of the ions. The latter is defined:

$$\mathbf{i}_I(t) = \frac{1}{V}\mathbf{J}_I(t) = \frac{1}{V}\sum_{i \in I}(q_i(t)\mathbf{v}_i(t)) \tag{5}$$

where $\mathbf{v}_i$ is the velocity of particle i and $\mathbf{J}_I$ the current of the ion-component.

One important implication of the formulation of the constitutive relations (1), (3), and (4) in conjunction with the definitions (2) and (5) is the fact that the frequency-dependent dielectric properties depend on the concentration of the respective component in the system! This is a direct consequence of a theoretical formulation that is compatible with the usual situation in dielectric experiments (Appendix I).

Linear response theory

We shall use linear response theory to describe the reaction of a system to an externally applied, homogeneous electric field $\mathbf{E}_0(t)$ that may vary with time. These are exactly the conditions under which measurements of the dielectric properties of matter are performed. Linear response theory (Kohler, 1972; Caillol *et al.*, 1986, 1989a; Neumann & Steinhauser, 1983a,b; Neumann *et al.*, 1984; Löffler *et al.*, 1997) allows us to express the polarization of the water-component and the protein, as well as the current-density of the ions in terms of collective correlation functions of the unperturbed system. The polarizations and the current-density can then be substituted into the phenomenological equations of matter.

In general, linear response theory states that the expectation values $\langle\tilde{\mathbf{O}}(\omega)\rangle$ of the frequency-components of an observable $\mathbf{O}$ are directly proportional to the frequency-components $\tilde{\mathbf{E}}_0(\omega)$ of he external field:

$$\langle\tilde{\mathbf{O}}(\omega)\rangle = \chi_{\mathbf{OP}}(\omega)\tilde{\mathbf{E}}_0(\omega) \tag{6}$$

The proportionality factor $\chi_{\mathbf{OP}}(\omega)$ is called generalized susceptibility and captures the coupling of $\mathbf{O}$ to the total polarization $\mathbf{P}$ of the system. The generalized susceptibility is defined as:

$$\chi_{\mathbf{OP}}(\omega) = \frac{V}{3kT}\int_0^\infty \langle\mathbf{O}(0)\dot{\mathbf{P}}(-t)\rangle e^{-i\omega t}\, dt \tag{7}$$

where a term of the form $\langle\mathbf{A}(0)\mathbf{B}(t)\rangle$ denotes the time-correlation function of the vectorial quantities $\mathbf{A}$ and $\mathbf{B}$ (and is of course identical to $\langle\mathbf{B}(0)\mathbf{A}(t)\rangle$).

For our special case of the three-component system of a protein immersed in an aqueous ionic solution, the total polarization $\mathbf{P}$ is of course:

$$\mathbf{P}(t) = \frac{1}{V}[\mathbf{M}_W(t) + \mathbf{M}_P(t) + \mathbf{M}_I(t)] \tag{8}$$

The dipole-moment of a charge distribution is independent of the origin of the coordinate-system only if the charge distribution is electrically neutral. Hence the only meaningful total polarization in this context results from a neutral simulation system, something that is required for molecular dynamics simulations, anyway. Similarly, each component of the system has to be electrically neutral itself. This is always true for the water-component, since each water molecule is neutral. Electrical neutrality of the protein, however, simply has to be assumed for this derivation. The Appendix will show what to do if the protein is not neutral. Given a neutral protein, the ion-component of the system has to be neutral, too, if total charge neutrality is to be preserved.

Using the special equation (8) for the total polarization, performing partial integration and taking into account that $\dot{\mathbf{M}}_I = \mathbf{J}_I$, the generalized susceptibility becomes:

$$\chi_{\mathbf{OP}}(\omega) = \frac{1}{3kT}\Bigg[\langle\mathbf{O}(0)\mathbf{M}_W(0)\rangle$$
$$- i\omega\int_0^\infty \langle\mathbf{O}(0)\mathbf{M}_W(t)\rangle e^{-i\omega t}\, dt$$
$$+ \langle\mathbf{O}(0)\mathbf{M}_P(0)\rangle - i\omega\int_0^\infty \langle\mathbf{O}(0)\mathbf{M}_P(t)\rangle e^{-i\omega t}\, dt$$
$$+ \int_0^\infty \langle\mathbf{O}(0)\mathbf{J}_I(t)\rangle e^{-i\omega t}\, dt\Bigg] \tag{9}$$

Using this formula, the polarizations and the current-density appearing in the phenomenological equations (1), (3) and (4) can be calculated from linear response theory: The polarization of the water-component of the system:

$$\langle\tilde{\mathbf{P}}_W(\omega)\rangle = \chi_{\mathbf{P}_W\mathbf{P}}(\omega)\tilde{\mathbf{E}}_0(\omega) \tag{10}$$

with

$$\chi_{\mathbf{P}_W\mathbf{P}}(\omega) = \frac{1}{3VkT}\Bigg[\langle\mathbf{M}_W(0)\mathbf{M}_W(0)\rangle$$
$$- i\omega\int_0^\infty \langle\mathbf{M}_W(0)\mathbf{M}_W(t)\rangle e^{-i\omega t}\, dt$$
$$+ \langle\mathbf{M}_W(0)\mathbf{M}_P(0)\rangle - i\omega\int_0^\infty \langle\mathbf{M}_W(0)\mathbf{M}_P(t)\rangle$$
$$\times e^{-i\omega t}\, dt + \int_0^\infty \langle\mathbf{M}_W(0)\mathbf{J}_I(t)\rangle e^{-i\omega t}\, dt\Bigg] \tag{11}$$

the polarization of the protein:

$$\langle \tilde{\mathbf{P}}_P(\omega) \rangle = \chi_{\mathbf{P}_P\mathbf{P}}(\omega) \tilde{\mathbf{E}}_0(\omega) \qquad (12)$$

with:

$$\chi_{\mathbf{P}_P\mathbf{P}}(\omega) = \frac{1}{3VkT}\Bigg[\langle \mathbf{M}_P(0)\mathbf{M}_W(0) \rangle$$

$$- i\omega \int_0^\infty \langle \mathbf{M}_P(0)\mathbf{M}_W(t) \rangle \, e^{-i\omega t} \, dt$$

$$+ \langle \mathbf{M}_P(0)\mathbf{M}_P(0) \rangle - i\omega \int_0^\infty \langle \mathbf{M}_P(0)\mathbf{M}_P(t) \rangle$$

$$\times e^{-i\omega t} \, dt + \int_0^\infty \langle \mathbf{M}_P(0)\mathbf{J}_I(t) \rangle \, e^{-i\omega t} \, dt \Bigg] \quad (13)$$

and the current-density of the ion-component of the system:

$$\langle \tilde{\mathbf{i}}_I(\omega) \rangle = \chi_{\mathbf{I}_P\mathbf{P}}(\omega) \tilde{\mathbf{E}}_0(\omega) \qquad (14)$$

with:

$$\chi_{\mathbf{i}_I\mathbf{P}}(\omega) = \frac{1}{3VkT}\Bigg[-i\omega \int_0^\infty \langle \mathbf{J}_I(0)\mathbf{M}_W(t) \rangle \, e^{-i\omega t} \, dt - i\omega$$

$$\times \int_0^\infty \langle \mathbf{J}_I(0)\mathbf{M}_P(t) \rangle \, e^{-i\omega t} \, dt$$

$$+ \int_0^\infty \langle \mathbf{J}_I(0)\mathbf{J}_I(t) \rangle \, e^{-i\omega t} dt \Bigg] \qquad (15)$$

In this last expression for $\chi_{\mathbf{i}_I\mathbf{P}}$ we made use of the fact that the static value (at time $t = 0$) of cross-correlation functions involving the current of the ions vanishes†, i.e. $\langle \mathbf{M}_W(0)\mathbf{J}_I(0) \rangle = 0$ and $\langle \mathbf{M}_P(0)\mathbf{J}_I(0) \rangle = 0$.

One of the central tasks in calculating the dielectric properties of a system consisting of a protein in an ionic solution will be the determination of the time-correlation functions appearing in the above equations from a molecular dynamics simulation in order to be able to calculate the generalized susceptibilities $\chi_{\mathbf{P}_W\mathbf{P}}(\omega)$, $\chi_{\mathbf{P}_P\mathbf{P}}(\omega)$, and $\chi_{\mathbf{i}_I\mathbf{P}}(\omega)$.

The relation of the external field to the Maxwell-field

The results from linear response theory involve the external electric field, while the phenomenological equations of matter are formulated in terms of the Maxwell field. Thus we need an expression for the relation between these two fields (Caillol *et al.*, 1986, 1989a; Neumann & Steinhauser, 1983a,b; Neumann *et al.*, 1984; Löffler *et al.*, 1997).

The electric field $\mathbf{E}(\mathbf{r}, t)$ at point $\mathbf{r}$ inside the system and at time t is determined by the external

---

† This can be seen by substituting $t = 0$ into the general equation $\langle \mathbf{A}(0)\mathbf{B}(-t) \rangle = -\langle \mathbf{A}(0)\mathbf{B}(t) \rangle$, which can only be true if $\langle \mathbf{A}(0)\mathbf{B}(0) \rangle = 0$. This applies to our case because $\mathbf{J}_I = \dot{\mathbf{M}}_I$.

field $\mathbf{E}_0(t)$ (which is assumed to be spatially homogeneous and hence does not depend on $\mathbf{r}$), a contribution from the charges and a contribution from the dipoles:

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0(t) - \int_V \nabla_{\mathbf{r}}\Phi(\mathbf{r} - \mathbf{r}')\rho_I(\mathbf{r}', t) \, d\mathbf{r}'$$

$$+ \int_V \nabla_{\mathbf{r}}\nabla_{\mathbf{r}}\Phi(\mathbf{r} - \mathbf{r}')[\mathbf{P}_W(\mathbf{r}', t) + \mathbf{P}_P(\mathbf{r}', t)] \, d\mathbf{r}'$$

$$(16)$$

The contribution of the charges in the system is formulated in terms of the charge-density $\rho_I(\mathbf{r}, t)$ of the ion-component of the system. Similarly, the contribution of the dipoles is formulated in terms of the polarization of the water and protein-components $\mathbf{P}_W(\mathbf{r}, t) + \mathbf{P}_P(\mathbf{r}, t)$ of the system. The electrostatic interaction which is employed in the simulation is characterized be the electrostatic potential $\Phi(\mathbf{r})$, the Ewald potential in our case. The double gradient of the electrostatic potential with respect to the position is called T-tensor, i.e. $\mathbf{T}(\mathbf{r}) \equiv \nabla_{\mathbf{r}}\nabla_{\mathbf{r}}\Phi(\mathbf{r})$.

To get the Maxwell field, we average $\mathbf{E}(\mathbf{r}, t)$ over the entire system volume:

$$\mathbf{E}(t) = \frac{1}{V}\int_V \mathbf{E}(\mathbf{r}, t) \, d\mathbf{r} \qquad (17)$$

and switch from $\mathbf{r}$ to $\mathbf{r}'' = \mathbf{r} - \mathbf{r}'$, which is, due to the periodic boundary conditions employed in the simulation, always inside the volume of the system (Neumann & Steinhauser, 1983a,b; Neumann *et al.*, 1984):

$$\mathbf{E}(t) = \mathbf{E}_0(t) - \frac{1}{V}\int_V \nabla_{\mathbf{r}''}\Phi(\mathbf{r}'') \, d\mathbf{r}'' \int_V \rho_I(\mathbf{r}', t) \, d\mathbf{r}'$$

$$+ \frac{1}{V}\int_V \mathbf{T}(\mathbf{r}'') \, d\mathbf{r}'' \int_V \mathbf{P}_W(\mathbf{r}', t) + \mathbf{P}_P(\mathbf{r}', t) \, d\mathbf{r}' \ (18)$$

We have already demanded that the sum of all ionic charges be zero, so:

$$\int_V \rho_I(\mathbf{r}', t) \, d\mathbf{r}' = 0 \qquad (19)$$

Furthermore:

$$\frac{1}{V}\int_V \mathbf{P}_W(\mathbf{r}', t) + \mathbf{P}_P(\mathbf{r}', t) \, d\mathbf{r}' = \mathbf{P}_W(t) + \mathbf{P}_P(t) \quad (20)$$

Lastly, as was shown before (Neumann & Steinhauser, 1983a,b; Neumann *et al.*, 1984), the integral over the T-tensor in an isotropic system is a multiple of the unit-tensor $\mathbf{I}$ and can thus be identified with the single scalar value ε_{RF} which exhibits formal equivalence to a static dielectric constant and is thus termed as such:

$$\int_V \mathbf{T}(r'') \, dr'' = \frac{4\pi}{3}\left[\frac{2(\varepsilon_{RF} - 1)}{2\varepsilon_{RF} + 1} - 1\right]\mathbf{I} \qquad (21)$$

How to calculate ε_{RF} has been discussed before

(Neumann *et al.*, 1984), and it suffices here to say that for an ideal implementation of the Ewald summation $\varepsilon_{RF} = \infty$. Indeed, as we will show when presenting the details of the simulation reported in this work, we came very close to this ideal implementation, so that we can safely assume that ε_{RF} is infinite.

After incorporating these equations, substituting the phenomenological equations of matter (1) and (3) for $\mathbf{P}_W$ and $\mathbf{P}_P$, respectively, and switching to frequency-components, we arrive at the relation of the Maxwell field to the external field:

$$\frac{\tilde{\mathbf{E}}(\omega)}{\tilde{\mathbf{E}}_0(\omega)} = \frac{2\varepsilon_{RF} + 1}{2\varepsilon_{RF} + \varepsilon_W(\omega) + \varepsilon_P(\omega) - 1} \quad (22)$$

For the case of a pure ionic solution, $\varepsilon_P(\omega)$ is 1 (just like for vacuum) and this relation nicely specializes to the expression derived elsewhere for ionic solutions (Löffler *et al.*, 1997).

Resulting equations

Using the results from linear response theory, the phenomenological equations of matter and the relation between external field and Maxwell field, we can give the expressions for the dielectric properties of the three components of a system consisting of a protein in an aqueous ionic solution.

The frequency-dependent dielectric constant $\varepsilon_W(\omega)$ of the water-component can be calculated by:

$$\frac{\varepsilon_W(\omega) - 1}{4\pi} = f(\omega)\chi_{\mathbf{P}_W\mathbf{P}}(\omega) \quad (23)$$

the frequency-dependent dielectric constant $\varepsilon_P(\omega)$ of the protein is accessible *via*:

$$\frac{\varepsilon_P(\omega) - 1}{4\pi} = f(\omega)\chi_{\mathbf{P}_P\mathbf{P}}(\omega) \quad (24)$$

and the frequency-dependent conductivity $\sigma_I(\omega)$ of the ion-component of the system is:

$$\sigma_I(\omega) = f(\omega)\chi_{\mathbf{i}_I\mathbf{P}}(\omega) \quad (25)$$

The common factor $f(\omega)$ in the above equations is:

$$f(\omega) = \left[1 - \frac{4\pi[\chi_{\mathbf{P}_W\mathbf{P}}(\omega) + \chi_{\mathbf{P}_P\mathbf{P}}(\omega)]}{2\varepsilon_{RF} + 1}\right] \quad (26)$$

The generalized susceptibilities $\chi_{\mathbf{P}_W\mathbf{P}}(\omega)$, $\chi_{\mathbf{P}_P\mathbf{P}}(\omega)$, and $\chi_{\mathbf{i}_I\mathbf{P}}(\omega)$ are defined by equations (11), (13) and (15), respectively.

It can easily be seen that in the case of an ideal implementation of the Ewald summation, as was achieved in the simulation reported below, the infinite ε_{RF} leads to $f(\omega) = 1$ and consequently the equations for the dielectric properties reduce to the following simple form:

$$\varepsilon_W(\omega) = 4\pi\chi_{\mathbf{P}_W\mathbf{P}}(\omega) + 1 \quad (27)$$

$$\varepsilon_P(\omega) = 4\pi\chi_{\mathbf{P}_P\mathbf{P}}(\omega) + 1 \quad (28)$$

$$\sigma_I(\omega) = \chi_{\mathbf{i}_I\mathbf{P}}(\omega) \quad (29)$$

These expressions show that Ewald summation leads to a decoupling of the dielectric properties of the system such that each dielectric property depends only on its own generalized susceptibility. Again, all the above equations are generalizations of the respective equations for pure ionic systems (Caillol *et al.*, 1986; Löffler *et al.*, 1997) and reduce to those if the protein is absent and thus $\chi_{\mathbf{P}_P\mathbf{P}}(\omega) = 0$.

Connection with the experiment

The individual dielectric properties ε_W, ε_P, and σ_I are important quantities from a theoretical point of view, but they cannot be determined experimentally. Only an overall dielectric response of the system can be extracted from experimental dielectric studies. It is the total conductivity that governs this overall dielectric behaviour, and the corresponding observable is (Caillol *et al.*, 1986):

$$\Sigma(\omega) = \frac{4\pi i}{\omega}\sigma(\omega) + \varepsilon_P(\omega) + \varepsilon_W(\omega) - 2 \quad (30)$$

We think it is a very important aspect of the theory presented in this section that it enables us to directly predict an experimental quantity.

Simulation

We performed a 13.1 ns (excluding equilibration time) molecular dynamics (MD) simulation of a system consisting of an electrically neutral 18-residue zinc finger peptide (Summers *et al.*, 1990), a zinc ion as a ligand to this peptide, two chloride ions as counter-ions to the zinc ion, and 2872 SPC/E water molecules in a periodic box of the size $(4.5 \text{ nm})^3$. Thus the concentration of the protein is 18.2 mmol 1^{-1}. The united-atom Charmm19 force-field (Brooks *et al.*, 1983) was used for non-water-water interactions. The simulation was performed using a home-grown MD-program (Löffler & Maccallum, 1995; Löffler & Schreiber, 1996) on Edinburgh Parallel Computing Centre's Cray T3D and on an SGI Power Challenge equipped with eight R8000 CPUs.

Electrostatic interactions were calculated using a conventional, albeit highly parallel, implementation of Ewald summation. The relevant parameters were: (1) 399 k-vectors; (2) twin-range cut-off of 0.9 nm and 1.2 nm, respectively; (3) Ewald decay parameter $\eta = 1.8 \text{ nm}^{-1}$. These parameters lead to a *de facto* ideal implementation of Ewald summation for this system.

The peptide comprises the first zinc finger domain from the gag protein p55 from HIV1 (Summers *et al.*, 1990). This peptide contains a CCHC zinc finger motif, where a zinc ion is co-ordinated to three Cys and one His residue. Due to

this strong co-ordination, the peptide has a well-established structure, which was solved by NMR and served as the starting structure for this simulation. Since the actual structure of the peptide is irrelevant for the calculation of the dielectric properties of this system, it suffices here to say that no unfolding or other structural reorganizations occur during the simulation.

The peptide itself has no net charge. The zinc ion which is co-ordinated to the centre of the peptide is regarded as being part of the ion-component of the system and thus contributes to the dielectric constant of the protein only *via* the cross-correlation term $\langle \mathbf{M}_P(0)\mathbf{J}_I(t)\rangle$, which is negligible in our case, as we show later. Using the formalism developed in the Appendix, we would have been able to handle proteins of arbitrary charge and thus consider the zinc ion as being part of the protein, which would then have a net charge of +2. However, studying the Appendix shows that this would have led to an almost identical result as our straightforward and simple approach (because the zinc ion stays close to the geometry centre of the peptide throughout the simulation).

Results and Discussion

To calculate the frequency-dependent dielectric constants of the water-component and the peptide as well as the frequency-dependent conductivity of the ion-component of the system we have to calculate, according to equations (27), (28) and (29), the generalized susceptibilities $\chi_{P_WP}(\omega)$, $\chi_{P_PP}(\omega)$, and $\chi_{\mathbf{i}_I P}(\omega)$. For these we need in turn the six possible collective correlation functions $\langle \mathbf{M}_W(0)\mathbf{M}_W(t)\rangle$, $\langle \mathbf{M}_W(0)\mathbf{M}_P(t)\rangle$, $\langle \mathbf{M}_W(0)\mathbf{J}_I(t)\rangle$, $\langle \mathbf{M}_P(0)\mathbf{M}_P(t)\rangle$, $\langle \mathbf{M}_P(0)\mathbf{J}_I(t)\rangle$, and $\langle \mathbf{J}_I(0)\mathbf{J}_I(t)\rangle$, as can be seen from equations (11), (13) and (15).

Figure 1 shows the auto-correlation function $\langle \mathbf{M}_W(0)\mathbf{M}_W(t)\rangle$ of the dipole moment of the water-component of the system and a bi-exponential fit to it. The bi-exponential fit $2139e^{-t/7} + 137e^{-t/159}$ is a very good approximation of this correlation func-

tion and helps us to identify two relaxation times for the dielectric relaxation of the water-component, namely $\tau_1 = 7$ ps and $\tau_2 = 159$ ps. τ_1 is similar to both the experimental relaxation time of water (9.3 ps) and the relaxation time determined by MD simulations of similar solutions (Yang *et al.*, 1995) and of pure water using a different water model (Neumann, 1986). τ_2 on the other hand is much larger and thus indicates a retardation effect in the dielectric relaxation of a portion of the water molecules. The respective weights of these two relaxation times (2139, or 94% for τ_1 compared to 137, or 6% for τ_2) together with the rather large value of τ_2 strongly suggest that τ_2 represents the relaxation times of those water molecules that are co-ordinated to either the peptide or the ions and thus are seriously restricted in their ability to rotate freely in response to an applied electric field.

The inset in Figure 1 just proves that the collective vibrations (known as librations) that are observed in the short-time region of the dipole auto-correlation function of pure water can also be seen in this simulation.

Figure 2 shows the auto-correlation function $\langle \mathbf{M}_P(0)\mathbf{M}_P(t)\rangle$ of the dipole moment of the peptide and a bi-exponential fit to it. The first aspect to note is the very slow relaxation of the peptide molecule: The simulation time of 13.1 ns gives rise to statistically reliable time-correlation functions of no more than 3 ns length. This length is not enough to follow the decay of $\langle \mathbf{M}_P(0)\mathbf{M}_P(t)\rangle$ to zero (as was easily achievable for $\langle \mathbf{M}_W(0)\mathbf{M}_W(t)\rangle$, for instance), but we are able to sample a good portion of this decay. The second observation is that a bi-exponential fit is not able to describe all the complex features of $\langle \mathbf{M}_P(0)\mathbf{M}_P(t)\rangle$. It is obvious that the dielectric relaxation of the peptide (or any molecule as complex as a peptide, for that matter) evolves around the many diverse relaxation channels of individual residues and groups of the molecule, a behaviour that cannot be fully described by a simple fit. However, the fit $94e^{-t/162} + 501e^{-t/4318}$ allows us to identify a fast initial decay ($\tau_1 = 162$ ps) of
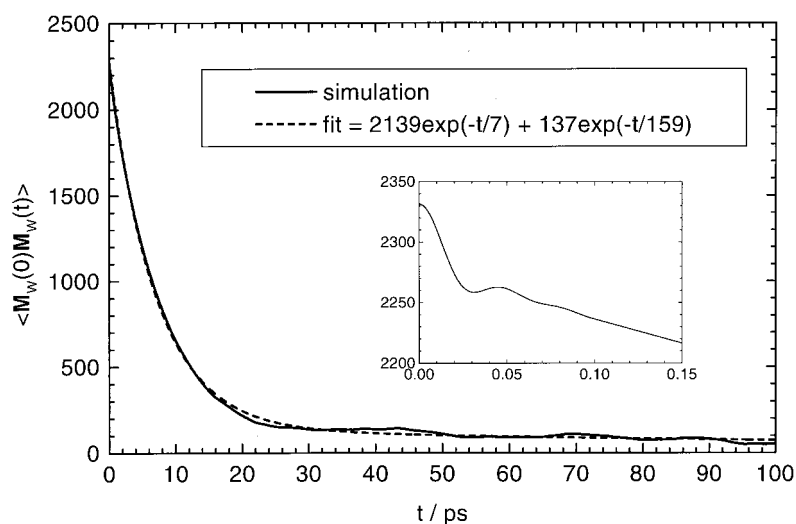

Figure 1. The auto-correlation function $\langle \mathbf{M}_W(0)\mathbf{M}_W(t)\rangle$ of the dipole moment of the water-component and a bi-exponential fit to it. The inset shows the librations in the short-term region of the correlation function.
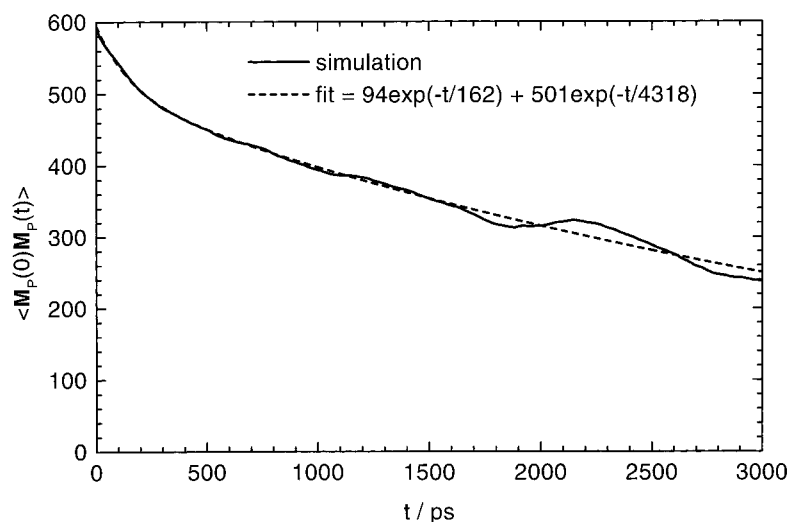
Figure 2. The auto-correlation function $\langle \mathbf{M}_P(0)\mathbf{M}_P(t) \rangle$ of the dipole moment of the peptide and a bi-exponential fit to it.

$\langle \mathbf{M}_P(0)\mathbf{M}_P(t) \rangle$ caused by the rapid relaxation of the most mobile groups of the peptide and an ensemble of slow decays collectively approximated by the correlation time $\tau_2 = 4318$ ps. The long relaxation time (4.318 ns) is a bit higher than the values (1.8 ns and 3.6 ns) found by Smith *et al.* (1993), which means that the dielectric relaxation of our (small) peptide is slower than the relaxation of the (large) proteins examined by Smith *et al.* (1993). The overall rotation of the peptide must be one of the dominant slow dielectric relaxation modes of the peptide. We also note that the short relaxation time of the peptide (162 ps) is very similar to the long relaxation time of the dipole moment correlation function of the water-component (159 ps), which would be the case if, for instance, the fastest relaxing residues of the peptide were on the surface of the peptide, where they would interact with water molecules in such a way that these water molecules exhibit the same relaxation time as the residues.

Figure 3 presents the cross-correlation function $\langle \mathbf{M}_W(0)\mathbf{M}_P(t) \rangle$ of the dipole moment of the water-component and the dipole moment of the peptide. This cross-correlation function is far from negligible, although its magnitude is considerably smaller than that of either $\langle \mathbf{M}_W(0)\mathbf{M}_W(t) \rangle$ or $\langle \mathbf{M}_P(0)\mathbf{M}_P(t) \rangle$. The cross-correlation function is very complex, however the reasonable bi-exponential fit $14e^{-t/18} + 151e^{-t/2876}$ identifies a relatively short relaxation time ($\tau_1 = 18$ ps) and a considerably longer one ($\tau_2 = 2876$ ps). All in all, this correlation function decays on the same time-scale as $\langle \mathbf{M}_P(0)\mathbf{M}_P(t) \rangle$, which means that its contribution to $\varepsilon_P(\omega)$, which is based on $\langle \mathbf{M}_P(0)\mathbf{M}_P(t) \rangle$, will be much stronger than its contribution to $\varepsilon_W(\omega)$. If we assume that the dipole moment of the water-component is the sum of the dipole moment of a slowly relaxing portion (being part of the hydration shell around peptide or ions) and the dipole moment of a fast relaxing portion (''free'' water), the cross-correlation function will be a sum
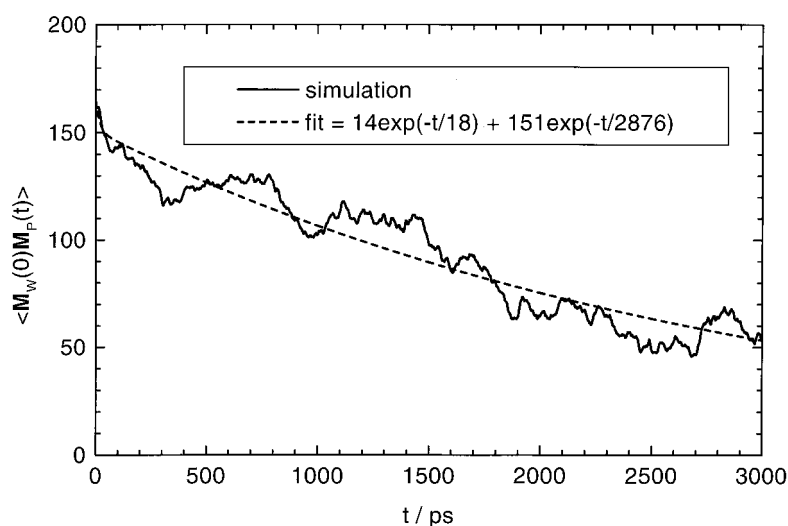

Figure 3. The cross-correlation function $\langle \mathbf{M}_W(0)\mathbf{M}_P(t) \rangle$ of the dipole moment of the water-component and the dipole moment of the peptide and a bi-exponential fit to it.
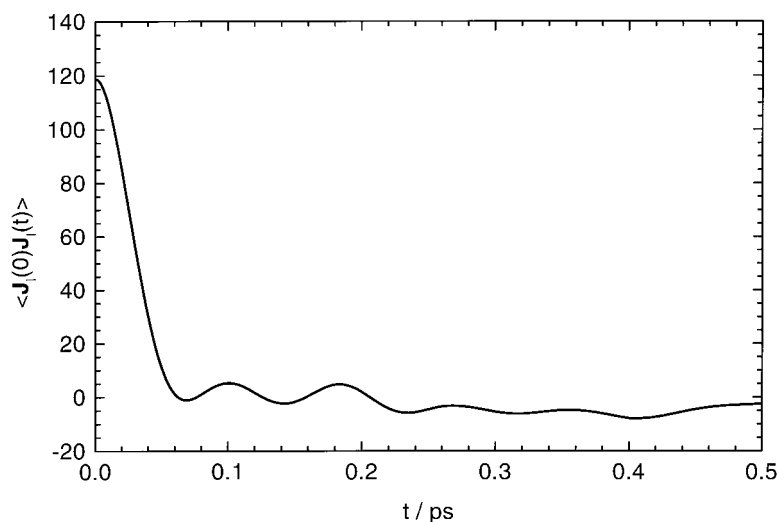
Figure 4. The auto-correlation function $\langle \mathbf{J}_I(0)\mathbf{J}_I(t) \rangle$ of the current of the ion-component.

of the cross-correlation functions between the dipole moment of each portion and the dipole moment of the peptide. The large value of τ_2 comes close to the longer relaxation time of $\langle \mathbf{M}_P(0)\mathbf{M}_P(t) \rangle$ (4318 ps) and thus suggests that the water molecules that are part of the hydration shell around the peptide or the ions are so tightly bound that the dielectric relaxation does not separate them readily from the peptide or the ions. On the other hand, the short relaxation time (18 ps) is very similar to the short relaxation time of the dipole moment correlation function of the water-component (7 ps).

Figure 4 shows the auto-correlation function $\langle \mathbf{J}_I(0)\mathbf{J}_I(t) \rangle$ of the current of the ion-component. As was observed in many other MD simulations (Caillol *et al.*, 1986, 1987, 1989a,b; Löffler *et al.*, 1997), this correlation function is extremely short-ranged. The oscillatory behaviour stems from ions captured in a co-ordination cage. The peptide, which co-ordinates the zinc ion with three Cys and one His residue, certainly plays the role of such a

cage. On the other hand, the two chloride ions are hydrated by water molecules, and thus are trapped inside a co-ordination cage, too, albeit not as perfectly as the zinc ion is trapped inside the peptide.

In the theory section we argued on theoretical grounds that the values for $t = 0$ of the cross-correlation functions $\langle \mathbf{M}_W(0)\mathbf{J}_I(t) \rangle$ and $\langle \mathbf{M}_P(0)\mathbf{J}_I(t) \rangle$ are zero. This is verified by Figures 5 and 6, which show these cross-correlation functions. Apart from the fact that the value for $t = 0$ is zero in these functions, they are very small at all other values of the frequency, too. On one hand this is caused by the small number of ions in the system. On the other hand, the low ion concentration and the great difference in the time regime in which the dielectric relaxation of the ions and that of the water-component or the peptide take place leads to a negligible coupling of these quantities. This decoupling of current and dipole moment is more perfect for the case of $\langle \mathbf{M}_W(0)\mathbf{J}_I(t) \rangle$ than for the case of $\langle \mathbf{M}_P(0)\mathbf{J}_I(t) \rangle$. The latter is obviously due to the
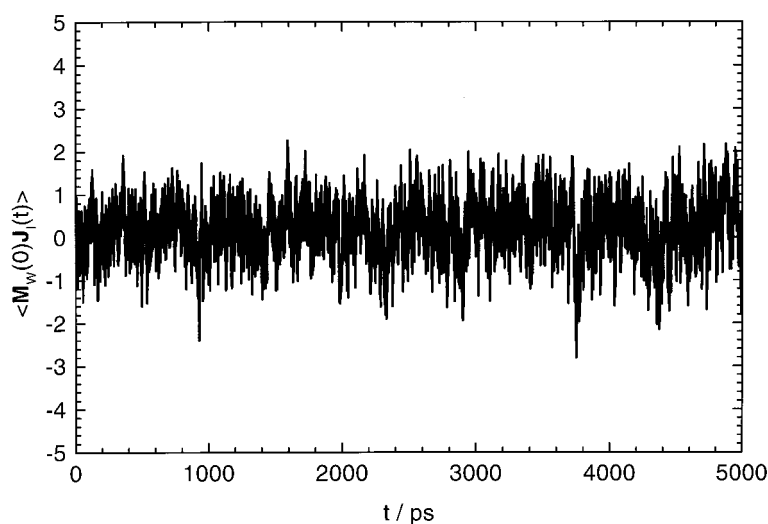

Figure 5. The cross-correlation function $\langle \mathbf{M}_W(0)\mathbf{J}_I(t) \rangle$ of the dipole moment of the water-component and the current of the ion-component.
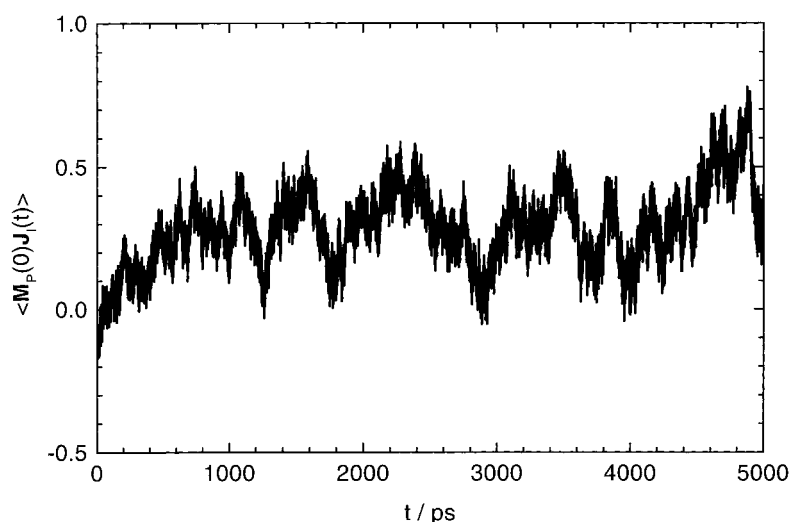
Figure 6. The cross-correlation function $\langle \mathbf{M}_P(0)\mathbf{J}_I(t)\rangle$ of the dipole moment of the peptide and the current of the ion-component.

fact that the zinc ion always moves together with the peptide, so there has to be some coupling of the current of the zinc ion to the dipole moment of the peptide at finite frequencies. This coupling can be seen in $\langle \mathbf{M}_P(0)\mathbf{J}_I(t)\rangle$, but since there is just one zinc ion, the magnitude of $\langle \mathbf{M}_P(0)\mathbf{J}_I(t)\rangle$ is still so small that it does not play any role in the calculation of the dielectric properties of the system.

The frequency-dependent dielectric constant $\varepsilon_W(\omega)$ of water is shown in Figure 7, both in the correct form and neglecting the cross-correlation function $\langle \mathbf{M}_W(0)\mathbf{M}_P(t)\rangle$. As can be seen from this comparison, omitting the cross-correlation function does not change the resulting dielectric constant considerably. This can be understood by realizing that the cross-correlation function decays much more slowly (see Figure 3) than the auto-correlation function $\langle \mathbf{M}_W(0)\mathbf{M}_W(t)\rangle$ of the dipole moment of the water-component (Figure 1), that otherwise determines $\varepsilon_W(\omega)$ according to equations (27) and (11). Consequently, the contribution of the cross-

correlation function is only important in the extreme low-frequency range (including the static value) of $\varepsilon_W(\omega)$, which results in a static dielectric constant of 47 *versus* 44 (6.4% too low), with and without the cross-correlation function, respectively. Apart from that, the static dielectric constant $\varepsilon_W(0)$ of water is with a value of 47 much lower than 71 for pure SPC/E water (Reddy & Berkowitz, 1989). Of course, this is related to the slower dielectric relaxation that was already observed in Figure 1.

The frequency-dependent dielectric constant $\varepsilon_P(\omega)$ of the peptide along with the $\varepsilon_P(\omega)$ that results if the cross-correlation function $\langle \mathbf{M}_W(0)\mathbf{M}_P(t)\rangle$ is neglected are shown in Figure 8. For the case of the peptide, the omission of the cross-correlation term has a major effect, which is a result of the fact that $\langle \mathbf{M}_P(0)\mathbf{M}_P(t)\rangle$ and $\langle \mathbf{M}_W(0)\mathbf{M}_P(t)\rangle$ decay at similar time scales (Figures 2 and 3 and equations (28) and (13)). The peptide shows two distinct dispersion regions: A relatively high-dielectric region up to approximately $10^{-4}\,\mathrm{ps}^{-1}$ ($10^2\,\mathrm{MHz}$) and a
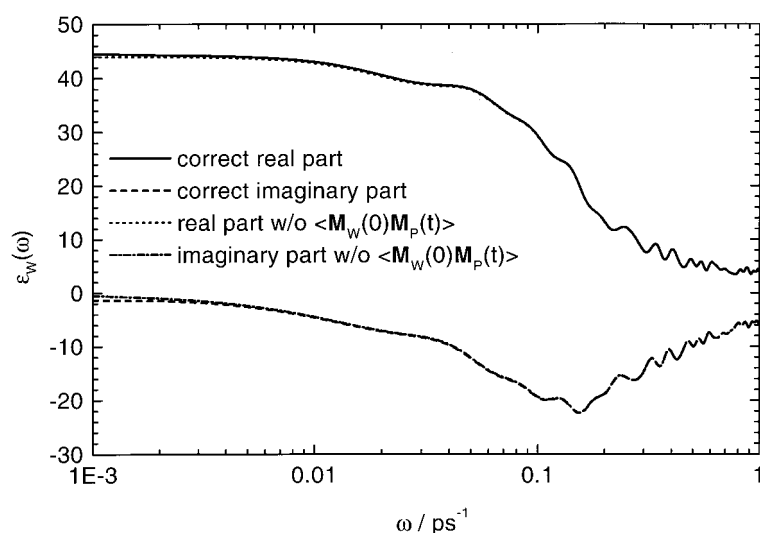

Figure 7. Real and imaginary part of the frequency dependent dielectric constant $\varepsilon_W(\omega)$ of the water-component. Also included in this Figure are the real and imaginary part of the $\varepsilon_W(\omega)$ that results if the cross-correlation term $\langle \mathbf{M}_W(0)\mathbf{M}_P(t)\rangle$ is neglected.
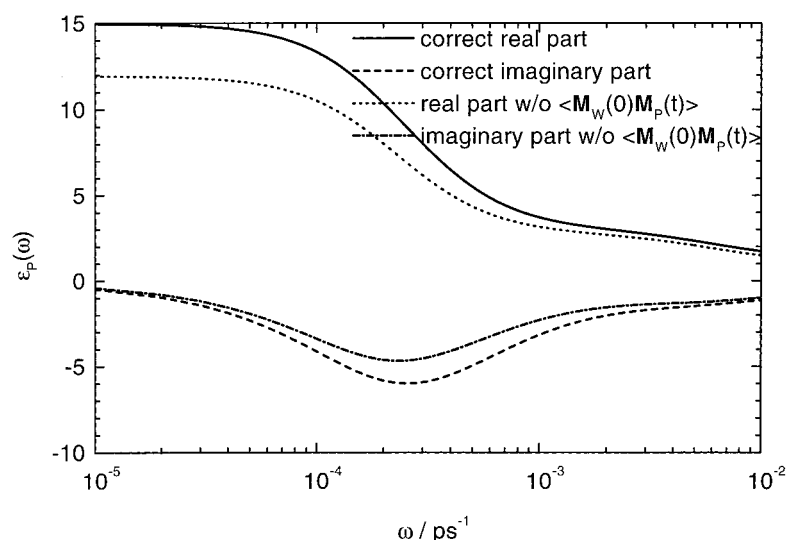
Figure 8. Real and imaginary part of the frequency-dependent dielectric constant $\varepsilon_P(\omega)$ of the peptide. Also included in this Figure are the real and imaginary part of the $\varepsilon_P(\omega)$ that results if the cross-correlation term $\langle \mathbf{M}_W(0)\mathbf{M}_P(t)\rangle$ is neglected.

low-dielectric region at frequencies higher than that. This is in general accordance with results from computer simulations by Smith *et al.* (1993), although the intensity of the overall dielectric response of the peptide is weaker in our case. This is even more noteworthy, since a small peptide like the one in our simulation is expected to be more flexible in its dielectric response than a big protein like in Smith *et al.*'s simulations. The static dielectric constant $\varepsilon_P(0)$ of the peptide is 15, which is considerably higher than what is traditionally used in solving the Poisson-Boltzmann equation (Gilson & Honig, 1987) but is in good accordance with several other computational studies (Nakamura *et al.*, 1988; King *et al.*, 1991; Simonson *et al.*, 1991; Antosiewicz *et al.*, 1994; Simonson & Perahia, 1995). It must be stressed that omitting the cross-correlation function $\langle \mathbf{M}_W(0)\mathbf{M}_P(t)\rangle$ as was done explicitly by several authors (Smith *et al.*, 1993; Yang *et al.*, 1995) (and implicitly by the rest) would result in a static dielectric constant of only 12 (20% too

low). It was noted by several authors who calculated the dielectric constant of a biomolecule from an MD simulation including explicit water molecules that the inclusion of water during the simulation phase is essential in obtaining a reasonable (reasonably high) value for the dielectric constant. While we certainly support this, we go one step further in saying that the inclusion of the coupling term between the dipole moment of the peptide and the dipole moment of the water-component is also fundamental to correctly describe the dielectric relaxation of a solvated peptide. Of course, being able to include this coupling term requires a dielectric theory, like the one developed in the theory section of this paper, that correctly describes the role the coupling term plays in determining $\varepsilon_P(\omega)$.

Figure 9 shows the frequency-dependent conductivity $\sigma_I(\omega)$ of the ion-component of the system. We see considerable static conductivity and a broad band of dielectric relaxation channels at the
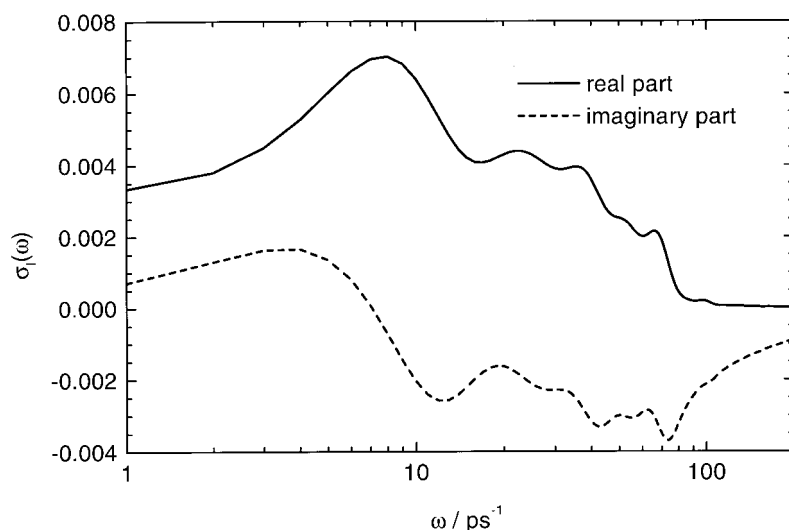

Figure 9. Real and imaginary part of the frequency dependent conductivity $\sigma_I(\omega)$ of the ion-component.

very high frequencies that are characteristic for conductivities of small ions (Löffler *et al.*, 1997).

Summary and Outlook

Based on linear response theory, the phenomenological equations of matter and a computer-adaptation of dielectric theory, ingredients that already proved successful in the case of one-component systems and ionic solutions, we were able to derive a theory of the frequency-dependent dielectric properties of systems consisting of a protein in an aqueous solution containing ions. This theory only works in the case of simulations done under periodic boundary conditions, which we consider a necessity, anyway. Using correlation functions determined by MD simulations, it is possible to calculate the frequency-dependent dielectric constant of the protein, the frequency-dependent dielectric constant of the water-component, and the frequency-dependent conductivity of the ions, which are all dependent on the concentration of the respective component in the system. Our theory requires not only the calculation of the auto-correlation functions of the dipole-moment or current of each component, respectively, but also the calculation of the cross-correlation functions between these properties. Indeed, the cross-correlation function between the dipole moments of the protein and the water-component proved important for the calculation of the dielectric constant of the protein.

Using this theory, we evaluated a 13.1 ns MD simulation of an 18-residue zinc finger peptide and three ions in aqueous solution, where the concentration of the peptide was 18.2 mmol 1^{-1}. The frequency-dependent dielectric properties of the three components of this system are given in the main part of this paper. The static dielectric constant of the peptide was found to be 15, which is in accordance to more recent investigations. Neglecting the coupling to the water-component would have underestimated the static dielectric constant by three units. The static dielectric constant of the water-component was found to be 45, compared to 71 for pure SPC/E water. The difference logically stems from the reduced mobility of those water molecules that are co-ordinated to the peptide or the ions.

This work presented a single (global) frequency-dependent dielectric function for each component in the system. Several authors (see Introduction) have calculated site-dependent (local) dielectric constants for proteins based on various approaches. A simple generalization of the theory presented in this paper allows the rigorous calculation of such local dielectric constants, including the necessary coupling terms between the dipole moments of the various sites, that are supposedly very important in such a case. Furthermore, this generalization need not be restricted to site-dependent dielectric constants of the peptide, but can also include local dielectric constants for the water-component. Specifically, it seems interesting to calculate separate dielectric constants for the first few hydration shells around the peptide, since their dielectric relaxation is bound to differ (as suggested by the results in this paper) from that of bulk water.

However, there is a caveat. While decomposition of linear response theory results is always possible down to any desired level of granularity, the respective constitutive relations are limited by the fact that each component must behave like a piece of dielectric matter. Of course this limit does not prohibit theoretical considerations concerning site-dependent (local) dielectric constants. But as granularity decreases the corresponding results become less and less related to those of a macroscopic dielectric.

Other questions that await further investigation concern the dependence of the dielectric properties on the concentration and volume fraction of each component. Furthermore, it seems promising to calculate the wave-vector dependent dielectric properties as outlined in Appendix I.

References

Antosiewicz, J., McCammon, J. A. & Gilson, M. K. (1994). Prediction of ph-dependent properties of proteins. *J. Mol. Biol.* **238**, 415–436.

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

Caillol, J. M. (1987). The dielectric constant and the conductivity of an electrolyte solution at finite wave-lengths and frequencies. *Europhys. Letters,* **4**, 159–166.

Caillol, J. M., Levesque, D. & Weis, J. J. (1986). Theoretical calculation of ionic solution properties. *J. Chem. Phys.* **85**, 6645–6657.

Caillol, J. M., Levesque, D., Weis, J. J., Kusalik, P. G. & Patey, G. N. (1987). A comparison between computer simulation and theoretical results for ionic solutions. *Mol. Phys.* **62**, 461–473.

Caillol, J. M., Levesque, D. & Weis, J. J. (1989a). Electrical properties of polarizable ionic solutions. I. theoretical aspects. *J. Chem. Phys.* **91**, 5544–5554.

Caillol, J. M., Levesque, D. & Weis, J. J. (1989b). Electrical properties of polarizable ionic solutions II. Computer simulation results. *J. Chem. Phys.* **91**, 5555–5566.

Darden, T., York, D. & Pedersen, L. (1993). Particle mesh ewald: an $n \cdot \log(n)$ method for ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092.

de Leeuw, S. W., Perram, J. W. & Smith, E. R. (1980a). Simulation of electrostatic systems in periodic boundary conditions. I. Lattice sums and dielectric constants. *Proc. Roy. Soc. ser. A*, **373**, 27–56.

de Leeuw, S. W., Perram, J. W. & Smith, E. R. (1980b). Simulation of electrostatic systems in periodic boundary conditions. II. Equivalence in boundary conditions. *Proc. Roy. Soc. ser A*, **373**, 57–66.

Gilson, M. K. & Honig, B. H. (1987). *Nature,* **330**, 84–86.

King, G., Lee, F. S. & Warshel, A. (1991). Microscopic simulations of macroscopic dielectric constants of solvated proteins. *J. Chem. Phys.* **95(6)**, 4366–4377.

Kohler, F. (1972). *The Liquid State,* chapt. 9.2, pp. 157–167, Verlag Chemie, Weinheim.

Löffler, G. & Maccallum, P. H. (1995). Molecular dynamics of solvated proteins on the Cray T3D: algorithms, implementation, benchmarks and results, ParCo95 Conference.

Löffler, G. & Schreiber, H. (1996). Parallel biomolecular simulation: an overview and analysis of important algorithms. In *High-Performance Computing and Networking: Lecture Notes in Computer Science* (Liddell, H., Colbrook, A., Hertzberger, B. & Sloot, P., eds), vol. 1067, pp. 929–934, Springer, Berlin, Heidelberg, New York.

Löffler, G., Schreiber, H. & Steinhauser, O. (1997). The frequency-dependent conductivity of a saturated solution of $ZnBr_2$ in water. A molecular dynamics simulation. *J. Chem. Phys.,* in print (scheduled for August 15, 1997).

Madden, P. & Kivelson, D. (1984). A consistent molecular treatment of dielectric phenomena. *Advan. Chem. Phys.* **56**, 467.

Nakamura, H., Sakamoto, T. & Wada, A. (1988). *Protein Eng.* **2**, 177–183.

Neumann, M. (1986). Dielectric relaxation in water computer simulations with the tip4p potential. *J. Chem. Phys.* **85**, 1567–1580.

Neumann, M. & Steinhauser, O. (1983a). On the calculation of the dielectric constant using the ewald-kornfeld tensor. *Chem. Phys. Letters,* **95**, 417–422.

Neumann, M. & Steinhauser, O. (1983b). On the calculation of the frequency-dependent dielectric constant in computer simulations. *Chem. Phys. Letters,* **102**, 508–513.

Neumann, M., Steinhauser, O. & Pawley, G. S. (1984). Consistent calculation of the static and frequency-dependent dielectric constant in computer simulations. *Mol. Phys.* **52**, 97–113.

Reddy, M. R. & Berkowitz, M. (1989). Dielectric constant of spc/e water. *Chem. Phys. Letters,* **155**, 173–176.

Simonson, T. & Perahia, D. (1995). Internal and interfacial dielectric properties of cytochrome *c* form molecular dynamics in aqueous solution. *Proc. Natl Acad. Sci. USA,* **92**, 1082–1086.

Simonson, T., Perahia, D. & Brünger, A. T. (1991). Microscopic theory of the dielectric properties of proteins. *Biophys. J.* **59**, 670–690.

Smith, P. E., Brunne, R. M., Mark, A. E. & van Gunsteren, W. F. (1993). Dielectric properties of trypsin inhibitor and lysozyme calculated from molecular dynamics simulations. *J. Phys. Chem.* **97**, 2009–2014.

Summers, M. F., South, T. L., Kim, B. & Hare, D. R. (1990). High-resolution structure of an HIV zinc fingerlike domain *via* a new nmr-based distance geometry approach. *Biochemistry,* **29**, 329–340.

Tironi, I. G., Sperb, R., Smith, P. E. & van Gunsteren, W. F. (1995). A generalized reaction field method for molecular dynamics simulations. *J. Chem. Phys.* **102**, 5451–5459.

Yang, L., Weerasinghe, S., Smith, P. E. & Pettitt, B. M. (1995). Dielectric response of triplex DNA in ionic solution from simulation. *Biophys. J.* **69**, 1519–1527.

Appendix I

In equations (1) and (3) we assumed the existence of a local, linear relationship between the frequency-components of the Maxwell field and the frequency-components of the polarization of the water-component and the protein, respectively. These equations were called phenomenological equations of matter or (local) constitutive relations. Of course, the dependency of the polarization on the Maxwell field is a function of the frequency-dependent dielectric constant of the water-component and the protein, respectively. Since it is not obvious at first sight that this dependency is a local, linear relationship, we shall briefly outline the more general theory in this Appendix.

In fact, the general expression for the dependency of the polarization somewhere in the system on the internal electric field (Maxwell field), the so-called non-local constitutive relation, is given by (Madden & Kivelson, 1984):

$$\mathbf{P}(\mathbf{r},\, \omega) = \int_V \frac{\varepsilon(\mathbf{r}-\mathbf{r}',\, \omega) - \delta(\mathbf{r}-\mathbf{r}')}{4\pi} \mathbf{E}(\mathbf{r}',\, \omega)\, d\mathbf{r}'$$

(AI1)

In other words, the frequency-component $\mathbf{P}(\mathbf{r}, \omega)$ of the polarization at any point $\mathbf{r}$ inside the system depends on the frequency-component $\mathbf{E}(\mathbf{r}', \omega)$ of the electric field at any other point $\mathbf{r}'$ inside the system and a generalized dielectric constant $\varepsilon(\mathbf{r}-\mathbf{r}', \omega)$ that itself depends on the distance vector $\mathbf{r}-\mathbf{r}'$. While the same arguments given here apply to the conductivity (cf. equation (4)) as well, we think it suffices to concentrate on the dielectric constants.

The ubiquitous assumption in this paper is that the dielectric properties of the system protein/water/ions can be described by two (generalised) dielectric constants, that is $\varepsilon_W(\mathbf{r}-\mathbf{r}', \omega)$ for the water-component and $\varepsilon_P(\mathbf{r}-\mathbf{r}', \omega)$ for the protein (we are omitting the conductivity of the ions here). Hence we separate equation (AI1) into a contribution $\mathbf{P}_W(\mathbf{r}, \omega)$ from the water and a contribution $\mathbf{P}_P(\mathbf{r}, \omega)$ from the protein:

$$\mathbf{P}(\mathbf{r},\, \omega) = \mathbf{P}_W(\mathbf{r},\, \omega) + \mathbf{P}_P(\mathbf{r},\, \omega) \qquad \text{(AI2)}$$

$$\mathbf{P}_X(\mathbf{r},\, \omega) = \int_V \frac{\varepsilon_X(\mathbf{r}-\mathbf{r}',\, \omega) - \delta(\mathbf{r}-\mathbf{r}')}{4\pi}$$
$$\times\, \mathbf{E}(\mathbf{r}',\, \omega)\, d\mathbf{r}', \quad X = W \text{ or } P$$

(AI3)

However, both the protein version and the water version of equation (AI3) employ the spatially resolved Maxwell field $\mathbf{E}(\mathbf{r}', \omega)$ over the entire system, since the polarization at a point, e.g. inside the protein naturally depends on the electric field outside the protein, too.

We shall now perform a spatial Fourier transformation of equation (AI3) going from distance vectors $\mathbf{r}$ in real space to wave vectors $\mathbf{k}$ in Fourier space. Note that this step is exact:

$$\frac{1}{V} \int_V \mathbf{P}_X(\mathbf{r}, \omega)\, e^{i\mathbf{kr}}\, d\mathbf{r} = \frac{1}{V} \int_V \int_V$$

$$\times \frac{\varepsilon_X(\mathbf{r} - \mathbf{r}', \omega) - \delta(\mathbf{r} - \mathbf{r}')}{4\pi}$$

$$\times \mathbf{E}(\mathbf{r}', \omega)\, e^{i\mathbf{kr}'}\, e^{i\mathbf{k}(\mathbf{r} - \mathbf{r}')}\, d\mathbf{r}'\, d\mathbf{r}$$

$$(AI4)$$

On the right-hand side we switch from $\mathbf{r}$ to $\mathbf{r}'' = \mathbf{r} - \mathbf{r}'$, which is, because of the periodic boundary conditions, always inside the volume of the simulation system (see earlier, where we have used the same argument):

$$\frac{1}{V} \int_V \mathbf{P}_X(\mathbf{r}, \omega)\, e^{i\mathbf{kr}}\, d\mathbf{r} = \frac{1}{V} \int_V \mathbf{E}(\mathbf{r}', \omega)\, e^{i\mathbf{kr}'}\, d\mathbf{r}' \int_V$$

$$\times \frac{\varepsilon_X(\mathbf{r}'', \omega) - \delta(\mathbf{r}'')}{4\pi}\, e^{i\mathbf{kr}''}\, d\mathbf{r}''$$

$$(AI5)$$

After introducing the wave vector dependent spatial Fourier transforms:

$$\tilde{\mathbf{P}}_X(\mathbf{k}, \omega) = \frac{1}{V} \int_V \mathbf{P}_X(\mathbf{r}, \omega)\, e^{i\mathbf{kr}}\, d\mathbf{r} \qquad (AI6)$$

$$\varepsilon_X(\mathbf{k}, \omega) - 1 = \int_V \varepsilon_X(\mathbf{r}'', \omega) - \delta(\mathbf{r}'')\, e^{i\mathbf{kr}''}\, d\mathbf{r}'' \quad (AI7)$$

and:

$$\tilde{\mathbf{E}}(\mathbf{k}, \omega) = \frac{1}{V} \int_V \mathbf{E}(\mathbf{r}', \omega)\, e^{i\mathbf{kr}'}\, d\mathbf{r}' \qquad (AI8)$$

of the familiar real space quantities we arrive at the local constitutive relation in Fourier space:

$$\tilde{\mathbf{P}}_X(\mathbf{k}, \omega) = \frac{\varepsilon_X(\mathbf{k}, \omega) - 1}{4\pi}\, \tilde{\mathbf{E}}(\mathbf{k}, \omega) \qquad (AI9)$$

Here, the dielectric properties are calculated on a level that is compatible with dielectric measurements, which means that we are considering the dielectric response to an externally applied, homogeneous electric field. The theory achieves this by relating the polarization as calculated by the constitutive relation to the polarization as calculated by linear response theory (see earlier).

Linear response theory too can be formulated in a general, wave vector dependent way, just like we have done it for the constitutive relation above. However, in the case of a homogeneous external field, which is the usual experimental situation that we want to describe, both linear response theory and the constitutive relation specialize to the case $\mathbf{k} = 0$. Thus in the case of a homogeneous external field equation (AI9) simplifies to:

$$\tilde{\mathbf{P}}_X(\mathbf{k} = 0, \omega) = \frac{\varepsilon_X(\mathbf{k} = 0, \omega) - 1}{4\pi}$$

$$\times \tilde{\mathbf{E}}(\mathbf{k} = 0, \omega), \quad X = W \text{ or } P. \quad (AI10)$$

As can easily be seen, equation (AI10) is identical with equations (1) and (3), respectively!

In the light of this derivation, equations (1) and (3) appear as the appropriate expressions for the theoretical treatment of the usual experimental situation of an homogeneous external electric field.

References

Madden, P. & Kivelson, D. (1984). A consistent molecular treatment of dielectric phenomena. *Advan. Chem. Phys.* **56**, 467.

Appendix II

In the main part of this paper we developed the theory of the calculation of the dielectric properties of an electrically neutral protein in an aqueous ionic solution. However, proteins usually carry an electric net charge. This Appendix presents a physically rigorous strategy for dealing with charged proteins in the context of the calculation of the dielectric properties.

Linear response theory tells us that the total dipole moment of the systems couples to the externally applied electric field that we apply to induce a dielectric response in the system. In our case, the total dipole moment $\mathbf{M}$ is made up of contributions from the water-component of the system ($\mathbf{M}_W$), the ions ($\mathbf{M}_I$), and the protein ($\mathbf{M}_P$). Since the system as a whole must be electrically neutral, $\mathbf{M}$ is independent of the origin of the coordinate system (in other words: $\mathbf{M}$ is physically reasonable). The same is true for $\mathbf{M}_W$, because every water molecule itself is electrically neutral. Consequently, the sum of the net charge of the protein and the net charge of the ions must vanish, so that $\mathbf{M}_I + \mathbf{M}_P$ is a physically reasonable quantity, too. If the protein has a net charge, both $\mathbf{M}_P$ and $\mathbf{M}_I$ depend on the origin of the coordinate system. This would not be a problem for the ion-component, because the theory section of the paper shows that we only use the time-derivative of $\mathbf{M}_I$, which is independent of the origin of the coordinate system even for a non-neutral change distribution. However, $\mathbf{M}_P$ is used directly for the calculation of the dielectric constant $\varepsilon_P(\omega)$ of the protein, so that for a protein bearing a net charge not only $\mathbf{M}_P$ but also $\varepsilon_P(\omega)$ would depend on the choice of the origin of the coordinate system, which is clearly not acceptable.

We are looking for a way to transform the individual dipole moments $\mathbf{M}_W$, $\mathbf{M}_I$, and $\mathbf{M}_P$ of the components of the system in a way that keeps the total dipole moment of the system constant. The total dipole moment must remain constant because it governs the physical response of the system as a whole to the external electric field. Please note that all "transformations" described here only affect the way the system is treated during the calculation of the dielectric properties of the system, after the MD simulation has been finished just as usual. A protein with an electrical net charge is basically a giant ion, albeit an ion whose dielectric relaxation is based considerably on orientational relaxation (which is measured by the dielectric constant) and not so much on simple translation (which is measured by the conductivity). So it is no surprise that if we are looking for a way to transform the individual dipole moments, we are striving for a re-assignment of charges between the protein and the ions:

$$\mathbf{M} = \mathbf{M}_W + \mathbf{M}_I + \mathbf{M}_P = \mathbf{M}_W + \mathbf{M}'_I + \mathbf{M}'_P \quad \text{(AII1)}$$

(We add a prime to the quantities that result from the transformation that we are developing here.)

In other words:

$$\sum_{i \in I} q_i \mathbf{r}_i + \sum_{j \in P} q_j \mathbf{r}_j = \sum_{i \in I'} q'_i \mathbf{r}_i + \sum_{j \in P'} q'_j \mathbf{r}_j \quad \text{(AII2)}$$

In writing this expression without primes attached to the position vectors, we already ruled out the possibility of changing the position of any particle because this is bound to introduce serious artifacts into our evaluation. What remains is changing the charges of particles and/or re-assigning particles from the set of the ions to the set of the protein atoms, or *vice versa*. Again, any changing or re-assigning of charges is done only for the purpose of the analysis of configurations of the system and by no means for the actual calculation of these configurations in an MD simulation!

The idea now is to encapsulate the net charge of the protein into a pseudo-ion, so that the remaining protein is electrically neutral. The pseudo-ion is then considered to be just another part of the ion-component of the system. The charges of all "real" ions remain unchanged and constitute the remaining part of the ion-component of the system. We propose the following straightforward algorithm

for separating the net charge $Q_P = \Sigma_{j \in P} q_j$ of the protein into a pseudo-ion: This algorithm distributes the negative net charge of the protein evenly over all N_P atoms of the protein:

$$q'_j = q_j - \frac{Q_P}{N_P} \quad \text{(AII3)}$$

This means a very small change in the charge of each atom of the protein which makes the protein electrically neutral. For instance, a net charge of $+2$ for a protein consisting of 2000 atoms means the subtraction of 0.001 from the charge of each protein atom. To compensate for the subtraction of Q_P, a pseudo-ion of the charge Q_P is created. Where this pseudo-ion has to be placed (its position $\mathbf{r}_{QP}$) can be calculated from the requirement that the total dipole moment of the system must not change through this transformation:

$$\mathbf{M}_W + \mathbf{M}_I + \mathbf{M}_P = \mathbf{M}_W + \mathbf{M}'_I + \mathbf{M}'_P \quad \text{(AII4)}$$

$$\sum_{i \in I} q_i \mathbf{r}_i + \sum_{j \in P} q_j \mathbf{r}_j = \sum_{i \in I} q_i \mathbf{r}_i + Q_P \mathbf{r}_{QP}$$
$$+ \sum_{j \in P} \left(q_j - \frac{Q_P}{N_P} \right) \mathbf{r}_j \quad \text{(AII5)}$$

$$\mathbf{r}_{QP} = \frac{1}{N_P} \sum_{j \in P} \mathbf{r}_j \quad \text{(AII6)}$$

In other words, the pseudo-ion has to be placed at the geometric centre of the protein!

This is in contrast to the technique of calculating the dipole moment of a charged protein by choosing the centre of mass of this protein as the origin of the coordinate system, as was done by Smith *et al.* (1993), and Yang *et al.* (1995).

References

Smith, P. E., Brunne, R. M., Mark, A. E. & van Gunsteren, W. F. (1993). Dielectric properties of trypsin inhibitor and lysozyme calculated from molecular dynamics simulations. *J. Phys. Chem.* **97**, 2009–2014.

Yang, L., Weerasinghe, S., Smith, P. E. & Pettitt, B. M. (1995). Dielectric response of triplex DNA in ionic solution from simulation. *Biophys. J.* **69**, 1519–1527.