Energy condensation and dipole alignment in protein dynamics

Alexander Tenenbaum ⬤[*]

Physics Department, Sapienza University, Piazzale Aldo Moro 5, 00185 Roma, Italy

The possibility that distant biomolecules in a cell interact via electromagnetic (e.m.) radiation was proposed many years ago to explain the high rate of encounters of partners in some enzymatic reactions. The results of two recent experiments designed to test the propensity of protein bovine serum albumin (BSA) to interact via e.m. radiation with other proteins were interpreted in a theoretical framework based on three main assumptions: (i) in order to experience this kind of interaction the protein must be in an out-of-equilibrium state; (ii) in this state there is a condensation of energy in low-frequency vibrational modes; and (iii) the hydration layers of water around the protein sustain the energy condensation. In the present paper we present the results of molecular dynamics simulations of BSA in four states: at equilibrium and out-of-equilibrium in water, and at room and high temperature in vacuum. By comparing physical properties of the system in the four states, our simulations provide a qualitative and quantitative assessment of the three assumptions on which the theoretical framework is based. Our results confirm the assumptions of the theoretical model showing energy condensation at low frequency and electretlike alignment between the protein's and the water's dipoles; they also allow a quantitative estimate of the contribution of the out-of-equilibrium state and of the water to the observed behavior of the protein. In particular, it has been found that in the out-of-equilibrium state the amplitude of the oscillation of the protein's dipole moment greatly increases, thereby enhancing a possible absorption or emission of e.m. radiation. The analysis of BSA's dynamics outlined in the present paper provides a procedure for checking the propensity of a biomolecule to interact via e.m. radiation with its biochemical partners.

I. INTRODUCTION

Many years ago Fröhlich proposed a quantum model of electromagnetic (e.m.) interaction among biomolecules with the aim of explaining the high speed of some enzymatic reactions [1–4]. In Fröhlich's model, molecular partners in such a reaction underwent a selective, long-range attractive e.m. interaction. This raised the frequency of their encounters, and thus the average speed of the reaction; this selective interaction was assumed to be produced by resonant dipole oscillations and to be sustained by an energy condensation in the lowest-frequency vibrational mode of the molecules. In recent years, Fröhlich's proposition has been revisited, his model revised in a classical version, and amended [5–8]. In particular, it has been shown that the energy condensation necessary to establish the resonant interaction among dipoles required the protein to be in an out-of-equilibrium state produced by a continuous energy intake from the environment [7].

Recently, a team led by Professor Pettini has combined two sets of experiments with a classical elaboration of the original

quantum model proposed by Fröhlich and has interpreted the experimental data in this new EMI (electromagnetic interaction) theoretical framework [8,9]. The experimental evidence of EMI is an absorption peak at 0.314 THz in a solution of bovine serum albumin (hereafter BSA) proteins in water. This peak has been held to prove the protein's ability to be excited in a low-frequency vibrational mode suitable for the absorption/emission of e.m. radiation. The interpretation by means of the new model is based on three assumptions: (i) In order to emit or absorb an e.m. signal in a frequency range of low absorption by the water surrounding a protein, the latter must be in an out-of-equilibrium state produced by the absorption of energy from an external source and its dissipation into the solvent. (ii) In this out-of-equilibrium state, the energy distribution among vibrational modes of the protein leads to an energy condensation in the mode of lowest frequency. (iii) The first hydration layers of the water solvent play an essential role in the process, coupling the protein's high-frequency modes with its low-frequency phonons.

In the experimental setup that we used as a model for our computer simulation, a solution of BSA molecules in water is subjected to an energy input provided by a laser beam. This light is absorbed by fluorochromes attached to lysine residues of the protein; the fluorochromes then emit a large fraction of the absorbed energy. The difference between the energies of one absorbed and one emitted photon is 0.198 eV; these are the elementary energy inputs which, accumulating in time, push the protein in an out-of-equilibrium state. It is believed that this energy input is funneled into the

lowest-frequency vibrational modes through an internal cascade of rapid structural motions referred to as protein quakes [10–12]. This deviation from the energy equipartition among vibrational modes is ascribed to a synchronization of the accordionlike modes of the 30 α-helices that compose the structure of BSA [9].

In this paper we report on an atomistic simulation of BSA aimed at imitating the experimental conditions of the above-mentioned experiments in order to check the three assumptions on which the new theoretical model is based. We performed molecular dynamics simulations of a BSA protein entailing 9220 atoms in four states: at equilibrium ($T = 300$ K) and out-of-equilibrium (solvent at $T = 300$ K, protein at $T = 600$ K) in water, and at room ($T = 300$ K) and at high temperature ($T = 600$ K) in vacuum. By comparing physical properties of the system in the four states, our simulations provide a qualitative and quantitative assessment of the three assumptions on which the EMI theoretical framework is based.

II. SYSTEM MODELING

A. Energy input

One quantitative feature of the revised EMI model can be derived from the experiments already carried out, namely, the energy input rate needed to keep the protein in an out-of-equilibrium state producing the EMI. In the experiment, the energy provided to the BSA protein originates in the absorption of photons of wavelength $\lambda = 488$ nm by fluorochromes attached to the protein (five fluorochromes per protein on average) [8]. The fluorochromes reemit part of the absorbed energy so that the energy actually transferred to the protein per absorbed photon is 3.18×10^{-20} J. It has been estimated that in the experimental setting in which the argon laser operates at 500 µW each protein receives between 120 and 300 photons per second, which amounts to a power input of between 3.82×10^{-18} and 9.54×10^{-18} J/s [8]. One can thus evaluate the amount of energy pumped into a single protein before the lowest-frequency vibrational modes are excited and the absorption of low-frequency radiation sets in. The time lag before the onset of this absorption, as measured in the experiment, ranges between 5 and 10 minutes approximately [8]. The total energy absorbed by a single protein during this time may thus be between 1.15×10^{-15} J (5 minutes, lowest power input) and 5.73×10^{-15} J (10 minutes, highest power input).

Our computer simulations are designed to reproduce the condition expected to hold in the revised, classical version of the Fröhlich model. In order to mimic the continuous input of energy into the protein population of the experiment in a viable computational setting, we simulated a system with a single BSA molecule. The solvent is thermalized by a thermostat at 300 K, but the protein is in contact with a thermostat at 600 K, which provides the desired energy input rate.

The GROMACS simulation program [13–18] allows the computation of the energy injected into the protein by the 600-K thermostat while the solvent is kept at 300 K by the other thermostat. In a 100-ns simulation this energy is about

8.38×10^{-13} J, a factor of between 146 and 728 times the energy absorbed in the experiment by a single protein before the onset of the low-frequency absorption. Thus, the length of the trajectory should be sufficient to trigger the condensation of energy as the protein reaches the lowest threshold of energy injection (1.15×10^{-15} J) after about 140 ps and exceeds the highest threshold (5.73×10^{-15} J) after about 685 ps. The amount of energy of 3.18×10^{-20} J transferred to the protein per absorbed photon, reported above, might be overestimated and could possibly only be a fraction of about 10% of this [19]. If this were the case, the time needed in our simulation to trigger the condensation of energy would be 10 times shorter.

B. Simulation layout

At and above the temperature of 300 K, the trajectory of the system in its phase space is well represented in the framework of classical dynamics. The classical simulations presented in this paper, with their acronyms, are as follows: (a) An equilibrium simulation in which the protein and the solvent are in contact with a thermostat at 300 K: BSa. (b) An out-of-equilibrium simulation in which the solvent is in contact with a thermostat at 300 K, but the protein is in contact with a thermostat at 600 K: BSb. (d) A simulation of the protein in vacuum at 300 K: BSd. (e) A simulation of the protein in vacuum at 600 K: BSe. (f) A simulation of pure water in the same conditions, i.e., number of molecules, temperature, pressure, as in BSa and BSb: H2O.

Trajectory (a) provides the reference levels for the quantities computed in the other cases. Trajectory (b) mimics, as explained above, the reference experimental setting. Trajectories (d) and (e) are designed to reveal the role of the (absent) solvent in the dynamics of the protein. Trajectory (f) is designed to reveal the role of the (absent) protein in the dynamics of the solvent.

In the simulations BSb and BSe the protein is at 600 K, that is, above its denaturation temperature. However, even if its secondary structure is strongly reduced, the overall configuration of the protein is still rather compact. The protein state at 600 K may thus be considered as a premolten globule whose average gyration radius (R_g) can be compared to that of the protein at 300 K. The average number of intramolecular H bonds of all trajectories is given in Table I.

The simulations were performed with the GROMACS atomistic simulation package and the VMD program for the visualization of the molecule [20]. The all-atom OPLS-AA force field was used [21]. For the systems simulated in explicit solvent, the SPC/E water model was used [22]. The technical details of the simulation methods are given in Appendix A 1 [23–28].

III. SIMULATION RESULTS

We report those results of the simulation that are relevant for the analysis of the EMI model. All results given below (with the exception mentioned in Sec. IV B) were computed after the rototranslational motion of the protein was removed from the recorded trajectories. The time evolution of the protein's temperature, as well as the system's kinetic energy, potential energy, and pressure, was monitored for all 100-ns

TABLE I. R_g of BSA in various states. First column: System and trajectory. Second column: Temperature of the protein. Third column: Average number of intramolecular H bonds. Fourth column: Average gyration radius R_g. Fifth column: Standard deviation of $R_g(t)$. Sixth column: Energy condensation parameter C.

System	T (K)	H bonds	R_g (nm)	s.d. R_g (nm)	C(%)
BSa	300	438	2.70	0.018	11.9
BSb	600	385	2.89	0.059	17.7
BSd	300	698	2.50	0.012	6.8
BSe	600	600	2.44	0.019	8.3

trajectories. All these curves show a regular behavior and have stable fluctuations with amplitudes appropriate to the temperature of each trajectory.

A. Time evolution: Gyration radius

The gyration radius $R_g(t)$ of the protein was computed for all trajectories of BSA. The results in solvent and in vacuum are shown in Fig. 1. The average value and the standard deviation (s.d.) of each of the 100-ns trajectories in solvent and in vacuum are given in Table I.

The R_g value of BSa is in line with the analogous result of a previous molecular dynamics (MD) simulation of BSA, performed in a solvent similar (but not identical) to the one used here [29]. Table I shows that trajectories of BSA in solvent produce a larger R_g and fluctuation than trajectories in vacuum at the same temperature. Moreover, $R_g(t)$ in trajectory BSb is characterized by a larger fluctuation than in trajectory BSa and shows a large and steep decrease between $t = 55$ ns and $t = 62$ ns. In the previous MD simulation of BSA, a negative jump in $R_g(t)$ of similar depth and duration is reported at about $t = 375$ ns in the trajectory of the protein in the excited state [29]. This time greatly exceeds any reasonable equilibration time for such a system. Therefore the jump in our simulation

of BSb can be ascribed to an intrinsic instability of the protein rather than to an equilibration that is too short.

The average values of the protein's R_g for BSd and BSe show a thermal compaction when the temperature is raised. This phenomenon is attributable to entropic and structural factors [30]. In order to check whether this phenomenon also takes place in solvent, the R_g of BSa must be compared with that of a trajectory of BSA at equilibrium with a solvent at 600 K. This trajectory was computed over 100 ns, yielding $R_g = 2.54$ nm, which is smaller than $R_g = 2.70$ nm of BSa at 300 K. This confirms the propensity of BSA to undergo thermal compaction.

B. Time evolution: Dipole moment

GROMACS allows the separate computation of the dipole moment $\boldsymbol{M}(t)$ of the protein, of the water, and of the whole system (protein + water + counterions). This moment was computed for all 100-ns trajectories. The results for the protein in solvent and in vacuum are shown in Fig. 2. The average value of M and its s.d. in each 100-ns trajectory are given in Table II, where we also report the values for the water and the system. In the same table we report the data of a 100-ns run of pure water (H2O) in the same conditions as in BSa and BSb (26 401 water molecules, $T = 300$ K, and $P = 1$ bar).
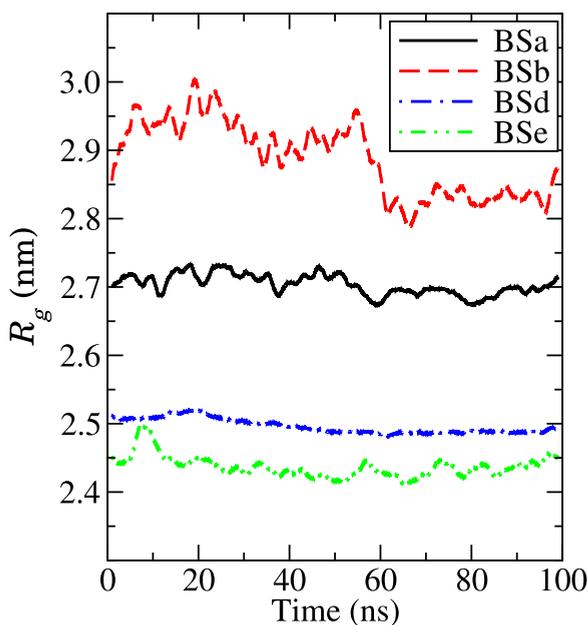

FIG. 1. Time evolution of the gyration radius $R_g(t)$ of protein BSA in four trajectories. The points are running averages over 2 ns.
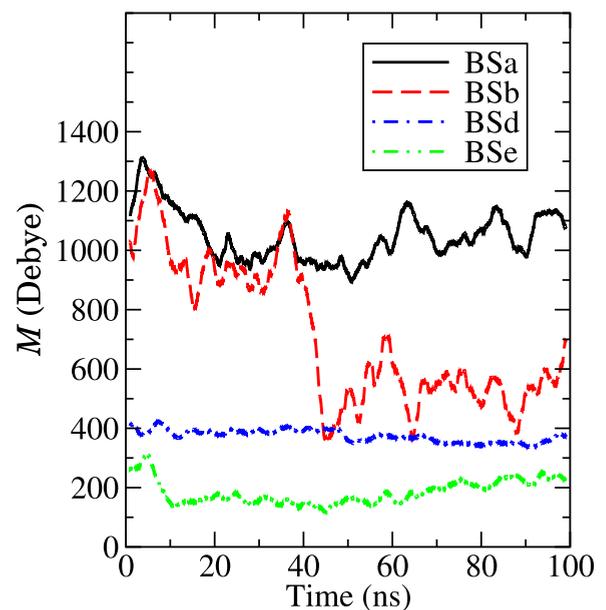

FIG. 2. Time evolution of the dipole moment $M(t)$ of protein BSA in four trajectories. The points are running averages over 2 ns.

TABLE II. Dipole moment of BSA in various states. First column: System and trajectory. Second column: Average dipole moment M (Debye). Third column: Standard deviation of $M(t)$ (Debye). Fourth column: Energy condensation parameter C. Fifth column: Dipole alignment parameter A.

System	M (D)	s.d. (D)	$C(\%)$	$A(\%)$
BSa (sys)	1516	411	2.1	43
BSa (prot)	1051	100	11.0	
BSa (H2O)	757	311	1.2	
BSb (sys)	1103	431	2.9	27
BSb (prot)	719	256	15.3	
BSb (H2O)	684	290	0.9	
BSd (prot)	375	27	6.0	
BSe (prot)	185	53	6.3	
H2O	688	288	0.7	

The average M value of BSa(prot) is in line, within one s.d., with the analogous result of the previous MD simulation of BSA cited above [29]. Our simulation produces two noteworthy results: (i) trajectories in solvent produce a much larger M and fluctuation for the protein compared with trajectories in vacuum at the same temperature, and (ii) in BSb $M(t)$ exhibits a large, steep decrease between $t = 37$ ns and $t = 40$ ns. The amplitude of this step lowers its values from the range of the trajectories in solvent to the range of the trajectories in vacuum and makes the average value of the protein's dipole in BSb significantly lower than in BSa, which is in line with the MD simulation of BSA previously mentioned [29]. No major variation of $R_g(t)$ takes place in conjunction with this large negative step of $M(t)$. This suggests that the large shift in $M(t)$ is due to a repositioning of charges of light mass, that is, H atoms.

If $M(t)$ of the protein and $M(t)$ of the water molecules were completely uncorrelated, the average angle between them would be $\pi/2$ and the dipole moment M of the system would be the Pythagorean sum of the dipole moments of the two subsystems. On the other hand, if the two dipoles were aligned and synchronized, the dipole moment M of the system would be the sum of the two dipole moments M. We introduced the parameter A to measure the alignment between the protein's dipole and the dipole of the water in each trajectory. Each system's actual dipole lies at a percentage A of the range between complete lack of correlation and complete alignment. This parameter allows a quantitative estimate of the extent of an oriented-dipole electret in the water layers surrounding the protein (see Sec. IV B) and is reported in Table II for the simulated systems.

If in BSa the two dipoles were completely uncorrelated, the dipole moment M of the system would be 1295 Debye (D). The sum of the two dipole moments M is 1808 D. The actual dipole moment of BSa(sys) (1516 D) is thus at $A = 43\%$ of the range between complete lack of correlation and complete alignment; this is the highest alignment achieved among all simulations and goes with the highest value of the protein's dipole moment (1051 D).

If in BSb the two dipoles were completely uncorrelated, the dipole moment M of the system would be 992 D. The sum of the two dipole moments M is 1403 D. The actual dipole of BSb(sys) (1103 D) is thus at $A = 27\%$. Shifting from BSa to BSb results in a loss of alignment in the total trajectory.

Notwithstanding the decrease of M when the temperature of BSA is increased there is no decrease in the amplitude of its oscillation. In contrast, when measured by the s.d., it increases by a factor 2.6 from BSa (prot) to BSb (prot), which is definitely larger than a factor of the order of $2^{1/2}$, which would be expected in harmonic approximation when the temperature is doubled. This data points to the active role of the out-of-equilibrium state in enhancing the dipole's fluctuation of the protein. Equally significant is the increase by a factor of 3.7 in the protein's dipole s.d. at 300 K in the solvent [BSa (prot)] compared to the same quantity at 300 K in vacuum (BSd) and the increase by a factor of 4.8 in the protein's dipole s.d. at 600 K in the solvent [BSb (prot)] compared to the same quantity at 600 K in vacuum (BSe). These last data highlight the active role of the solvent in enhancing the dipole's fluctuation of the protein.

As for the water, the average M of BSa(H2O) is larger than that of the free water H2O, while the value in BSb(H2O) is almost equal to that of H2O.

C. Low-frequency range

In order to analyze the frequency dependence of the system dynamics, a discrete Fourier transform (DFT) was computed for various dynamical variables. For the 100-ns trajectories the DFT covers the range 10 MHz–100 GHz. As explained in Appendix A 2, the limits of the range of the DFT are determined by the time length of the trajectory (its inverse being the lower limit) and by the time interval between recorded frames (its inverse being twice the higher limit). The low-frequency structured profile of the DFTs, which are displayed and discussed below, depends on the lower limit of the DFTs. If the latter were computed on a trajectory of 50 ns instead of 100 ns, their first data, and their first peak, would move from 10 to 20 MHz, maintaining a similar profile. Therefore, in the diagrams reporting the low-frequency DFTs, the frequency is given in arbitrary units (a.u.) even though its nominal units are GHz.

The DFTs of variables not depending on the configuration of the protein, such as its temperature and the system pressure, produce a uniform spectrum over the whole frequency range (up to 100 GHz) without particular intensity at the lowest frequency, see Figs. 3 and 4. These spectra increase uniformly when the protein's temperature increases from 300 to 600 K. The average increase in BSb vs BSa is a factor of the order of 2.5 for the DFT spectrum of the protein's temperature and a factor of the order of 1.3 for the DFT spectrum of the pressure of the system.

On the other hand, dynamical variables that depend on the configuration of the protein, such as $R_g(t)$, $M(t)$, and the system's potential energy, have a profile with high peaks at the beginning of the low-frequency range (0–0.5 GHz), the first peak being usually the highest. This profile is followed by a low, flat profile in the rest of the frequency range. In Figs. 5 and 6 we report the DFT at low frequency of $R_g(t)$ and of $M(t)$ of BSA, respectively, highlighting the values of the highest peaks in all 100-ns runs. In these two figures the highest first peak is found when the protein is in BSb, that
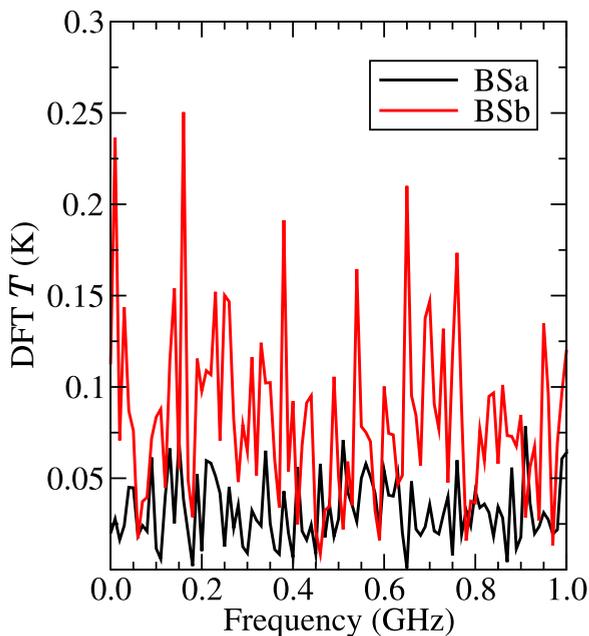
FIG. 3. Low-frequency range of the discrete Fourier transform of the protein's temperature computed on trajectories BSa and BSb.
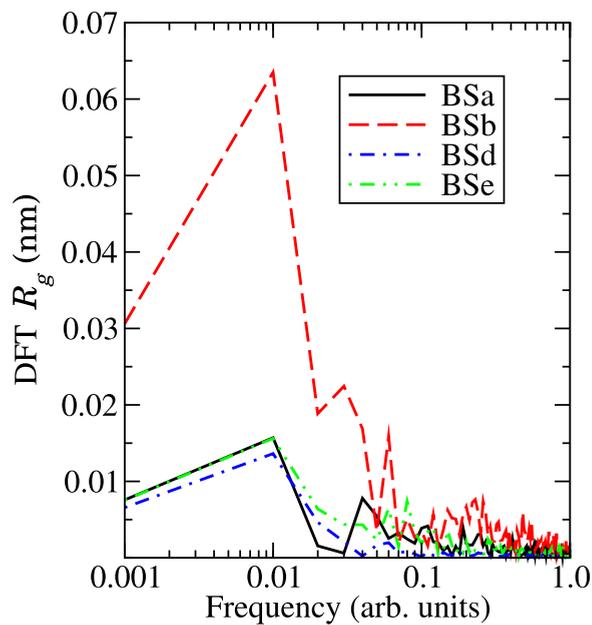

FIG. 5. Low-frequency discrete Fourier transform of the gyration radius $R_g(t)$ of protein BSA in four trajectories. Frequency nominal units: GHz.

is, in the out-of-equilibrium state. The DFT peaks of $R_g(t)$ and of $M(t)$ in the BSb trajectory dwarf those of the other trajectories.

The height of the first peak of the DFT of $R_g(t)$ and of $M(t)$ provides qualitative information about the amount of vibration condensed in the low-frequency region. However, due to the differences in the peaks' pattern in the various systems, it is difficult to use them for a quantitative estimate of

this condensation. In order to compute a quantity comparable in different states of the system, it is convenient, as explained in Appendix A 2, to measure the area under the spectrum in both the low-frequency range (0–0.5 GHz) and in the whole frequency range (0–100 GHz). The ratio of the former area to the latter gives a parameter C that describes the fraction of vibrational energy concentrated in the low-frequency region; these values are reported in Table I (R_g) and Table II (M).

FIG. 4. Discrete Fourier transform of the protein's temperature computed on trajectories BSa and BSb. The points are running averages over 2 GHz.
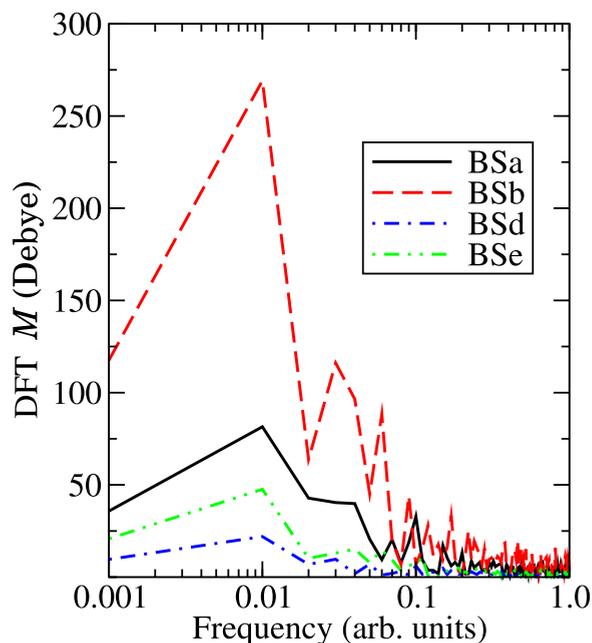

FIG. 6. Low-frequency discrete Fourier transform of the dipole moment $M(t)$ of protein BSA in four trajectories. Frequency nominal units: GHz.

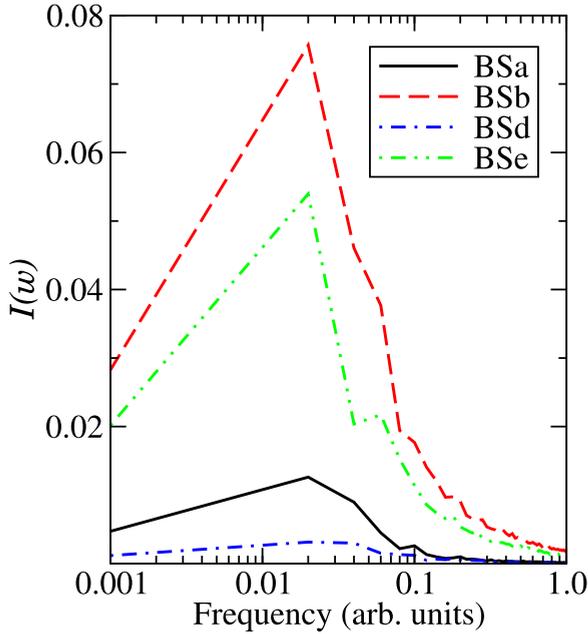FIG. 7. Low-frequency absorbance $I(\omega)$ of protein BSA in four trajectories. Frequency nominal units: GHz.

FIG. 8. Eigenvalues of the covariance matrix of the four trajectories of BSA. The lines are guides to the eye.

For variables having a uniform DFT profile, C should equal the ratio of the extensions of the two ranges, that is, 0.5%. Indeed, from the DFT of the protein's temperature one finds $C = 0.5\%$ both in BSa and BSb. Any C value larger than 0.5% points, therefore, to a vibrational condensation in low-frequency modes.

The values of C in Table I (R_g) for BSA in solvent are notably high: 11.9% at 300 K and 17.7% at 600 K; this is 24 times and 35 times higher, respectively, than that of a uniform DFT profile. As for the protein's M values in Table II, these also show a notably high C in solvent. C increases from 11.0% at 300 K to 15.3% at 600 K.

The low-frequency condensation of vibrational activity of protein BSA affects its absorbance $I(\omega)$, which is proportional to the DFT of the autocorrelation function (ACF) of the protein's dipole $\boldsymbol{M}(t)$ [31]:

$$I(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-i\omega t) \; <\boldsymbol{M}(0) \cdot \boldsymbol{M}(t)> \; dt.$$

We computed this ACF, as well as its DFT, for the four trajectories of BSA, the results of which are given in Fig. 7. The time range of the ACF is half the length of the trajectory, 50 ns; therefore, the first peak of the absorbance moved from 0.01 to 0.02 nominal GHz. With the previous caveat about the "true" value of the frequency on the horizontal axis, the large increase in the absorbance of protein BSA in the out-of-equilibrium state BSb compared to BSa mirrors the increase of the absorption found in the two experiments previously mentioned [8]. If the length of the original trajectory were limited to 6.37 ps, the peak of the absorbance would be exactly at 0.314 THz, as shown in Appendix A 3; this is the frequency where an absorption peak has been located in those experiments.
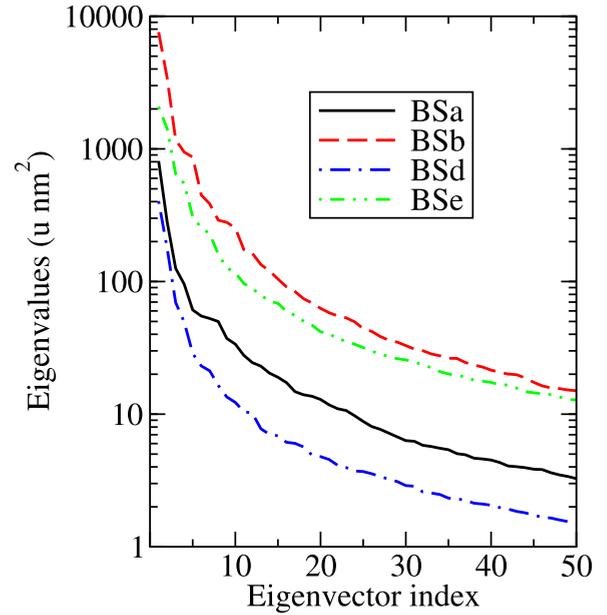
D. Covariance analysis

A further insight into the particular features of a protein's dynamics in the out-of-equilibrium state and into the role of the solvent can be obtained by resorting to a covariance analysis [32,33]. We computed the covariance matrix and its eigenvectors for each of the 100-ns trajectories entailing the protein after removing the rototranslational motion. The first eigenvectors of the covariance matrix correspond to the largest eigenvalues; these are known as the principal components (PCs), and they represent, better than the normal modes, the dynamics of systems governed by an anharmonic potential [32,34,35]. And the low-frequency principal components span the same subspace as the low-frequency normal modes [36,37].

For each PC an effective frequency ω can be defined through the formula $\sigma^2 = k_B T/\omega^2$ [32,38], where the fluctuation σ is its eigenvalue. As the PCs are ordered by decreasing eigenvalues, they are ordered by increasing effective frequency ω and the first PC has the lowest frequency. However, it should be noted that the effective frequencies assigned to the PCs acquire their physical content only for a harmonic potential surface and in the absence of solvent, conditions that are not met in the systems presented here [38].

In Fig. 8 we report the eigenvalues of the first 50 eigenvectors of BSA. It is well known that a limited number of the first PCs account for a large fraction of the total dynamics of a protein [32,34,35,39]. In our systems the sum of the first ten eigenvalues computed over the 100-ns trajectories amounts to 67.9% of the sum of all eigenvalues for BSa, 82.3% for BSb, 69.3% for BSd, and 67.2% for BSe.

The BSa, BSb, BSd, and BSe trajectories were projected on the respective first ten PCs. The inspection of the time evolution of the first ten PCs of BSa and BSb shows a similar pattern in the two trajectories, but the amplitude in BSb is about double that in BSa, as expected. While PCs 2–10
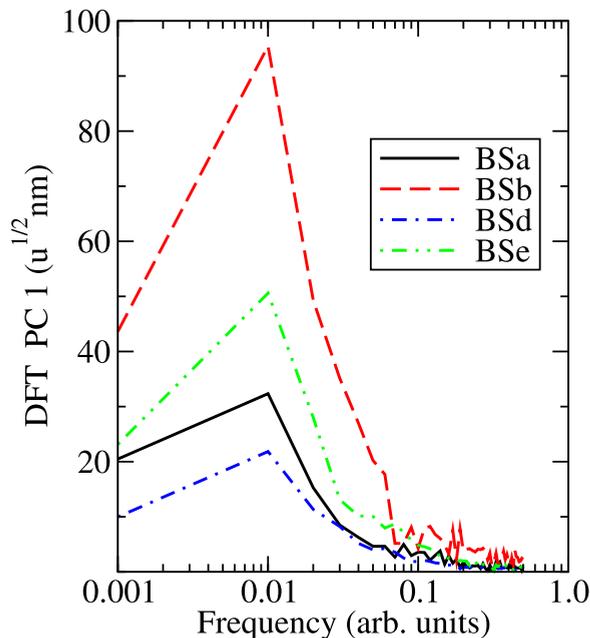
FIG. 9. Low-frequency discrete Fourier transform of the amplitude of the principal component PC 1 of four trajectories of BSA. Frequency nominal units: GHz. u is the atomic mass unit.
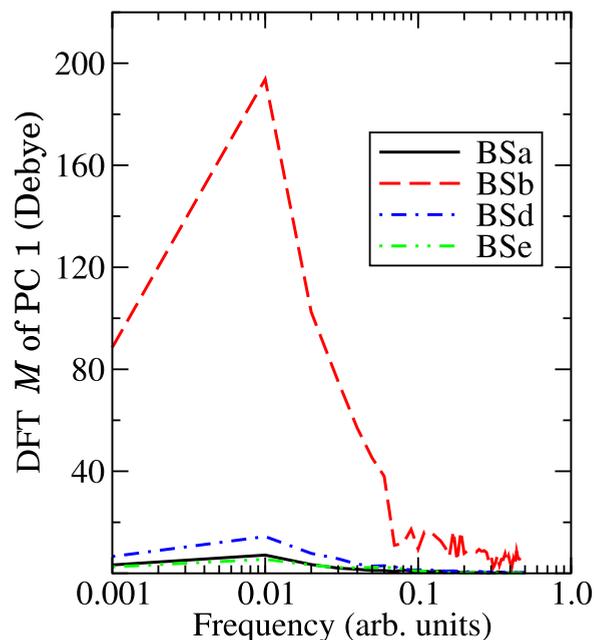

FIG. 10. Low-frequency discrete Fourier transform of the dipole moment $M(t)$ of the projection of four trajectories of BSA on their respective principal component PC 1. Frequency nominal units: GHz.

oscillate around their average value, in both cases PC 1 shows a monotonic decrease that eventually overturns its initial value, hinting at an oscillation period of the order of 200 ns, corresponding to a frequency of about 5 MHz. The amplitude of the PC oscillation in vacuum (BSd, BSe) is similar to, or lower than, the amplitude of the oscillation in solvent at the corresponding temperature.

We computed the DFT of the amplitude of the first ten PCs of the four trajectories of protein BSA. Even though the PCs computed on different trajectories are different, a comparison among equal-ranking PCs yields interesting insights. Figure 9 highlights the highest peaks of the DFT of PC 1, all found at the low-frequency end of the spectrum. The first DFT peak of BSb's PC 1 is much higher than the first peaks of the other trajectories. The results for PC 2 and PC 3 show a similar hierarchy of the protein's states, but the height of the peaks gradually diminishes.

In order to selectively attribute to single PCs the time evolution of $M(t)$ in the low-frequency range, we computed it on the first ten PCs of each trajectory. The decrease in M's average value in BSb compared to BSa can be ascribed mainly to PC 1. We have also computed the DFT of $M(t)$ of the first ten PCs of the four trajectories of protein BSA. The results for PC 1 are given in Fig. 10. The first DFT peak of $M(t)$ of BSb's projection on PC 1 dwarfs the peaks of the other trajectories, highlighting the contribution of this low-frequency PC to $M(t)$'s overall oscillation in the out-of-equilibrium state. The results for PC 2 and PC 3 show a pattern that is similar to the pattern of PC 1, with the peak heights gradually diminishing.

E. High-frequency range

The vibrational power spectrum of a protein can be estimated via a DFT of the ACF of its velocities. The results for

BSa and BSb are given in Fig. 11. The spectrum of BSb shows a lower profile than BSa, with the remarkable exception of a new, high and narrow peak at around 113 THz.

The spectrum of BSa is very similar to the vibrational density of states computed by means of a normal mode analysis, as shown in Fig. 12.

In order to analyze how the change in the vibrational power spectrum translates into the dynamical behavior of
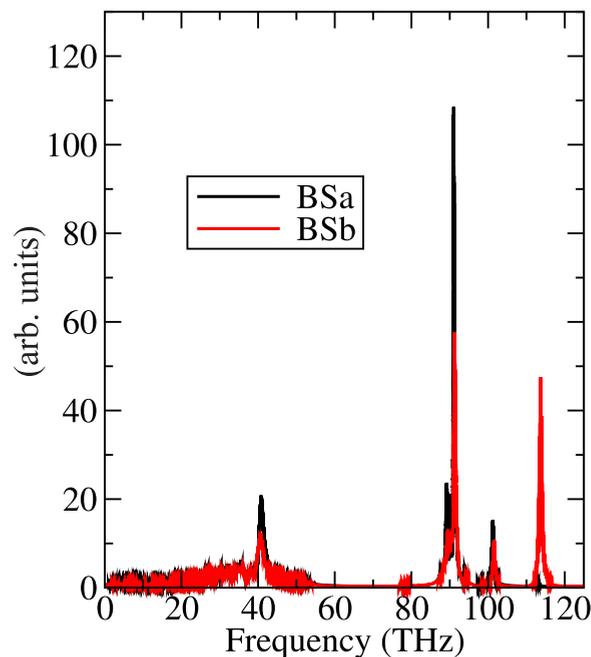

FIG. 11. Vibrational power spectrum of protein BSA at $T = 300$ K (BSa) and $T = 600$ K (BSb).
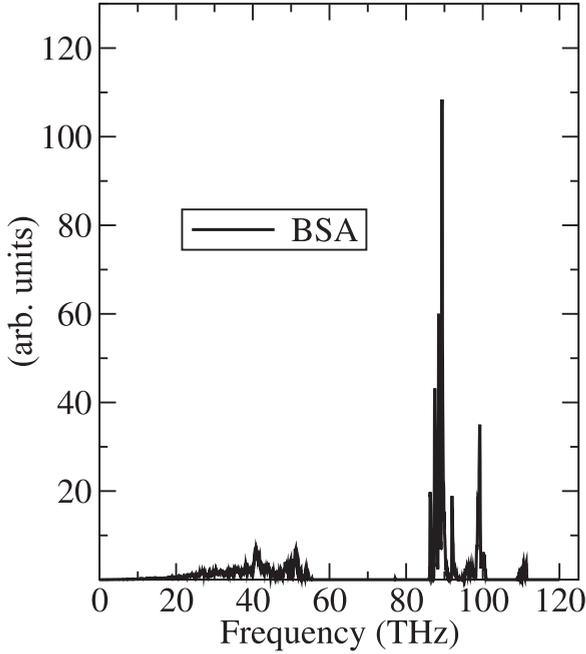
FIG. 12. Vibrational density of states of the protein.

FIG. 13. High-frequency DFT of the temperature of the BSA protein in trajectories BSa and BSb.

BSA at higher frequencies, the DFT of various quantities was computed in the range 0–125 THz using 1-ns trajectories with frames recorded every 4 fs (Appendix A 2). These high-frequency spectra are lower than their homologous low-frequency spectra, reflecting the lesser amplitude of the high-frequency oscillation. The DFT of BSA's $R_g(t)$ does not exhibit any remarkable feature in this whole frequency range other than a uniform increase of the spectrum of BSb compared to the spectrum of BSa caused by the increase in the temperature.

But other quantities in BSb exhibit localized enhanced peaks characterized by an increase that significantly exceeds the one expected when the temperature of the protein is doubled. As shown in Fig. 13, the high-frequency DFT of the protein's temperature exhibits in BSb two enhanced peaks compared to BSa: one around 23 THz (increase factor about 13.2 compared to BSa) and one around 93 THz (increase factor about 6.6). Also, the DFT of the system's potential has two enhanced peaks at the same frequencies as the protein's temperature. And the DFT of protein's $M(t)$ exhibits in BSb a strongly enhanced peak at around 113 THz (increase factor about 6.5), as shown in Fig. 14; this is related to the new high peak in the vibrational power spectrum previously mentioned.

All the DFTs so far mentioned in this section were computed also for the trajectories in vacuum BSd and BSe; these do not show any anomalous behavior when the temperature of the protein is doubled. In particular, the DFT of protein's $M(t)$ computed on the trajectory at 600 K in vacuum (BSe) does not exhibit any unexpected localized increase of the spectrum, as shown in Fig. 15.

IV. DISCUSSION

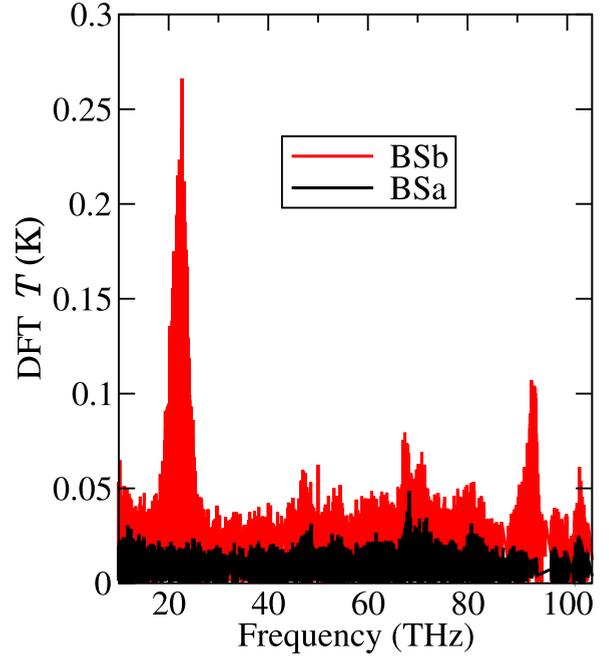Three essential ingredients of the EMI theoretical model are the energy input into the protein, the solvent, and the

out-of-equilibrium state [8]. In the present simulation we mimic the experimental input of energy by an increase ΔT of the protein's temperature in relation to the solvent temperature. The role of these ingredients in the dynamics of the protein can be elicited by comparing the data derived from the different trajectories in solvent, in vacuum, at equilibrium, and in the out-of-equilibrium state. The relations among the
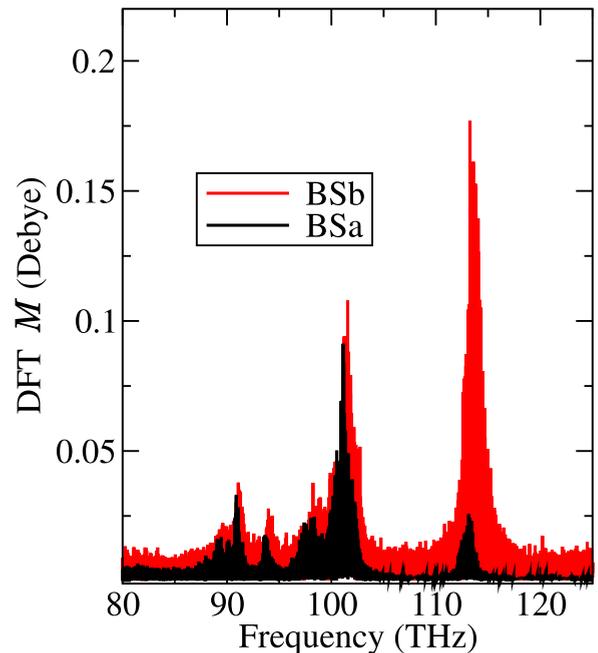
FIG. 14. High-frequency discrete Fourier transform of the dipole moment $M(t)$ of two trajectories of BSA in solvent.
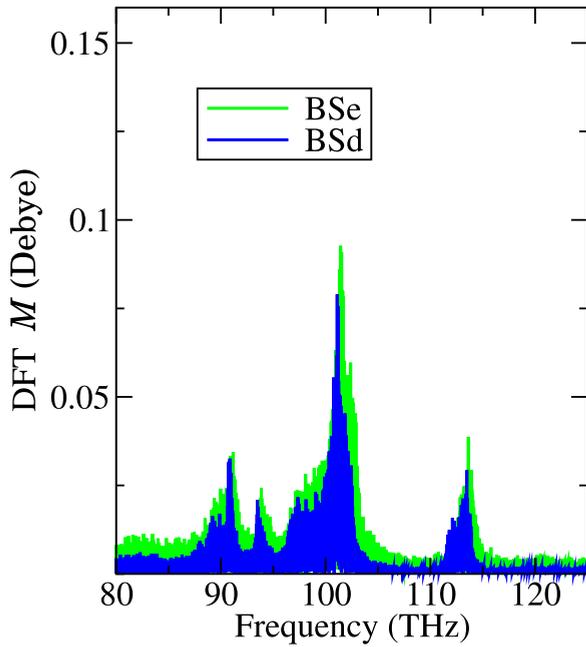
FIG. 15. High-frequency discrete Fourier transform of the dipole moment $M(t)$ of two trajectories of BSA in vacuum.

different states simulated in the present work, and the three mentioned ingredients, are depicted in Fig. 16.

The procedure for a quantitative estimate of the ingredients mentioned is as follows. Starting from BSd (d: in vacuum, 300 K), one can switch to BSa (a: in solvent, 300 K) or to BSe (e: in vacuum, 600 K). The modification $d \rightarrow a$ entails the addition of solvent; the modification $d \rightarrow e$ entails a ΔT
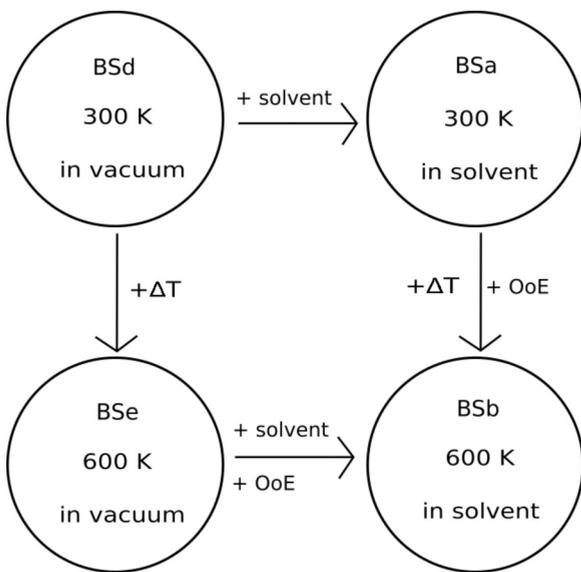

FIG. 16. Block diagram of the four states in which the protein has been simulated, highlighting the changes needed to shift the system from one state to another. OoE is an acronym for out-of-equilibrium. ΔT is the increase in the protein's temperature from 300 K to 600 K. The temperatures in the diagram are those of the protein; the solvent is always at $T = 300$ K.

increase of the protein's temperature from 300 K to 600 K. Then two more modifications are possible: $a \rightarrow b$, entailing ΔT plus the creation of the out-of-equilibrium (OoE) state, and $e \rightarrow b$, entailing the addition of solvent plus the creation of the OoE state (b: 600 K, in solvent at 300 K). Every quantity X [R_g, M, their s.d., $I(\omega)$] computed in state b thus has three contributions: ΔT, the solvent, and the OoE state. In order to compute the contribution of the latter, one has to subtract from the value Xb the contribution to that value of ΔT and of the solvent. (i) The contribution of ΔT can be estimated from $Xe - Xd$, the temperature being the only parameter that changes in that modification. (ii) The contribution of the solvent can be estimated from $Xa - Xd$, its presence being the only parameter that changes in that modification. (iii) The contribution of the OoE state is therefore $(Xb - Xa) - (Xe - Xd)$ or, which is equivalent, $(Xb - Xe) - (Xa - Xd)$. These three contributions can be computed using the data of Tables I and II, in the approximation that the three contributions are independent.

In Table III we report the three contributions to R_g and its s.d., and to M and its s.d. We have also computed, by the same procedure described above, the three contributions to the height of the first peak of the absorbance shown in Fig. 7, taken as an indicator of its intensity. The significant features are as follows: (i) The solvent is always an increasing factor both for R_g and M, as well as for the amplitude of their oscillations. (ii) The out-of-equilibrium state is an increasing factor for R_g and its oscillation's amplitude. (iii) The out-of-equilibrium state is a decreasing factor for M but an increasing factor for the amplitude of its oscillation. (iv) As for the absorbance, the main contribution to its increase from BSa to BSb comes from the increase in temperature; the increases due to the other two factors have similar magnitude, with a prevalence of the OoE state. We now discuss, within the framework of Table III, the results reported in Sec. III.

A. Solvent and out-of-equilibrium state

As reported in the Introduction, the EMI theoretical model assigns a fundamental role in promoting the condensation of energy in the low-frequency vibrational modes to the solvent. This role in the dynamics of the protein can be brought into focus by comparing its trajectories in solvent with those in vacuum at the same temperature: BSa vs BSd, and BSb vs BSe.

The role of solvent in the dynamics of BSA is clearly visible in Table I: its presence increases significantly the average value of R_g in BSa compared to BSd (both at 300 K), and in BSb compared to BSe (both at 600 K). The solvent also increases the amplitude of the oscillation of $R_g(t)$, compared to the trajectories in vacuum at the same temperature, by a factor of 1.5 (BSa vs BSd) and 3.1 (BSb vs BSe), respectively. As mentioned in Sec. III D, the effect of the solvent can also be seen in the comparison of the amplitude of the PC oscillation in solvent and in vacuum, the former being generally larger than the latter. This effect is counterintuitive, as the solvent usually dampens a protein's oscillation and therefore diminishes its amplitude; indeed, it is in contrast with the findings of a molecular dynamics simulation of melittin, where the amplitude of the oscillations of the PCs in vacuum was larger

TABLE III. Contributions to the average values of R_g and M, and to the amplitude of their oscillations, from the temperature difference between protein and solvent, from the solvent and from the out-of-equilibrium state. First column: Increase or decrease factor. ΔT is the 300-K difference between the protein's temperature and temperature of the solvent, Solvent is the introduction of solvent, and OoE is the shift of the protein into the out-of-equilibrium state. Second column: Variation of R_g. Third column: Variation of the s.d. of $R_g(t)$. Fourth column: Variation of M. Fifth column: Variation of the s.d. of $M(t)$. Sixth column: Variation of the maximum absorbance in Fig. 7.

Factor	ΔR_g (nm)	Δ s.d. R_g (nm)	ΔM (D)	Δ s.d. M (D)	$I(\omega)$ (arb. units)
ΔT	−0.059	+0.007	−190	+26	+0.0508
Solvent	+0.208	+0.006	+676	+73	+0.0097
OoE	+0.243	+0.033	−142	+130	+0.0122

than their amplitude in solvent by a factor up to 4 [38]. Here the presence of solvent, together with the out-of-equilibrium state, contributes to the increase by 6.8% of R_g's average in the BSb trajectory compared to the BSa trajectory. This positive role is confirmed by contrast with the two trajectories in vacuum, BSd (300 K) and BSe (600 K), where doubling the temperature of the protein decreases the average value of R_g by 2.4%. Moreover, the solvent, together with the out-of-equilibrium state, also enhances the increase of the amplitude of the oscillation of $R_g(t)$ when the temperature is doubled. As measured by the standard deviation, the amplitude increases by a factor of 3.3 in BSb vs BSa, more than double the factor of the order of $2^{1/2}$ expected in harmonic approximation when the temperature is doubled. In contrast, in vacuum the increase factor of the standard deviation of $R_g(t)$ in BSe vs BSd is 1.6, that is, within the order of the expected $2^{1/2}$.

The presence of the solvent significantly increases the average value of M in BSa compared to BSd (factor 2.8) and in BSb compared to BSe (factor 3.9); it also increases the amplitude of its oscillation, measured by the standard deviation, by a factor of 3.7 (BSa vs BSd) and 4.8 (BSb vs BSe), respectively. The factor 3.7 in the first case must be attributed entirely to the solvent, which enhances the oscillation of the protein's dipole moment due to the electret effect. The larger value 4.8 of the increase factor in the second case must therefore be attributed, beyond the role of the solvent, to the positive contribution of the out-of-equilibrium state of the protein in BSb. The oscillation of $M(t)$ increases by a factor of 2.6 from BSa(prot) to BSb(prot), which is significantly larger than the $2^{1/2}$ expected in the harmonic approximation; this is due to the combined contribution of ΔT and of the out-of-equilibrium state. The same increase factor in vacuum from BSd to BSe, due only to ΔT, is 2.0.

Because of the large jumps in the time evolution of both $R_g(t)$ and $M(t)$ of the BSb trajectory, visible in Figs. 1 and 2, one could ask whether the results before and after the jumps differ. While the average values R_g and M are obviously different, all results of BSb relevant to the analysis performed in this paper (parameters A and C, low- and high-frequency DFTs, absorbance, vibrational power spectrum) are approximately unchanged when measured in the time ranges before or after the jump. The study of these instabilities goes beyond the scope of the present work.

B. The electret

The thickness of the hydration shell of protein BSA has been estimated to be around 2.5 nm [40]. In our simulation

the initial distance between the protein and the wall of the box is equal to 1.0 nm or slightly larger. Therefore, the hydration shell provided by the water in the simulation box is only a fraction of the experimental one; notwithstanding this limit, one can ask whether the water in our simulation is sufficient to create a coordination effect around the protein.

Comparing Figs. 17 and 18, where the $M(t)$ data were taken before the BSa trajectory was corrected to eliminate the rototranslational motion of the protein, one can detect, beyond the local oscillations of the curves, a synchronization between the long-term trend of the homologous components of the two dipoles, with the components of the protein's $M(t)$ following those of the water's $M(t)$ with a delay of about 2–3 ns. The synchronization of the dipoles can thus be described as the water dragging the protein's dynamics; this is evidence of a pattern called the solvent-slaved protein [41–43]. This interpretation is confirmed by the correlation time of the ACF of the water's dipole moment. It is 5.1 ps in H2O and 8.7 ps (70% higher) in BSa(H2O), highlighting the temporal organization of the hydration water induced by the presence of the protein. This behavior can be attributed to an oriented-dipole electret, where the coordination layer of the water molecules around a protein forms a semistable structure whose dipole coordinates with that of the protein. As already noted, this electretlike effect enhances the protein's M [BSa(prot)]; it also slightly enhances the water's M [BSa(H2O)] by a factor of 1.1 compared to its value in pure water in the same condition (H2O). The alignment of the dipoles in this state is summarized in the value $A = 43\%$ for trajectory BSa, the strongest alignment recorded in the present simulation. Therefore, it appears that the alignment of the dipoles and their mutual strengthening are concurrent phenomena.

A similar extended synchronization pattern cannot be observed in the analogous figures for BSb, as the higher temperature of the protein weakens the coordination, as shown in Figs. 19 and 20. Also here the data were taken before the BSb trajectory was corrected to eliminate the rototranslational motion of the protein. At best, some synchronization can be tentatively detected for limited stretches of time. This can be interpreted as an unstable water electret intermittently loosening its structure and/or weakening the alignment between its dipole and the protein's dipole. Here again, this interpretation is confirmed by the correlation time of the ACF of the water's dipole moment: 7.3 ps in BSb(H2O), still 43% higher than in H2O. Notwithstanding the more chaotic dynamics of the protein at 600 K, the latter is still able to induce a temporal organization of the hydration water, albeit less than in BSa. When the temperature of the protein is doubled, M of the
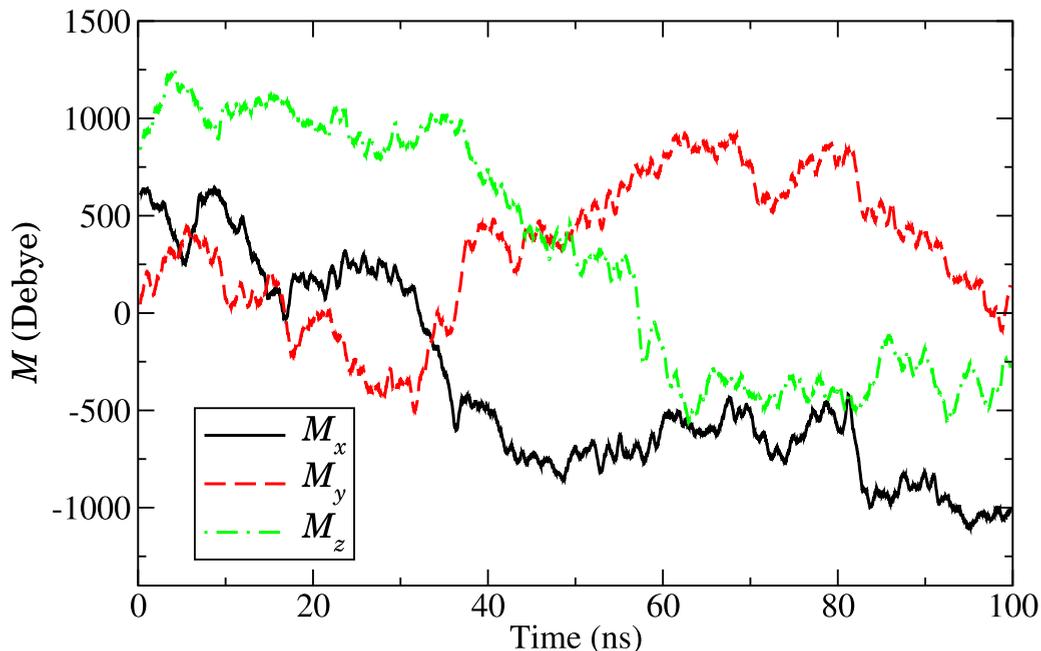
FIG. 17. Components of the dipole moment $M(t)$ of protein BSA at equilibrium at 300 K (trajectory BSa). The points are running averages over 0.5 ns.

water decreases from 757 D [BSa(H2O)] to 684 D [BSb(H2O)] oscillating near the value of 687 D of pure water (H2O). On the other hand, the $M = 719$ D value of the protein at 600 K in the solvent [BSb(prot)] is larger by a factor of 3.9 than the $M = 185$ D value of the protein at the same temperature in vacuum (BSe). This factor is the combined effect of the solvent and of the OoE state. A similar comparison at 300 K between BSa(prot) and BSd yields a factor of 2.8 due to sole solvent.

C. Condensation of energy

The condensation of vibrational energy in the low-frequency vibrational domain has been extensively proven by the results in Sec. III C. The data for BSb show that the condensation of energy in the out-of-equilibrium state of the protein strongly prevails over the other states. The amount of condensed energy, as measured by parameter C, clearly distinguishes between physical properties where this condensation takes place (Figs. 5–7: gyration radius, dipole moment,
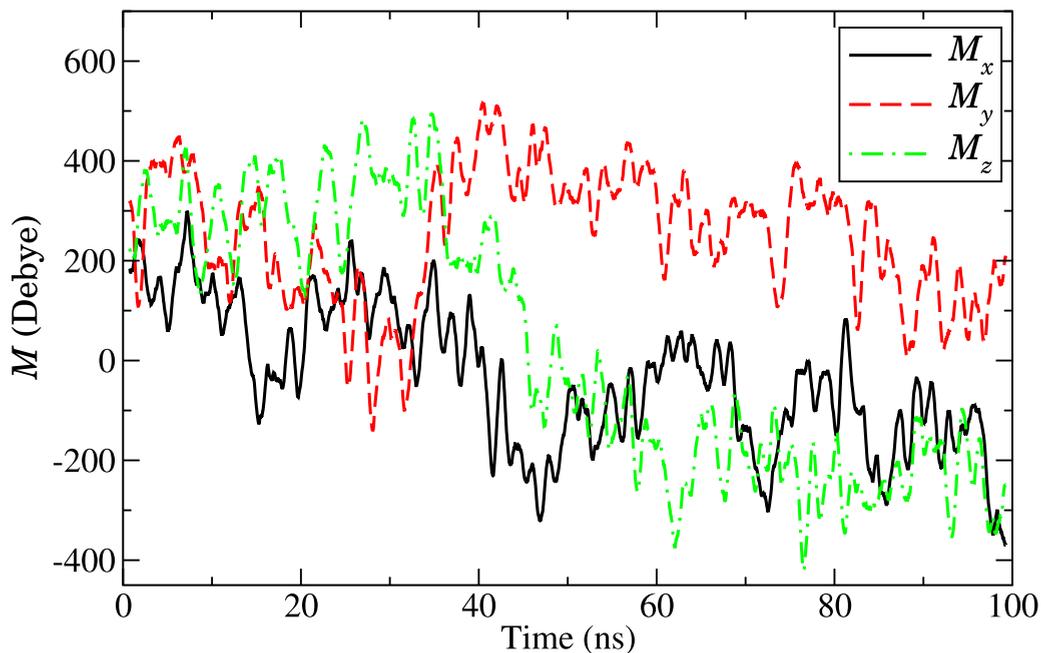

FIG. 18. Components of the dipole moment $M(t)$ of the water (around protein BSA) at equilibrium at 300 K (trajectory BSa). The points are running averages over 1.0 ns.
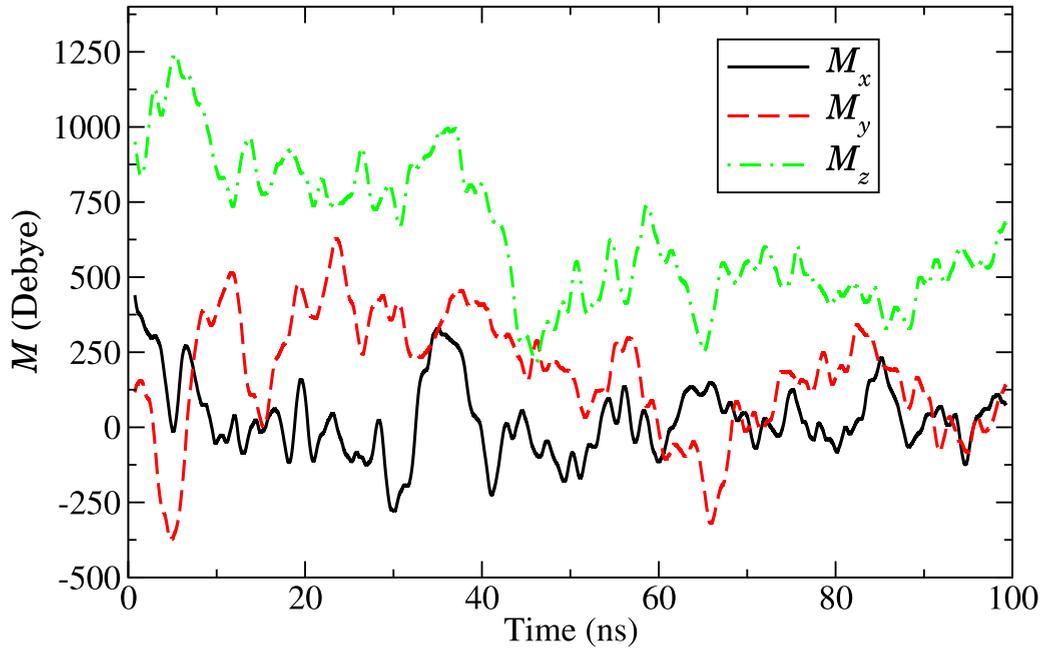
FIG. 19. Components of the dipole moment $M(t)$ of protein BSA in the out-of-equilibrium state at 600 K (trajectory BSb). The points are running averages over 0.5 ns.

absorbance) and physical properties lacking condensation (Figs. 3 and 4: temperature, system pressure). Considering in Table I the energy condensation in R_g, when the temperature is doubled and the protein is pushed into the OoE state, the protein's C in solvent increases from 11.9% (BSa) to 17.7% (BSb). Doubling the temperature in vacuum, the protein's C has a smaller increase from 6.8% (BSd) to 8.3% (BSe). Considering in Table II the energy condensation in M, when the temperature is doubled and the protein is pushed

into the OoE state, the protein's C in solvent increases from 11.0% to 15.3%. Doubling the temperature in vacuum, the protein's C has a smaller increase, from 6.0% to 6.3%. Both patterns therefore show that the energy condensation at low frequencies is sustained by the presence of the solvent and of the out-of-equilibrium state. These findings support the basic assumptions of the EMI model, namely, that the solvent and the OoE state enhance the low-frequency condensation of vibrational energy that is necessary for the protein to absorb
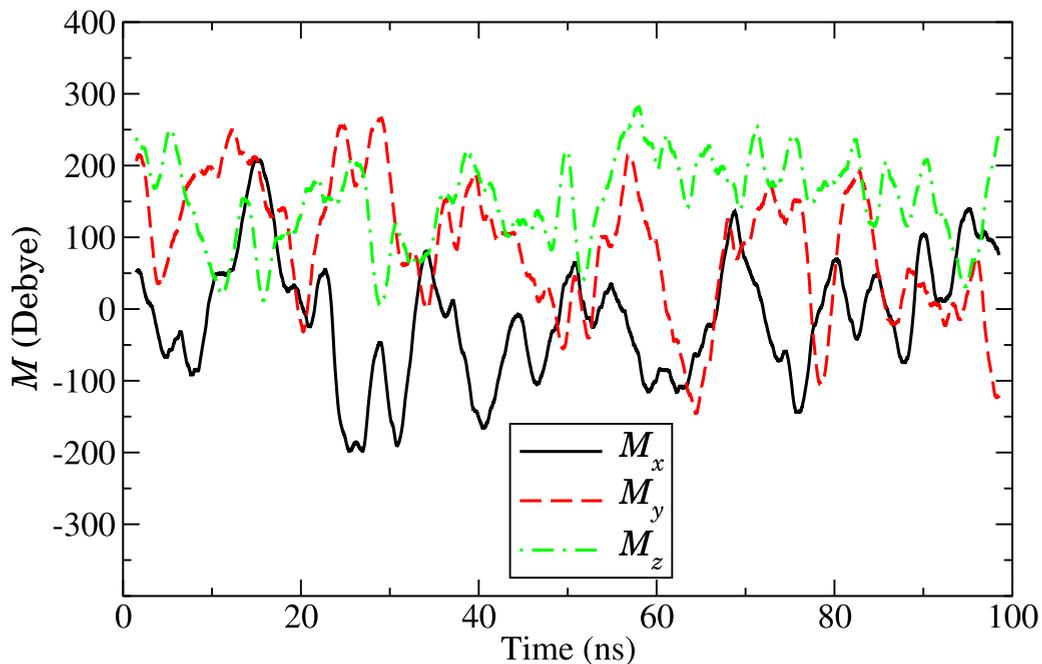

FIG. 20. Components of the dipole moment $M(t)$ of the water (around protein BSA) in the out-of-equilibrium state (trajectory BSb). The points are running averages over 1.0 ns.

or emit an e.m. radiation supporting the long-range interaction with other proteins.

D. Principal components

As described is Sec. III D, Fig. 9 for PC 1, and similar figures for PC 2 and PC 3 (not shown), highlight the highest peaks of the amplitude DFT of the first three PCs, all found at the low-frequency end of the spectrum. As in the case of $R_g(t)$ and $M(t)$ of the original trajectories, the most remarkable feature of these three figures is the increase of the first peak in BSb compared to BSa. The conjunctive role of the solvent and of the out-of-equilibrium state in BSb is evident in these first three PCs, as the increase in the first peak found shifting in vacuum from BSd to BSe is smaller. As far as the amplitude of the oscillations of the PCs is concerned, the hierarchy BSb > BSe > BSa > BSd of the peaks' height emphasizes the prevailing role of the higher temperature in determining the height of these peaks. The presence of the solvent gives a further significant contribution to the hierarchy at 600 K, while at 300 K the difference between BSa and BSd is small in PC 1 and negligible in PC 2 and PC 3.

Figure 10 for PC 1, and similar figures for PC 2 and PC 3 (not shown), highlight the highest peaks of the DFT of the dipole's moment $M(t)$ of the first three PCs. Again, and even more than in the previous figures relative to the PCs' amplitudes, the most remarkable feature is the huge increase of the first peak in PC 1 and PC 3 shifting form BSa to BSb. These peaks dwarf those of the other trajectories, highlighting the contribution of these low-frequency PCs to the overall $M(t)$ oscillation in the out-of-equilibrium state. This huge difference is due to the combination of the out-of-equilibrium state and of the solvent. The increase of the first peak in PC 1 and PC 3 shifting in vacuum from BSd to BSe is due only to ΔT and is much smaller. Here the contribution of the OoE state to the peaks' height can be estimated by observing that in PC 1 and PC 3 there are small differences among BSa, BSd, and BSe, their peaks being much lower than BSb's peak; the first three systems lack the solvent and/or the higher temperature, i.e., they lack the OoE state. On the other hand, the peaks of PC 2 show that both the solvent (BSa > BSd, BSb > BSe) and the temperature (BSb > BSa, BSe > BSd) contribute independently to the height of the peaks.

The parameter C, measuring the ratio of the area under the low-frequency range of the DFT (0-0.5 GHz) to the area under the DFT of the whole frequency range (0-100 GHz), as previously defined, was computed both for the amplitudes and for the $M(t)$ of the first ten PCs. These values are reported in Figs. 21 and 22. Both figures highlight the larger condensation of energy taking place in shifting from BSa to BSb as opposed to the much smaller change in C when shifting from BSd to BSe. In the first case the condensation is strongly increased due to the out-of-equilibrium state of the protein and to the presence of the solvent with its electretlike structure around the protein. The two figures also clearly show that the first principal component, characterized by the lowest effective frequency, provides the highest energy condensation.

To summarize this section, the results pinpoint the role of the solvent and of the out-of-equilibrium state in the
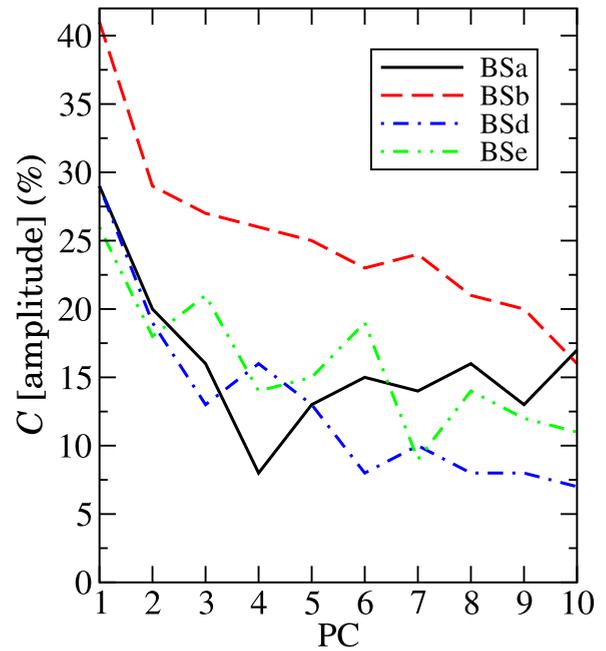

FIG. 21. Parameter C, measuring the condensation of energy in the low-frequency range of the amplitude oscillation's DFT of the first ten principal components. The lines are guides to the eye.

condensation of the vibrational energy in the lowest-frequency mode defined by the covariance analysis.

E. High-frequency spectrum

As mentioned in Sec. III E, the DFTs computed for the trajectories in vacuum (BSd and BSe) do not show any peculiar

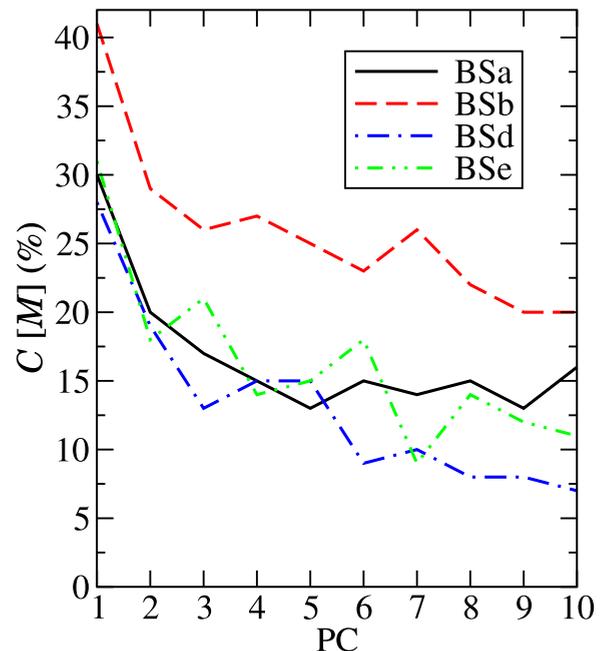
FIG. 22. Parameter C, measuring the condensation of energy in the low-frequency range of the dipole moment's DFT of the first ten principal components. The lines are guides to the eye.

behavior when the temperature of the protein is doubled, as they do not exhibit any unexpected localized increase of the high-frequency spectrum. This is a sign that the enhancement of the DFTs of the proteins' $M(t)$ described in that section can by ascribed to the presence of the solvent.

The frequencies of 93 THz (3100 cm^{-1}) and 113 THz (3770 cm^{-1}) of the peaks found in the DFT of $M(t)$ in BSb fall into the opposite outer wings of the vibrational spectrum of a water molecule. But in the present simulation a rigid model (SPC/E) of water was used; therefore, the peaks at those frequencies cannot be ascribed to the vibration of the water molecules. There could still be a contribution by a network of H bonds forming among water molecules [44]; it has been found that such clusters have vibrational frequencies between 3273 cm^{-1} (98.2 THz) and 3400 cm^{-1} (102.0 THz) [45]. To check this possibility, we computed the DFT of the temperature, of the potential, of the dipole moment $M(t)$, and of the pressure of the sole water in trajectory H2O. None of these DFTs has a peak at the frequencies of 23 THz, 93 THz, or 113 THz; these frequencies correspond therefore to a high-susceptibility reaction of BSA to the out-of-equilibrium condition and to the electretlike structure of water around the protein.

In BSb the 113-THz frequency of the proteins' $M(t)$ DFT peak (Fig. 14) is very near to 111 THz, the vibrational frequency of an OH group, for example, in lysine or in tyrosine. Together with the other two peaks of BSb at lower frequencies (23 THz and 93 THz in the spectra of the protein's temperature and of the system's potential), this could be part of a pattern of intermodal coupling allowing the flow of energy from high-frequency vibrational modes to low-frequency accordionlike modes of α-helices (2–3 THz) [8,46] via internal friction [47].

Before proceeding to the conclusions, it is suitable to compare the present results to those of the previous computer experiment on BSA already mentioned [29]. There are some significant differences between the two simulations. (i) We do not use constraints in the system, while they constrain the bond lengths. (ii) The excited state of their system is realized by modifying the charges of two tryptophans in order to mimic a UV excitation of the system. (iii) Their dipole ACFs, from which various absorbances are computed, are the average of some 200 *NVE* stretches of the ground and of the excited trajectories (each stretch being 100 ps, for a total time of about 20 ns) while ours are computed over single *NPT* trajectories of 100 ns. Moreover, they do not provide an estimate of the energy input of the protein when the charges of the tryptophans are modified. Notwithstanding these differences, one observes the convergence of a few results. The average values at equilibrium ($T = 300$ K) of R_g and M are similar; however, because of point (ii) above, it is not surprising that the same variables in the excited state have different values. And the 23-THz-high peak in the DFT of the protein's temperature, found here in the OoE state (Fig. 13), is near a major peak at 21 THz found in the absorbance in their work. There are no other results to compare because they analyze a large variety of structural details related to their excitation of the system; on the other hand, the focus of the present paper is on results supporting the validity of the premises of the revised EMI theoretical model.

V. CONCLUSIONS

The present paper describes the results obtained in an atomistic simulation of protein BSA, an experimental candidate for an e.m. long-range interaction with cognate biomolecules. Our simulation mimics the experimental setup and succeeds in supporting the relevant assumptions of the classical theoretical model used to interpret those experiments under similar conditions: the condensation of energy in the lowest-frequency range, the relevance of the out-of-equilibrium state, and the role of the solvent, specifically, of the hydration layer. We found evidence of an electretlike alignment between the protein and water dipoles.

The reported results allow a quantitative estimate of the contribution of the out-of-equilibrium state and of the solvent to the observed condensation of vibrational energy of the protein and fit into the framework of Table III. The covariance analysis highlights the predominant contribution of the first principal components to the energy condensation. This characteristic produces a strong peak in the lowest-frequency absorbance of the protein in the out-of-equilibrium system; this reflects the core result of Fröhlich's theory as elaborated in the classical model. Taking into account the numerical peculiarity of the low-frequency range of the DFT, this peak could be put at 0.314 THz, where the experimental peak is located.

Some results of the analysis performed on the BSA dynamics may show generic properties of proteins, while others may distinguish proteins that are appropriate candidates for the conjectured long-distance e.m. interaction that has prompted the present study. The present paper provides a procedure for checking the propensity of a biomolecule to interact via e.m. radiation with its biochemical partners.

These properties of proteins are possibly generic features: (i) the role of the solvent in increasing the dipole moment of the protein and the amplitude of its oscillation; (ii) a moderate condensation of energy C, favored by the solvent, in the low-frequency vibrational range of $R_g(t)$ and $M(t)$ at equilibrium at 300 K; (iii) a strong reaction of the dipole moment, when energy is injected into the protein (the average M decreases markedly in BSA).

On the other hand, other properties may be characteristic of proteins able to enter an EMI process. These EMI-grade properties are (i) a strong alignment A of the protein's dipole with the water's dipole; (ii) a large increase in the amplitude of $M(t)$ oscillation in the out-of-equilibrium state compared to the equilibrium state; (iii) an enhancement of the condensation of energy C in the lowest-frequency vibrational range of $R_g(t)$ and $M(t)$ when shifting from equilibrium to the out-of-equilibrium state; and (iv) a large localized peak in the high-frequency DFT of the protein's dipole moment when the protein is forced into the out-of-equilibrium state (the 113-THz peak in BSA).

The susceptibility of the dipole, mentioned above, hints at an instability of the charge distribution of the protein and could be an important ingredient in the absorption or emission of e.m. radiation. From this point of view, EMI-grade property (ii) is the relevant physical property, as the total absorbed or radiated power is proportional to the square of the dipole moment's oscillation.

The distinction between generic properties and EMI-grade properties of proteins is therefore qualitative (presence of property) and quantitative (intensity of property). Properties of the two groups may be similar, but the intensity of the effects is different. EMI-grade property (ii) appears to strongly characterize those proteins that are capable of exchanging e.m. interaction with partners. We can therefore conclude that, within the framework of the EMI model discussed in this paper, protein BSA is a promising candidate for long-range e.m. interaction.

APPENDIX

1. Simulation methods

All simulations of BSA started from the corresponding pdb file provided by the PDB repository [23]; this structure of 9220 atoms was converted into the GROMACS format using the pdb2gmx tool. The protein was solvated in a suitable triclinic box filled with SPC/E water molecules extending in all directions to a distance of at least 1.0 nm from the solute. Periodic boundary conditions (pbc's) were applied to the box. The 1.0-nm distance between the protein and the nearest wall of the box is sufficient to prevent self-interactions of BSA once the system has relaxed; the distance between the protein and its image created by the pbc is, on average, about 2 nm. Using the GROMACS genion module, 16 positive Na counterions were added to the solution to neutralize the total charge on BSA.

The system was first relaxed in the solvent, minimizing its total energy. A 1-ns equilibration in a *NVT* ensemble at $T = 300$ K was then performed using V-rescale temperature coupling [24], followed by a 1-ns equilibration in an *NPT* ensemble at $T = 300$ K and $P = 1$ bar using the Berendsen barostat [25]. The final configuration of this equilibration was the initial one for the actual simulation of trajectory BSa. Starting from the same final configuration of the equilibration, a further 1-ns *NPT* equilibration was performed with two thermostats at $T = 300$ K (solvent) and $T = 600$ K (protein) to produce the initial configuration of trajectory BSb. These 1-ns equilibration runs, performed after the energy minimization of the protein in the solvent, were executed by applying positional restraints of 1000 kJ mol^{-1} nm^{-1} to the protein's atoms and by constraining the H bonds using the LINCS algorithm [26,27]. The runs in vacuum, BSd and BSe, began after the protein's PDB structure, cleaned of the crystallization water molecules, was subjected to a 1-ns *NVT* equilibration without the solvent at the respective temperatures; this equilibration was computed without restraints or constraints. All runs following the equilibrations were computed without restraints or constraints applied to the protein's atoms.
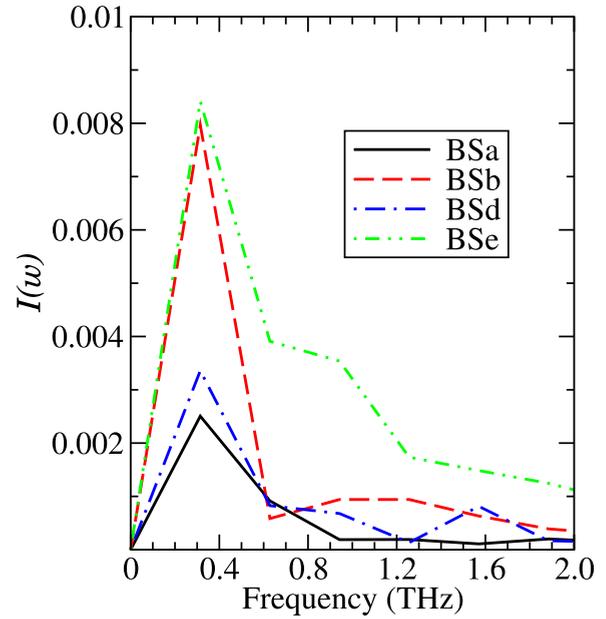

FIG. 23. Absorbance of BSA computed on a 6.37-ps stretch of the 1-ns trajectories.

In the simulation, the short-range interactions were computed using a cutoff of 1.0 nm, and the long-range electrostatic interactions were computed using the particle-mesh Ewald method [28] with a grid spacing of 0.16 nm. The rototranslational movement was removed from each trajectory before computing the dynamical and thermodynamic quantities, with the exception of those analyzed in Sec. IV B of the main text.

All trajectories were computed in two versions: a 100-ns one, during which the system's frames were recorded every 5 ps, and a 1-ns one, with a very dense framework recording, in order to inspect the behavior of the system at short times, that is, at high frequency. In these short runs the configurations were recorded every 4 fs, allowing the extension of the frequency range of the DFT up to 125 THz (see Appendix A 2).

When computing the dipole of the system, GROMACS puts a charge opposite to the charge of the molecule in the center of mass of each molecule. Therefore, to compute the dipole of the whole system (protein + water + counterions), the counterions are switched off while a point charge of $+16$ e is put in the center of mass of BSA.

2. Discrete Fourier transforms

The behavior of the system at low and high frequency was analyzed by discrete Fourier transforms (DFTs). The lowest-frequency limit is given by the inverse of the duration of the trajectory; the highest-frequency limit is given by half of the inverse of the time interval between contiguous recorded configurations of the system. Therefore, the DFTs of the 100 ns trajectories, where the frames were recorded every 5 ps, cover the range 10 MHz–0.1 THz. The DFT of quantities dependent on the structure of the protein, like potential energy, $R_g(t)$, and $M(t)$, display a first peak, sharp and large, usually at a frequency of 10 MHz. This frequency is related to the duration of the trajectory. If one computes these DFTs on a 50-ns trajectory, the first peak is at 20 MHz. On the other hand,

comparing DFTs of different quantities, computed on one and the same trajectory, allows the physical insight described in the present paper.

The peaks entail physical information: in the 50-ns trajectory the height of the major peak is approximately half the height of the major peak computed on the 100–ns trajectory, but its base is approximately double so that the area under this peak approximately keeps its value. Therefore, in Tables I and II we reported the ratio C between the area under the DFT in the low-frequency range 0.0–0.5 GHz, and the area under the DFT in the whole frequency range 0.0–0.1 THz. The ratio of the two ranges is 0.005; thus, values of C higher than 0.5% show a condensation of the oscillation in the lowest-frequency range. 0.5 GHz is the frequency below which the structure-dependent quantities of protein BSA, like $R_g(t)$, $M(t)$, and $I(\omega)$, show strongly enhanced peaks in their DFT, signaling the condensation of vibrational energy in low-frequency modes. Above 0.5 GHz, these DFTs are uniform up to the highest frequency of the DFT computed on the 100-ns trajectories, that is, up to 100 GHz. Therefore 0.5 GHz was chosen as the border between the very low frequencies, where the DFTs show a vibrational energy condensation, and the rest of the low-frequency range.

As for the high-frequency range, the 1-ns trajectories with frames recorded every 4 fs produce DFTs in the range 1 GHz– 125 THz.

3. Virtual absorption peak

Figure 23 displays the absorbance of four trajectories computed on a short 6.37-ps stretch of the 1-ns trajectories where the frames were recorded every 4 fs; it shows how this particular time length would position the peak of $I(\omega)$ exactly at the experimental frequency of 0.314 THz. Obviously, such a short trajectory would have very scant statistical content.

[1] H. Fröhlich, Long-range coherence and energy storage in biological systems, Int. J. Quantum Chem. **2**, 641 (1968).

[2] H. Fröhlich, Long range coherence and the action of enzymes, Nature (London) **228**, 1093 (1970).

[3] H. Fröhlich, Selective long range dispersion forces between large systems, Phys. Lett. A **39**, 153 (1972).

[4] H. Fröhlich, Long-range coherence in biological systems, Riv. Nuovo Cimento **7**, 399 (1977).

[5] J. Preto and M. Pettini, Resonant long-range interactions between polar macromolecules, Phys. Lett. A **377**, 587 (2013).

[6] J. Preto, E. Floriani, I. Nardecchia, P. Ferrier, and M. Pettini, Experimental assessment of the contribution of electrodynamic interactions to long-distance recruitment of biomolecular partners: Theoretical basis, Phys. Rev. E **85**, 041904 (2012).

[7] J. Preto, M. Pettini, and J. A. Tuszynski, Possible role of electrodynamic interactions in long-distance biomolecular recognition, Phys. Rev. E **91**, 052710 (2015).

[8] I. Nardecchia, J. Torres, M. Lechelon, V. Giliberti, M. Ortolani, P. Nouvel, M. Gori, Y. Meriguet, I. Donato, J. Preto, L. Varani, J. Sturgis, and M. Pettini, Out-of-equilibrium collective oscillation as phonon condensation in a model protein, Phys. Rev. X **8**, 031061 (2018).

[9] M. Lechelon, Y. Meriguet, M. Gori, S. Ruffenach, I. Nardecchia, E. Floriani, D. Coquillat, F. Teppe, S. Mailfert, D. Marguet, P. Ferrier, L. Varani, J. Sturgis, J. Torres, and M. Pettini, Experimental evidence for long-distance electrodynamic intermolecular forces, Sci. Adv. **8**, eabl5855 (2022).

[10] A. Ansari, J. Berendsen, S. F. Bowne, H. Frauenfelder, I. E. T. Iben, T. B. Sauke, E. Shyamsunder, and R. D. Young, Protein states and protein quakes, Proc. Natl. Acad. Sci. USA **82**, 5000 (1985).

[11] D. Arnlund, L. C. Johansson, C. Wickstrand, A. Barty, G. J. Williams, E. Malmerberg, J. Davidsson, D. Milathianaki, D. P. DePonte, R. L. Shoeman, D. Wang, D. James, G. Katona, S. Westenhoff, T. A. White, A. Aquila, S. Bari, P. Berntsen, M. Bogan, T. B. van Driel *et al.*, Visualizing a protein quake with time-resolved X-ray scattering at a free-electron laser, Nat. Methods **11**, 923 (2014).

[12] M. Levantino, G. Schirò, H. T. Lemke, G. Cottone, J. M. Glownia, D. Zhu, M. Chollet, H. Ihee, A. Cupane, and M. Cammarata, Ultrafast myoglobin structural dynamics observed with an x-ray free-electron laser, Nat. Commun. **6**, 6772 (2015).

[13] E. Lindahl, B. Hess, and D. van der Spoel, GROMACS 3.0: A package for molecular simulation and trajectory analysis, J. Mol. Model. **7**, 306 (2001).

[14] D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen, GROMACS: Fast, flexible and free, J. Comput. Chem. **26**, 1701 (2005).

[15] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl, GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation, J. Chem. Theory Comput. **4**, 435 (2008).

[16] S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess, and E. Lindahl, GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit, Bioinformatics **29**, 845 (2013)

[17] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX **1-2**, 19 (2015).

[18] GROMACS, https://www.gromacs.org.

[19] E. Faraji, R. Franzosi, S. Mancini, and M. Pettini, Energy transfer to the phonons of a macromolecule through light pumping, Nat. Sci. Rep. **11**, 6591 (2021).

[20] W. Humphrey, A. Dalke, and K. Schulten, VMD: Visual molecular dynamics, J. Mol. Graphics **14**, 33 (1996).

[21] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids, J. Am. Chem. Soc. **118**, 11225 (1996).

[22] P. Mark and L. Nilsson, Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K, J. Phys. Chem. A **105**, 9954 (2001).

[23] Protein Data Bank entry, https://www.rcsb.org/structure/3V03.

[24] G. Bussi, D. Donadio, and M. Parrinello, Canonical sampling through velocity rescaling, J. Chem. Phys. **126**, 014101 (2007).

[25] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Di Nola, and J. R. Haak, Molecular dynamics with coupling to an external bath, J. Chem. Phys. **81**, 3684 (1984).

[26] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije, LINCS: A linear constraint solver for molecular simulations, J. Comput. Chem. **18**, 1463 (1997).

[27] B. Hess, P-LINCS: A parallel linear constraint solver for molecular simulation, J. Chem. Theory Comput. **4**, 116 (2008).

[28] T. Darden, D. York, and L. Pedersen, Particle mesh Ewald: An N log(N) method for Ewald sums in large systems, J. Chem. Phys. **98**, 10089 (1993).

[29] K. Azizi, M. Gori, U. Morzan, A. Hassanali, and P. Kurian, Examining the origins of observed terahertz modes from an optically pumped atomistic model protein in aqueous solution, PNAS Nexus **2**, 1 (2023).

[30] A. Battisti, G. Ciasca, A. Grottesi, and A. Tenenbaum, Thermal compaction of the intrinsically disordered protein tau: Entropic, structural, and hydrophobic factors, Phys. Chem. Chem. Phys. **19**, 8435 (2017).

[31] R. G. Gordon, Correlation functions for molecular motion, Adv. Magn. Reson. **3**, 1 (1968).

[32] S. Hayward and N. Go, Collective variable description of native protein dynamics, Annu. Rev. Phys. Chem. **46**, 223 (1995).

[33] A. Kitao and N. Go, Investigating protein dynamics in collective coordinate space, Curr. Opin. Struct. Biol. **9**, 164 (1999).

[34] A. Amadei, A. B. M. Linssen, and H. J. C. Berendsen, Essential dynamics of proteins, Proteins Struct. Funct. Genet. **17**, 412 (1993).

[35] A. Amadei, A. B. M. Linssen, B. L. De Groot, and H. J. C. Berendsen, Essential degrees of freedom of proteins, in *Modelling of Biomolecular Structures and Mechanisms*, edited by A. Pullman, J. Jortner, and B. Pullman, The Jerusalem Symposia on Quantum Chemistry and Biochemistry Vol. 27 (Springer, Dordrecht, 1995).

[36] S. Hayward, A. Kitao, F. Hirata, and N. Go, Effect of solvent on collective motions in globular protein, J. Mol. Biol. **234**, 1207 (1993).

[37] S. Hayward, A. Kitao, and N. Go, Harmonic and anharmonic aspects in the dynamics of BPTI: A normal mode analysis and principal component analysis, Protein Sci. **3**, 936 (1994).

[38] A. Kitao, F. Hirata, and N. Go, The effects of solvent on the conformation and the collective motions of protein: Normal mode analysis and molecular dynamics simulations of melittin in water and in vacuum, Chem. Phys. **158**, 447 (1991).

[39] A. E. García, Large-amplitude nonlinear motions in proteins, Phys. Rev. Lett. **68**, 2696 (1992).

[40] O. Sushko, R. Dubrovka, and R. Donnan, Sub-terahertz spectroscopy reveals that proteins influence the properties of water at greater distances than previously detected, J. Chem. Phys. **142**, 055101 (2015).

[41] P. W. Fenimore, H. Frauenfelder, B. H. McMahon, and R. D. Young, Bulk-solvent and hydration-shell fluctuations, similar to α- and β-fluctuations in glasses, control protein motions and functions, Proc. Natl. Acad. Sci. USA **101**, 14408 (2004).

[42] J. E. Curtis, M. Tarek, and D. J. Tobias, Methyl group dynamics as a probe of the protein dynamical transition, J. Am. Chem. Soc. **126**, 15928 (2004).

[43] V. Conti Nibali, G. D'Angelo, A. Paciaroni, D. J. Tobias, and M. Tarek, On the coupling between the collective dynamics of proteins and their hydration water, J. Phys. Chem. Lett. **5**, 1181 (2014).

[44] E. D. Belega, M. N. Zakirov, A. I. Chulichkov, D. N. Trubnikov, and Y. V. Novakovskaya, Effective-mode analysis of the dynamics of weakly bound molecular systems by an example of hydrogen-bonded water clusters, Phys. Rev. A **107**, 032812 (2023).

[45] J. Guo, D. Cao, J. Chen, K. Bian, L.-M. Xu, E.-G. Wang, and Y. Jiang, Probing the intermolecular coupled vibrations in a water cluster with inelastic electron tunneling spectroscopy, J. Chem. Phys. **152**, 234301 (2020).

[46] N. Go, T. Noguti, and T. Nishikawa, Dynamics of a small globular protein in terms of low-frequency vibrational modes, Proc. Natl. Acad. Sci. USA **80**, 3696 (1983).

[47] K. Hinsen and G. R. Kneller, Solvent effects in the slow dynamics of proteins, Proteins Struct. Funct. Genet. **70**, 1235 (2008).