Simulation of protonic fluctuations in hydrated protein powders

Giorgio Careri
Dipartimento di Fisica, Università di Roma "La Sapienza," INFM Unità Roma Uno, 00185 Roma, Italy

Edoardo Milotti
Dipartimento di Fisica, Università di Udine, Via delle Scienze, 208, I-33100 Udine, Italy
and INFN—Sezione di Trieste, Via delle Scienze, 208, I-33100 Udine, Italy
(Received 27 November 2002; published 23 May 2003)

Protons migrating on the surface of weakly hydrated protein powders provide a percolating mesoscopic system which exhibits charge fluctuations near room temperature. In this paper, we describe a simple numerical model where the statistical redistribution of protons on a space distribution of identical side chains lying on a spherical protein surface is varied by random ionization-recombination process to investigate the noise power spectrum of the fluctuating dipole moment, the ergodicity of this system, and the occurrence of localized or extended proton distributions. The case of lysozyme is considered to this end.

DOI: 10.1103/PhysRevE.67.051923 PACS number(s): 87.14.Ee, 05.40.Ca, 87.15.Aa

I. INTRODUCTION

Protons migrating on the surface of weakly hydrated protein powders are relevant not only for protein function, but also for providing a mesoscopic system with a dynamical behavior which exhibits cooperativity, as already reviewed elsewhere [1,2]. The proton fluctuations in the frequency range from 100 Hz to 1 MHz, studied in the laboratory [3], are especially important because they show that dielectric relaxation originates from proton displacements on percolating clusters of water molecules adsorbed on the protein surface. More recently [4], the power spectrum of the dipole moment fluctuations of a single macromolecule was evaluated from dielectric data, and the high-frequency tail was found to follow a $1/f^{\alpha}$ power law with α from 1.4 to 1.8. These results require that the proton population must be temperature independent, at least in the range from room temperature to the glass temperature observed near 266 K [5]. Of course, this is possible only in poorly hydrated powders, since here strong interactions between side chains are assisted by the two-dimensional (2D) water layer bound to the protein surface.

The ac part of the electric dipole moment of a protein is associated to different physical processes as it contains contributions from rotational, vibrational, and ''breathing'' modes of the molecule. A further contribution in water comes from the dynamics of free protons on the protein surface. In this paper, we shall investigate the statistical redistribution of P semiclassical protons on a variable space distribution of S, identical side chains lying on the protein surface, as a relevant contribution to the fluctuation spectrum of the electric dipole of the globular protein. This is done by following a kinetic method [6] based on a population of identical random walkers distributed over a network of connected sites. In this method, neighboring proton configurations are those that are reachable with a single-proton jump to a free site. Here, we do not follow the proton paths on the protein surface, instead we consider the proton configuration as a sort of lattice gas or Ising spin model, where the occupation number or the spin value represents the presence or the absence of a proton in a given capture site at a given pH value. The algorithm does not include the fluctuations due to migrating ions distributed around the ionizable side chains nor the fluctuations of the protein's backbone dipoles, since these processes are known [1,2] to contribute to the higher-frequency tail of the processes involving migrating protons. The occupation probabilities at each lattice site satisfy a set of coupled Langevin equations and, according to Ref. [6], this suggests that the power spectrum of this occupation probability has a power spectral density characterized by a white spectrum at low frequency and by a $1/f^{\alpha}$ power law at higher frequencies. Indeed, we show at the end that pure ionization and recombination processes in which protons are treated as uncharged particles and the ionized side chains as empty sites is sufficient to explain qualitatively the occurrence of the $1/f^{\alpha}$ power law.

Broadly speaking, we treat the hydrated protein surface as a protonic amorphous semiconductor, where protons are donated and trapped by ionizable side chains and move on the hydrogen-bonded network provided by the nearby water molecules of the percolating cluster. This picture stems from previous studies, and here, to simulate this dynamical scenario in view of a future quantum mechanical treatment, we assume also that each protonic charge is distributed on its ionizable side chain in an s state, thus moving from that side chain to the next empty one, thanks to the overlapping s states centered on these two side chains. We are not aware of theoretical studies of proton transfer between ionizable side chains in poorly hydrated protein surfaces, and therefore we take the decay length Λ as a free parameter of the numerical model and we use it to adjust the transfer rate of protons in the 2D water matrix, to explore the behavior of the model. As we show below, the usable range of Λ values is close to the size of a few water molecules, and Λ controls the extension of the proton distribution. This is consistent with *ab initio* path integral simulation studies by Marx *et al.* [7] on the nature of the hydrated excess proton in liquid water, where the charged defects become delocalized over several hydrogen bonds because of the quantum delocalization, and the rate of proton diffusion is determined by the thermally

induced hydrogen-bond breaking in the second solvation shell. On the basis of this theoretical picture, it seems reasonable to take a range of Λ values near $\Lambda = 1.0$ nm for the charged defect migrating in the water layer adsorbed on the protein surface considered in this paper. Notice also that we simulate ionizable side chains as identical geometrical sites only, with no affinity towards protons, while in water solution a strong *p*H dependence has been detected [8]. Moreover, again for computational convenience, a spherical model for the lysozyme macromolecule will be assumed, as already proposed by Smith *et al.* [9] in the time domain shorter than 1 ns in their molecular dynamics study, where they neglected the very low-frequency contribution arising from the migrating protons that we are interested in.

Section II describes the simulation algorithm, Sec. III puts forward some general considerations on the Monte Carlo (MC) simulation, Sec. IV defines the ergodicity measure that we use in this paper, Sec. V reports the results of the simulation runs, Sec. VI compares the numerical results with the actual measurements performed on weakly hydrated lysozyme, and we draw our conclusions in Sec. VII.

II. THE SIMULATION ALGORITHM

The simulation program starts by defining an array of S capture sites randomly uniformly distributed over the protein surface: the number of sites is fixed for a given simulation run, and for each run we define a set of m different site configurations. In order to avoid unnatural clusterings of sites, we set from the start a minimum site distance D_{min}, so that the distance between any two sites is greater than D_{min}. The value $D_{min} = 0.4$ nm has been chosen in this work, since Heringa and Argos [10] in their study of side chain cluster in proteins found that half of the clusters display residue pairs within 0.4–0.5 nm of each other. The artificial molecule is assumed to be fixed in space because the sample is a polycrystalline powder, and the electrical dipole moment receives no contribution from orientational motion. Afterwards the transition probabilities λ_{jk} (= probability that a proton at site k steps into site j) are calculated, according to the rule explained below. P protons ($S > P$) are distributed randomly over the S capture sites and are allowed to step into the other sites according to the following rules: (a) the total transition probability p_k per MC step that the proton at site k steps into anyone of the free sites is calculated for all protons; (b) the total transition probabilities p_k are normalized and the proton that actually executes the step is drawn from the resulting distribution; (c) the final site j is drawn randomly from the normalized distribution of the λ_{jk}'s that correspond to transitions to free sites; (d) the dipole moment of the charge distribution is computed along each coordinate axis and is stored on disk; the program returns to step (a) and loops until the maximum number of Monte Carlo steps is reached. The whole procedure is repeated again for another random configuration of capture sites, until the maximum number of capture site configurations m is reached.

Since the true Hamiltonian of this complex physical system is not known, we can only surmise the dependence of the transition probabilities λ_{jk} on the actual physical parameters of the system: here, we assume that the wave function of the protons localized in the capture sites are *s*-state hydrogenoid wave functions. Then, from the first-order perturbation theory, we know that the transition probability per unit time is proportional to the square of the matrix element of a perturbing Hamiltonian, and if we assume that this perturbing Hamiltonian is slowly varying and approximately position independent, each transition probability λ_{jk} is roughly proportional to the square of the overlap integral of the wave functions centered in j and k, and thus is a simple exponential $\lambda_{jk} \approx A \exp(-L_{jk}/\Lambda)$, where L_{jk} is the distance between the two sites, measured along the surface of the molecule. The spectral density of each projection of the dipole moment is computed and all the resulting spectral densities are averaged together: this makes sense, since the molecule is assumed to be ''frozen'' in space, as it is in polycrystalline hydrated powders, but the configuration is random and rotationally invariant; moreover, the measured spectral density is the average response of many such hydrated molecules.

It is important to notice that we do not know how the perturbing Hamiltonian compares to the localizing Hamiltonian, and therefore we do not know the actual transition rate; this means that we must introduce still another parameter in the simulation program, namely, the probability λ_0 that the proton configuration remains fixed. The probability λ_0 and the coefficient A set the time scale of the simulation.

The computed dipole moment is defined by the centroid of the occupied sites, and since we do not actually know the permanent dipole moment, nor are we interested to evaluate it, this means that the dc component of the spectral density has no physical relevance.

In the simulation procedure, we set a realistic molecule size from the start (molecular radius = 1.5 nm and $D_{min} = 0.4$ nm), and we fix the parameters $A = 1$ and $\lambda_0 = 10$ (taken together these parameters set the overall time scale of the simulation). Then the only remaining free parameters are S, P, and Λ, and therefore we must adjust these parameters until the results of the simulation program reproduce the experimental data, i.e., the slope of the observed spectral density α and the integrated dipole fluctuation x_Δ^2,

$$x_\Delta^2 = \int_{f_1}^{f_2} x^2(f)\,df \approx \int_{f_1}^{\infty} x^2(f)\,df, \tag{1}$$

where $x^2(f)$ is the average spectral density of the electric dipole fluctuations, and (f_1, f_2) is the frequency range (the physical meaning of x_Δ^2 is discussed in Ref. [4]). The spectral density in Eq. (1) may be well approximated by a power law $x^2(f) \approx A/f^\alpha$ (so that in practice the upper integration limit does not matter much, because at very high frequency, and for $\alpha > 1$, the spectral density vanishes and the integration may be extended—at least formally—to infinity). Then the integrated dipole fluctuation may be approximated as follows:

$$x_\Delta^2 \approx \int_{f_1}^{f_2} \frac{C}{f^\alpha}\,df = \frac{C}{1-\alpha} f^{1-\alpha}\Big|_{f_1}^{f_2} = \frac{C}{1-\alpha}\left(f_2^{1-\alpha} - f_1^{1-\alpha}\right)$$

$$\approx \frac{C}{\alpha - 1} f_1^{1-\alpha}, \tag{2}$$

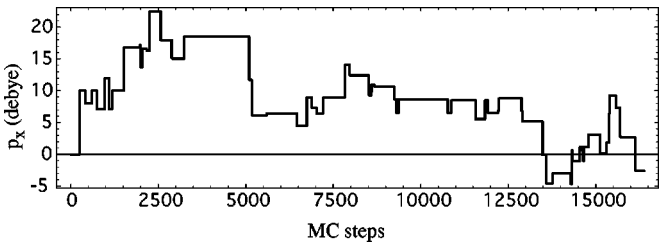

FIG. 1. This figure shows the time evolution of the *x* component of the electric dipole moment in one of the simulated molecules in run (40, 4, 1.0) (only the first 2^{14} steps are shown). The horizontal coordinate is time in Monte Carlo steps and the vertical coordinate is the dipole intensity in debyes. Notice that in this run the signal is visibly non-Gaussian.

and we see that it is very closely related to the C coefficient in the power law

$$\frac{C}{f_1^{\alpha-1}} \approx (\alpha-1)x_\Delta^2, \tag{3}$$

so that the quantity $C/(\alpha-1)$ can be used as a frequency-independent estimator of the integrated fluctuation in any given frequency band of the power-law fraction of the spectrum. We see also that if f_0 is the crossover frequency that marks the transition between the low-frequency white noise region and the high-frequency power-law region, then the actual integrated fluctuation in the power-law region is

$$x_\Delta^2(f_0) \approx \frac{C}{(\alpha-1)f_0^{\alpha-1}}, \tag{4}$$

and thus we take α, C, and f_0 as meaningful outputs from the simulation; we also compute $x_\Delta^2(0)$—the integrated fluctuation for all frequencies but dc—as an estimator of the total fluctuation (this includes the white noise plateau before the power-law region), and $x_\Delta^2(f_0)$. The program provides also some other interesting insights into the physics of the dipole fluctuation mechanism, in particular, it can be used to evaluate the ergodicity of the fluctuation process, as we show below.

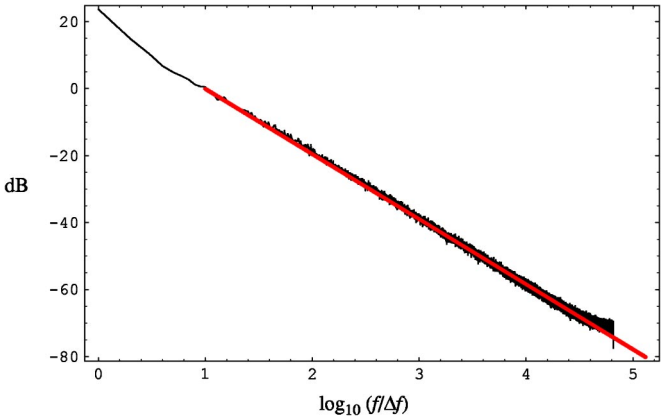

FIG. 2. Spectral density of the electrical dipole moment in run (40, 4, 1.0) obtained as the average of all the spectral densities of the *x*, *y*, and *z* components of the dipole moment in all the subruns. Δf is the frequency resolution in this simulation, and it is related to the Nyquist frequency f_N and to the number of samples N by the formula $\Delta f = 2f_N/N$; here $N = 2^{17}$. The thick line at the center shows a $1/f^\alpha$ fit which corresponds to $\alpha = 1.945$.

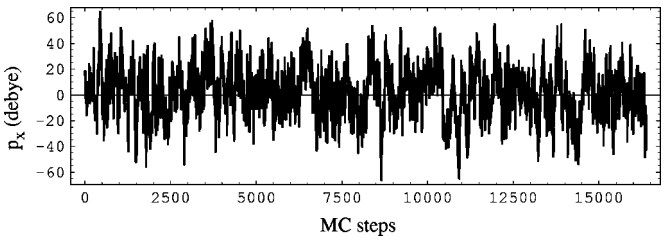

FIG. 3. This figure shows the time evolution of the *x* component of the electric dipole moment in one of the simulated molecules in run (40, 8, 5.0) (only the first 2^{14} steps are shown). The horizontal coordinate is time in Monte Carlo steps and the vertical coordinate is the dipole intensity in debyes. Notice that in this run the signal is very nearly Gaussian, in sharp contrast with the signal shown in Fig. 1, and this is confirmed by careful statistical analysis of the dipole signal.

III. GENERAL CONSIDERATIONS AND ESTIMATES

The structure of the MC program suggests some simple scaling laws and straightforward relations between the MC data, that we wish to discuss in this section before moving on to the analysis of the MC data themselves.

If the protons were independent and any two protons could visit the same site at the same time, then the spectral density would be proportional to the number of protons P, i.e., we should expect that $x_\Delta^2 \propto P$. This is not so, because even though protons do not have long range interactions, they behave like a lattice gas and cannot occupy the same site at the same time; therefore the ratio x_Δ^2/P, or also

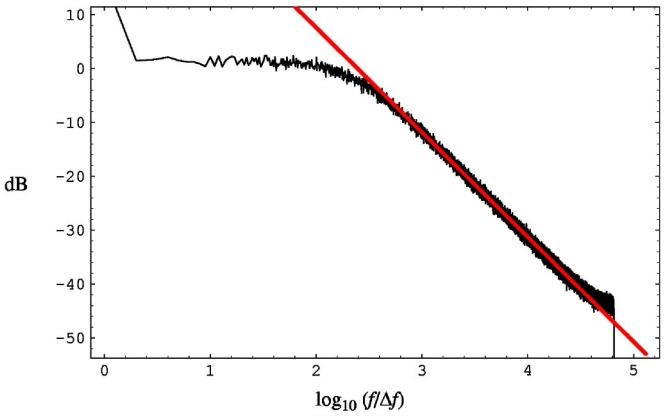

FIG. 4. Spectral density of the electrical dipole moment in run (40, 8, 5.0) obtained as the average of all the spectral densities of the *x*, *y*, and *z* components of the dipole moment in all the subruns. The thick line at the center shows a $1/f^\alpha$ fit which corresponds to $\alpha = 1.944$. Compare this figure with Fig. 2, and notice that in this run the spectral density flattens at low frequency; this means that the lowest transition rate is higher than the frequency resolution Δf of this plot.

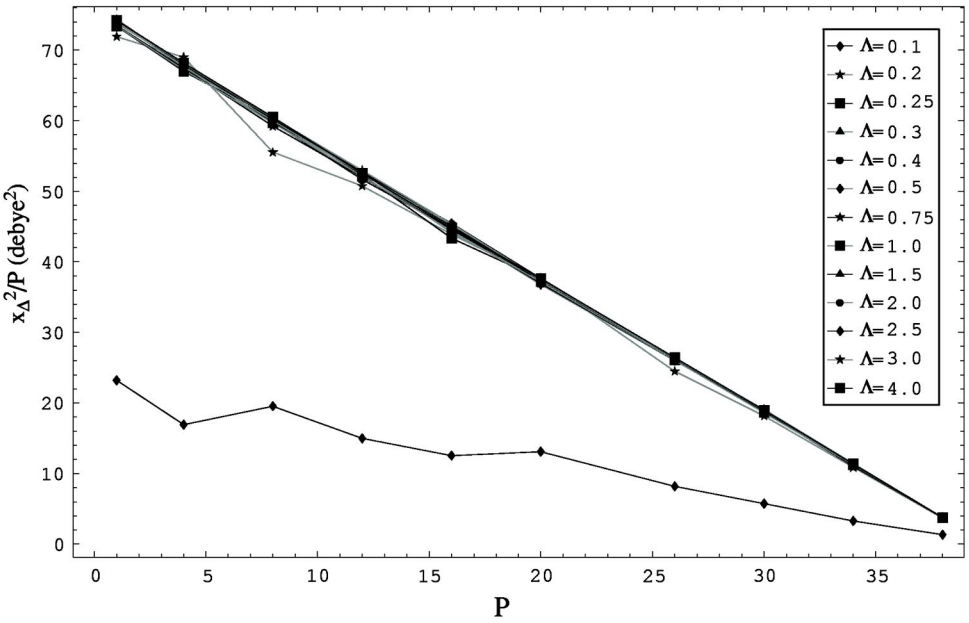

FIG. 5. Plot of $x_\Delta^2(0)/P$ vs P for different Λ's. The figure shows that the ac part of the total fluctuation behaves in a purely combinatorial way, as discussed in Sec. III. The inset shows the symbols used for different Λ values (in nanometer). Only the $\Lambda=0.1$ nm data deviate from the expected behavior: this is because the transition rates are too small and the number of MC events is insufficient; all the other MC data are quite satisfactory.

$x_\Delta^2(P)/x_\Delta^2(P=1)$ indicates the effective number of protons enacting the charge rearrangement on the surface of the molecule. Notice that if we assume that all transitions are equally likely and protons act in uncorrelated way, then the number of transitions per unit time should be proportional to $P(S-P)$ and therefore $x_\Delta^2 \propto P(S-P)$, a simple combinatorial rule.

An important feature of the experimentally observed $1/f^\alpha$ spectrum is that it spans several orders of magnitude, as is often the case with $1/f^\alpha$ spectra. The span of the power-law region in the Monte Carlo simulation can be estimated as follows: if we assume that the distribution of transition probabilities λ_{ij} is randomly sampled by the population of protons, then the total range of the power-law region is determined by $\min(\lambda_{ij})$ and $\max(\lambda_{ij})$. The value of $\max(\lambda_{ij})$ is determined by the parameter $D_{\min}$, while the value of $\min(\lambda_{ij})$ is determined by the radius of the molecule and by Λ, and has a weaker dependence on $D_{\min}$; the crossover frequency f_0 defined above is just the average of $\min(\lambda_{ij})$ calculated in the ensemble of simulated site configurations. The average transition probability is found as follows: since the total surface of the idealized spherical molecule is $4\pi R_0^2$, where R_0 is the radius of the spherical molecule, there is on average an area $4\pi R_0^2/S$ per capture site, which corresponds to a solid angle $4\pi/S=2\pi(1-\cos\theta)$ and therefore to a circle on the sphere with (angular) radius $\theta=\arccos(1-2/S)$, and nearest-neighbor distance $\rho \approx \theta R_0 = R_0 \arccos(1-2/S)$, so that the average transition probability is roughly proportional to $\exp(-\rho/\Lambda)$; these are only very rough approximations, but they help in estimating the correctness of the procedure.

Finally, we note that in the case of the simple exponential relaxation (Debye) following a sharp pulse, the average relaxation time can be estimated from the half width at half maximum of the spectral density, which in this case is a straightforward Lorentzian curve, as it is well known [13]. In our case the relaxation process is more complicated, instead of a single Lorentzian we have a superposition of Lorentzian curves, but at low frequency—below the lowest transition rate—the spectrum can again be well approximated by a single Lorentzian and at very low frequency this flattens out to an almost white spectrum. We define an average relaxation time τ by ignoring all the complexities of the spectrum and by fitting a single Lorentzian curve. In the previous paragraph, we have seen how the average transition probability may be estimated neglecting all proton interactions: the average relaxation time that we have just defined is actually a much better estimator of the average transition probability, inasmuch as it takes into account the excluded volume interaction between protons.

IV. ERGODICITY ESTIMATES

Sometimes protons are trapped in clusters of sites very close to one other and jump only occasionally to sites that are further away: in such cases, the behavior of the system is not ergodic for times shorter than the escape time. In order to study this nonergodicity, we have introduced a simple measure, similar to the one in Ref. [14]. Let $\mathbf{p}(n)=\{p_x(n),p_y(n),p_z(n)\}$ be the electric dipole vector at the nth step, and let

$$\overline{\mathbf{p}}=\frac{1}{N}\sum_{k=1}^{k=N}\mathbf{p}(k) \tag{5}$$

be the average value of $\mathbf{p}$ during one simulation run, where N is the total number of samples in the run, then we define the distance

$$d(n)=\left|\frac{1}{n}\sum_{k=1}^{k=n}\mathbf{p}(k)-\overline{\mathbf{p}}\right| \tag{6}$$

between the partial average $(1/n)\Sigma_{k=1}^{k=n}\mathbf{p}(k)$ and the global average $\overline{\mathbf{p}}$. If the simulation starts from a given value $\mathbf{p}_1$ then

$$d(n)=\left|\frac{1}{n}\left(\mathbf{p}_1+\sum_{k=2}^{k=n}\mathbf{p}(k)\right)-\overline{\mathbf{p}}\right|$$

$$=\left|\frac{1}{n}(\mathbf{p}_1-\overline{\mathbf{p}})+\frac{1}{n}\left(\sum_{k=2}^{k=n}\mathbf{p}(k)-\overline{\mathbf{p}}\right)\right|. \tag{7}$$

Successive values of $s(n)=(1/n)\Sigma_{k=2}^{k=n}\mathbf{p}(k)$ are highly correlated since the sum represents a random walk in dipole moment space: if we assume that $\mathbf{p}(n)$ is a white noise sequence, i.e., $\langle[\mathbf{p}(n)-\overline{\mathbf{p}}]\cdot[\mathbf{p}(m)-\overline{\mathbf{p}}]\rangle=\sigma_p^2\delta_{n,m}$, then the average distance is

$$\overline{d}(n)=\left\langle\left|\frac{1}{n}(\mathbf{p}_1-\overline{\mathbf{p}})+\frac{1}{n}\left(\sum_{k=2}^{k=n}\mathbf{p}(k)-\overline{\mathbf{p}}\right)\right|\right\rangle \tag{8}$$

$$=\frac{1}{n}\left\langle\sqrt{|\mathbf{p}_1-\overline{\mathbf{p}}|^2+2(\mathbf{p}_1-\overline{\mathbf{p}})\cdot\left(\sum_{k=2}^{k=n}\mathbf{p}(k)-\overline{\mathbf{p}}\right)+\left|\sum_{k=2}^{k=n}\mathbf{p}(k)-\overline{\mathbf{p}}\right|^2}\right\rangle \tag{9}$$

$$\approx\frac{1}{n}\sqrt{\left\langle|\mathbf{p}_1-\overline{\mathbf{p}}|^2+2(\mathbf{p}_1-\overline{\mathbf{p}})\cdot\left(\sum_{k=2}^{k=n}\mathbf{p}(k)-\overline{\mathbf{p}}\right)+\left|\sum_{k=2}^{k=n}\mathbf{p}(k)-\overline{\mathbf{p}}\right|^2\right\rangle} \tag{10}$$

$$=\frac{1}{n}\sqrt{\langle|\mathbf{p}_1-\overline{\mathbf{p}}|^2\rangle+\left\langle\left|\sum_{k=2}^{k=n}\mathbf{p}(k)-\overline{\mathbf{p}}\right|^2\right\rangle} \tag{11}$$

$$=\frac{1}{n}\sqrt{\langle|\mathbf{p}_1-\overline{\mathbf{p}}|^2\rangle+(n-1)\sigma_p^2} \tag{12}$$

$$\approx\sqrt{\frac{1}{n^2}\langle|\mathbf{p}_1-\overline{\mathbf{p}}|^2\rangle+\frac{1}{n}\sigma_p^2}, \tag{13}$$

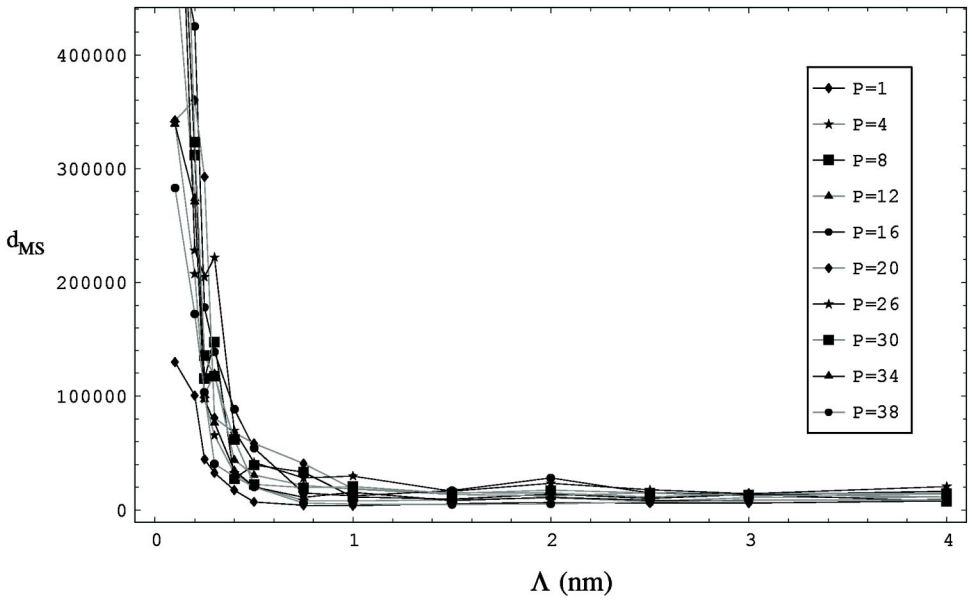

FIG. 6. Plot of d_{MS} vs Λ for different P's. The data gather around a common curve, and the ergodicity indicator d_{MS} falls off sharply at $\Lambda\approx 0.4$ nm.

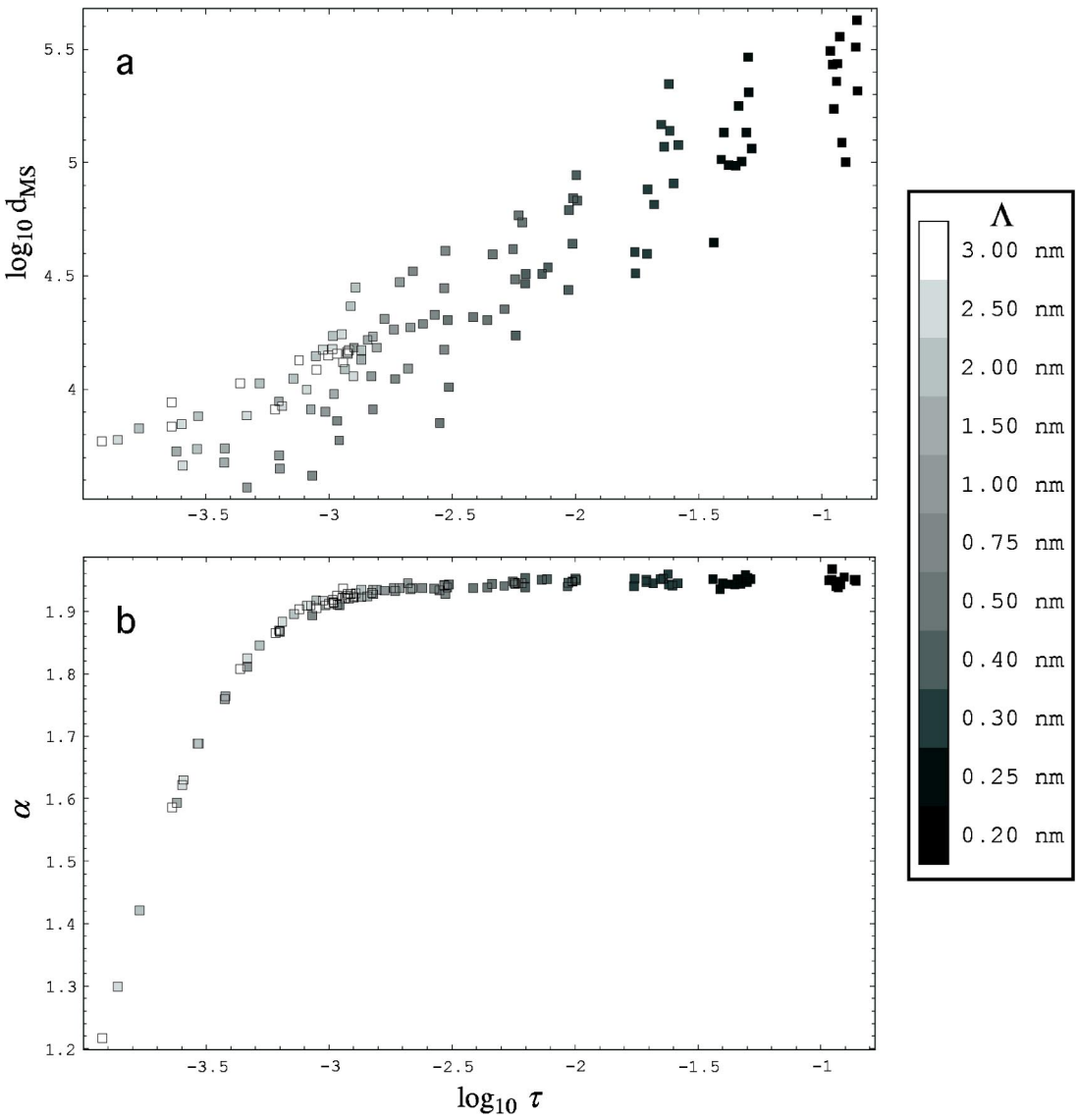

FIG. 7. This figure shows (a) plot of the logarithm of the ergodicity indicator $\log_{10} d_{MS}$ vs the logarithm of the average relaxation time $\log_{10}\tau$ (discussed in Sec. III) and (b) plot of α vs $\log_{10}\tau$; the figure includes all the data discussed in Sec. VI. Each data point is filled with a gray level that identifies the corresponding value of Λ (gray level scale on the right): notice that data points at different Λ's cluster in different regions of the two plots, and that there is a clear correlation among the three quantities.

clear deviations from average (13) should indicate an intrinsic nonergodicity. In the following, we use the square deviation from the average,

$$d_{MS} = \sum_{k=1}^{k=N} \langle [d(n) - \overline{d}(n)]^2 \rangle, \tag{14}$$

as a measure of nonergodicity.

V. RESULTS OF THE NUMERICAL SIMULATION

We have performed an extensive series of simulation runs to explore the parameter space; in each run ten different random site configurations have been generated, and for each site configuration we have taken 2^{17} Monte Carlo steps, so that each run corresponds to more than $2^{20}(\approx 10^6)$ Monte Carlo steps. As already explained above, the parameters in each run are (1) the number of sites S; (2) the number of protons P; (3) the exponential decay length Λ; (4) and (5) the constants A and λ_0 (these last two parameters are closely related), but to limit the complexity of the search we have decided to concentrate on the first three parameters, after the first trials showed that the spectral slope α was only very weakly dependent, if at all, on A and λ. We have labeled the runs by their (S,P,Λ) values, except for a few runs where we have also changed A and λ_0. The x, y, and z dipole signals are recorded in a file, and at the end they can be analyzed with the usual analysis tools: in particular, they are tested for Gaussianity and they are used to compute the final dipole spectral density. Some of the signals happen to be strongly non-Gaussian; this is so because transitions are rare and the dipole changes correspond to sharp steps. When tran-

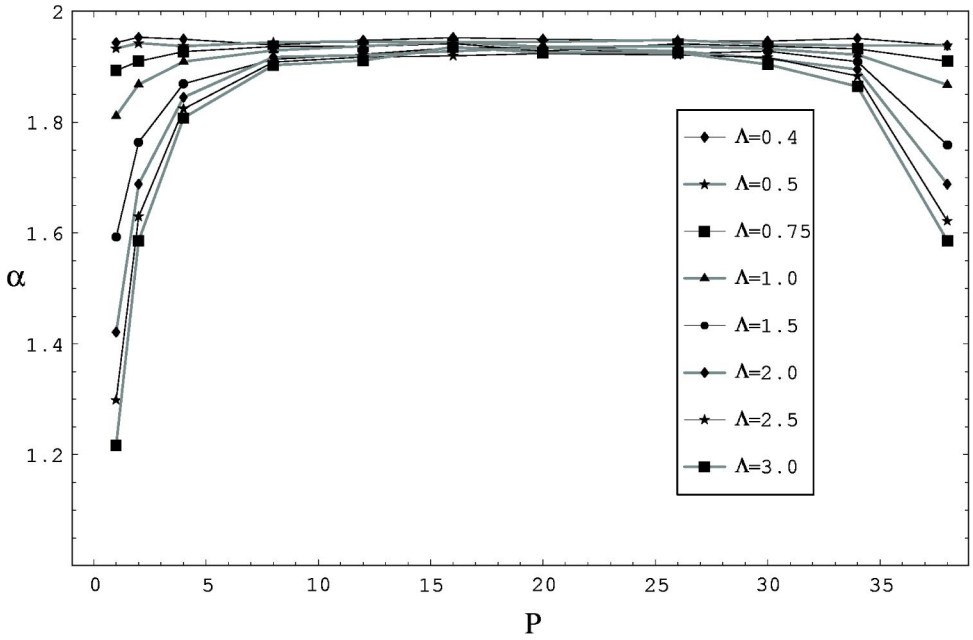

FIG. 8. Plot of α vs P for different Λ's. The inset shows the symbols used for different Λ values (in nanometer). The apparent left-right asymmetry is due to the uneven distribution of the values of P sampled in the simulation runs.

sitions are more frequent the signals have a smoother shape and the statistics is Gaussian, and when this happens the resulting spectral densities are meaningful also at the microscopic level. Indeed, the spectral density fully characterizes the statistics of a stochastic process only if it is Gaussian (see, e.g., Ref. [11]), and there are examples of interesting stochastic processes which have a $1/f$ power spectral density, while all the other statistics are very different from those observed in the laboratory (such as the process of Procaccia and Schuster described in Ref. [12]). At the macroscopic level all the dipole signals are always Gaussian because they result from the superposition of processes with finite variance and the central limit theorem applies.

Figures 1–4 show some of the signals and statistics obtained in some runs. Figures 2 and 4 show two spectral densities computed from the simulation data with a straightforward, unwindowed fast Fourier transform: the slope change at low frequency is due to the expected spectral whitening at very low frequency, while the slope change at high frequency is due to a slight amount of aliasing (in order to avoid alias-

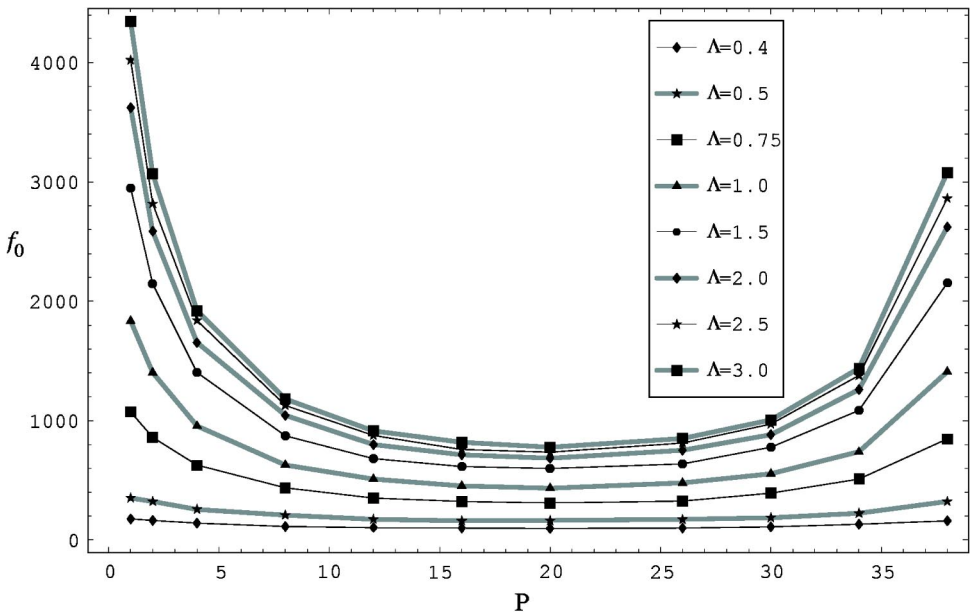

FIG. 9. Plot of the crossover frequency f_0 (arbitrary units) vs P. The inset shows the symbols used for different Λ values (in nanometers). As discussed in the text, at the end of Sec. II, f_0 corresponds to the lowest transition rate, i.e., to the longest proton path in a site-to-site jump, measured in units of Λ.

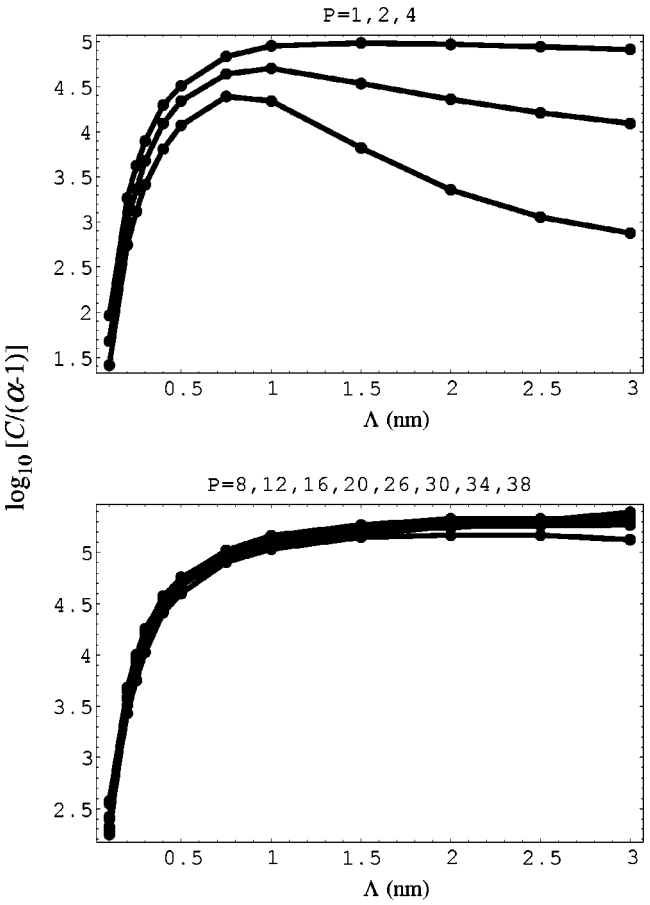

FIG. 10. Plot of the frequency-independent spectral density estimator $C/(\alpha-1)$ vs Λ for different values of P (see the discussion at the end of Sec. II). The curves bend upward and flatten for larger values of P.

ing, we oversample the electric dipole signal; the procedure is equivalent to a low pass filter, and what is left is a residual aliasing that affects very weakly only the highest frequencies and appears sometimes in the spectra as a slight upward bend of the high-frequency tail). The fits shown in the figures are not influenced by these different whitening trends because spectral information at very low and at very high frequencies is discarded before the fit.

VI. SIMULATION OF THE LYSOZYME MACROMOLECULE

To simulate the lysozyme macromolecule with the numerical model described in the previous sections, we set $S=40$ and $1\leqslant P\leqslant 39$. Since we assume that the protons bind to all sites in the same way and we neglect direct interactions among the protons themselves, the physical constants are factored out of the model and the transition probabilities are determined up to an undefined constant which incorporates all of the underlying physics. The number of protons P depends on the pH of the of the water solution before lyophilization, and in the lysozyme samples that have been studied experimentally so far $P=17$ at $p\mathrm{H}=10$ and $P=27$ at $p\mathrm{H}=3$ (see Ref. [15]). The protons interact with one another only indirectly, through an excluded volume effect (a proton cannot jump to a site that is already occupied by another proton): though indirect this interaction is quite powerful and modifies significantly the straightforward combinatorial behavior that expect from the simple considerations in Sec. III. However, this modification shows up only in the power-law region and we see in Fig. 5 that the integrated dipole fluctuation $x_0^2(0)$ behaves just as expected on simple combinatorial grounds; we use this property of the system to assess the quality of the Monte Carlo simulation, and since we find that only the set of $\Lambda=0.1$ nm runs display large deviations, we discard them from all the other analyses and future considerations (this is due to the extremely low values of the transition probabilities which would require a much larger number of Monte Carlo steps).

The numerical model lends itself to tests and measurements that would be awkward or impossible in the laboratory: in particular, it is possible to determine the ergodicity of the dynamics of the electric dipole moment in its configuration space at different values of P and Λ. The ergodicity measure d_{MS} is plotted in Fig. 6, and there we see that the data cluster around a common curve, i.e., they are independent of P, and that the ergodicity measure d_{MS} falls off sharply and ergodicity sets in only for values of $\Lambda>0.4$ nm. The nonergodicity occurs at large average relaxation times τ and corresponds to large values of the spectral index α (see Fig. 7): this is easily understood, because when Λ is very short each proton remains strongly localized in the vicinity of the original site, the electric dipole moment displays only occasionally large fluctuations, and repeated transitions between neighboring sites lead to a very small variance of the relaxation time, so that on the whole the dipole fluctuations roughly reproduce a simple Debye behavior.

The spectral index is markedly smaller than 2, and thus the system deviates from the simple Debye behavior, for large values of Λ and extreme values of P (either close to 1 or close to 40) as shown in Figs. 8 and 7(b). Lower values of α usually imply a wider distribution of relaxation times [16], and a likely explanation may be that for larger values of Λ the distribution of relaxation times is indeed wider, except when there are many protons that reduce the actual number of allowed single-proton jumps at each Monte Carlo step. A similar explanation holds when P is close to S, but applied to holes instead of protons.

We conclude by saying that two different space distributions of protons are possible, one localized at small Λ and one extended at large Λ, with a smooth change between the two when Λ corresponds to the size of about three water molecules.

The simulation program provides a wealth of other data, e.g., it yields both the crossover frequency f_0 that marks the transition between the white noise plateau and the power-law region (see Fig. 9) and the frequency-independent estimator of the $1/f$ like fluctuations [Fig. 10; see also Eqs. (3) and (4) and the related discussion at the end of Sec. II].

VII. CONCLUSIONS

In this paper, we have introduced a semiclassical kinetic model of the migration of protons on the surface of an ide-

alized protein molecule. We have studied in some detail the case $S=40$, because this is the best value to model the lysozyme macromolecule. We found that ergodicity holds for any $\Lambda \geqslant 0.4$ nm (about the size of one H_2O molecule) for any P. This is quite reasonable, since for small Λ each proton keeps close to its original site and the system evolves slowly and by jerks (from time to time, a proton suddenly jumps to a faraway site). We have also investigated the fluctuations of the electrical dipole moment, and we have found that the spectral index α is mostly insensitive to P in a wide range of Λ values, and it is very close to the Debye value $\alpha=2$, as expected for a process involving a single relaxation time. However, near $\Lambda=1$ nm the spectral index displays a change to lower values (especially at small or large P), pointing to a transition from a localized to a nonlocalized space distribution of protons; it is noteworthy that this value of Λ corresponds approximately to the size of three water molecules, and this is close to the size of the hydrated protonic defect in bulk water evaluated theoretically by Marx *et al.* [7]. Thus, each proton becomes actually delocalized over the whole protein surface for large values of Λ, unless there are many protons: in that case the proton populations acts together as a whole and the spectral index increases again near two. When the number of protons increases still more and nearly reaches S, the single particle behavior appears again, as the role of holes and protons is reversed.

We have seen that the $1/f^{\alpha}$ noise observed in hydrated protein powders in the 1 KHz–1 MHz frequency range [1,2] can be explained by a purely statistical model of the ionization-recombination processes. Power-law noises always appear in systems with memory and some form of scale invariance: in this case, the scale invariance is associated to the distribution of capture sites and thus to the distribution of transition rates, while the memory is provided by the excluded volume effect. Thus, the occurrence of power-law noise in these systems is largely independent from local microscopic features and forces, and depends only on statistics. This, we believe, is the major result of this work.

One may wonder whether the shape of the probability density of the dipole signal has a biological relevance, however, we believe that it is premature to derive conclusions of biological interest from this work. The occurrence of two different states with different proton distributions, localized and nonlocalized, should be relevant in enzymology because they also display different fluctuating dipole moment and they entail a different proton contribution to the total polarizability, and it is known that the environmental polarizability is crucial in the electron transfer reactions catalyzed by enzymes [17]. Nearly dry samples of proteins embedded in membranes should be considered as meaningful examples.

[1] G. Careri, Prog. Biophys. Mol. Biol. **70**, 223 (1998).
[2] M. Peyrard and G. Careri, Cell Biol. **47**, 745 (2001).
[3] G. Careri and G. Consolini (unpublished).
[4] G. Careri and G. Consolini, Phys. Rev. E **62**, 4454 (2000).
[5] A. Levstik *et al.*, Phys. Rev. E **60**, 7604 (1999).
[6] E. Milotti, Phys. Rev. E **51**, 3087 (1995).
[7] D. Marx, M.E. Tuckerman, J. Hutter, and M. Parrinello, Nature (London) **397**, 601 (1999).
[8] D.D.F. Shiao and J.M. Sturtevant, Biopolymers **15**, 1201 (1976).
[9] R.E. Smith *et al.*, J. Phys. Chem. **97**, 2009 (1993).
[10] J. Heringa and P. Argos, J. Mol. Biol. **220**, 151 (1991).
[11] W.R. Bennett, Proc. IRE **44**, 609 (1956); reprinted in M.S. Gupta, *Electrical Noise: Fundamentals and Sources* (IEEE Press, New York, 1977), p. 65.
[12] I. Procaccia and H. Schuster, Phys. Rev. A **28**, 1210 (1983).
[13] H. Fröhlich, *Theory of Dielectrics: Dielectric Constant and Dielectric Loss*, 2nd ed. (Clarendon Press, Oxford, 1958).
[14] J.E. Straub and D. Thirumalai, Proc. Natl. Acad. Sci. U.S.A. **90**, 809 (1993).
[15] G. Careri, M. Geraci, A. Giansanti, and J.A. Rupley, Proc. Natl. Acad. Sci. U.S.A. **82**, 5342 (1985).
[16] G. Careri, G. Consolini, Z. Kutnjak, C. Filipic, and A. Levstik, Phys. Rev. E **64**, 052901 (2001).
[17] R.A. Marcus and N. Sutin, Biophys. Biochem. Acta **811**, 265 (1985).