Science Advances
AAAS

Supplementary Materials for

Experimental evidence for long-distance electrodynamic intermolecular forces

Mathias Lechelon, Yoann Meriguet, Matteo Gori, Sandra Ruffenach, Ilaria Nardecchia, Elena Floriani, Dominique Coquillat, Frédéric Teppe, Sébastien Mailfert, Didier Marguet, Pierre Ferrier, Luca Varani, James Sturgis, Jeremie Torres*, Marco Pettini*

*Corresponding author. Email: jeremie.torres@umontpellier.fr (J.T.); marco.pettini@cpt.univ-mrs.fr (M.P.)

Published 16 February 2022, *Sci. Adv.* **8**, eabl5855 (2022)
DOI: 10.1126/sciadv.abl5855

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

Additional Data

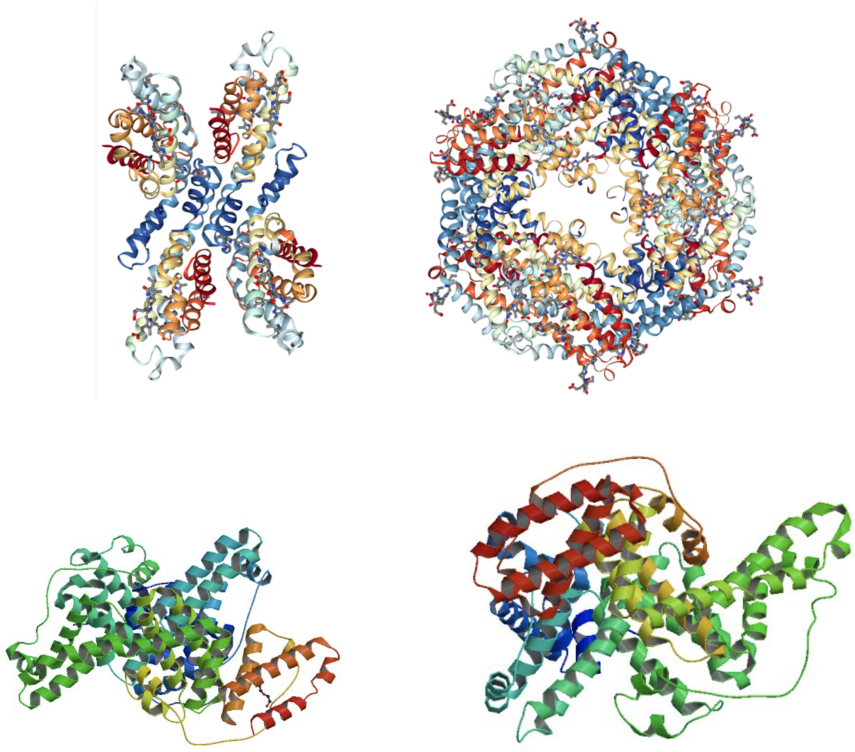

Figure S1: Images from Protein Data Bank. Top right, the assembly $\alpha_2\beta_2$ which forms a subunit (1EYX) of the R-PE protein. Top left, a unit of R-PE $(\alpha_2\beta_2)_3$. Bottom images: Two different projections of the BSA (3V03). It is clearly seen that both proteins are mainly made out of alpha-helices.

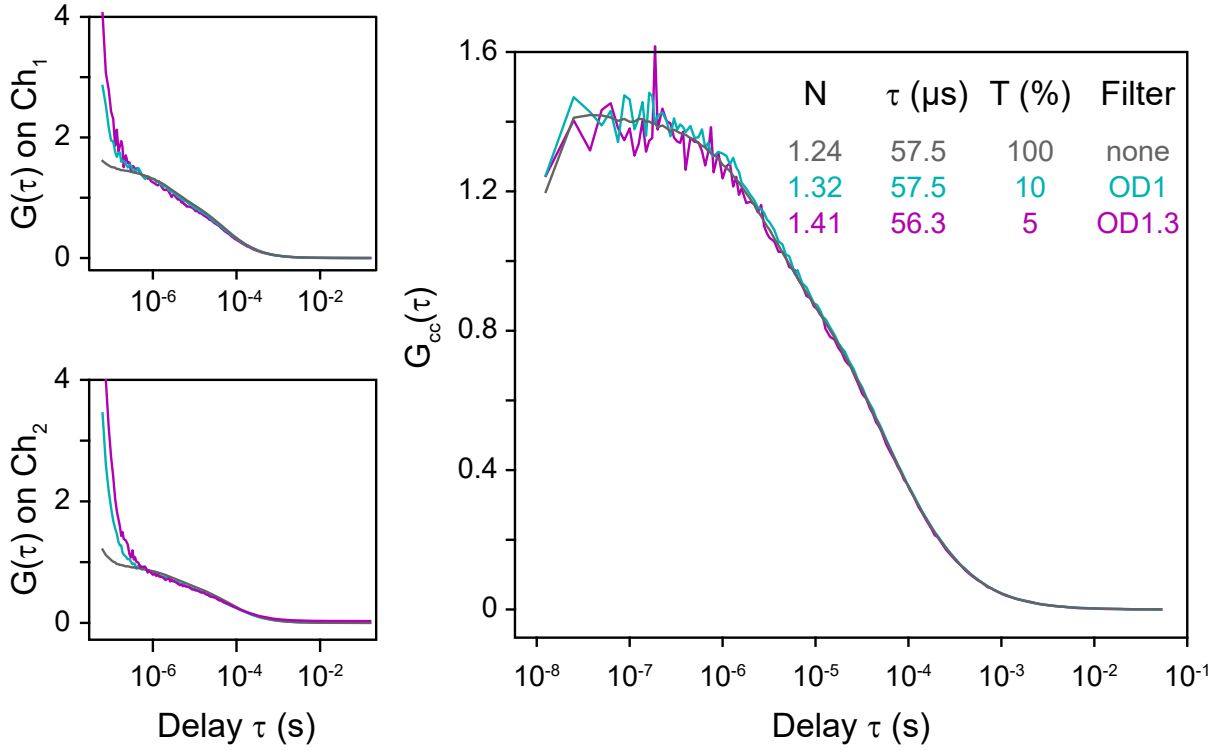

Figure S2: Cross-correlation Functions (CCFs). CCFs in the presence of micro-molar concentrations of fluorescent dyes obtained with density filters along the optical path combined with averaging during long acquisition times. The left panel shows the comparison of Cross-Correlation Functions (CCFs) of a single solution of the dye Atto 488 at the concentration of 1 nM, measured without OD (Optical Density) filter (no OD, black line) along the fluorescence path, with OD1 and with OD1.3 (cyan and magenta lines, respectively). Panels on the right show the Auto-Correlation Functions (ACFs) obtained with the signals of channels 1 and 2 for the same measurements. At short time-lag the ACFs clearly show a sharp increase, whereas this phenomenon is absent on CCFs. The phenomenon is consequently an artifact attributed to afterpulsing. Each CCF was obtained by averaging 72 CCFs, each one obtained from a record of 50 seconds, corresponding to an overall acquisition time of one hour.

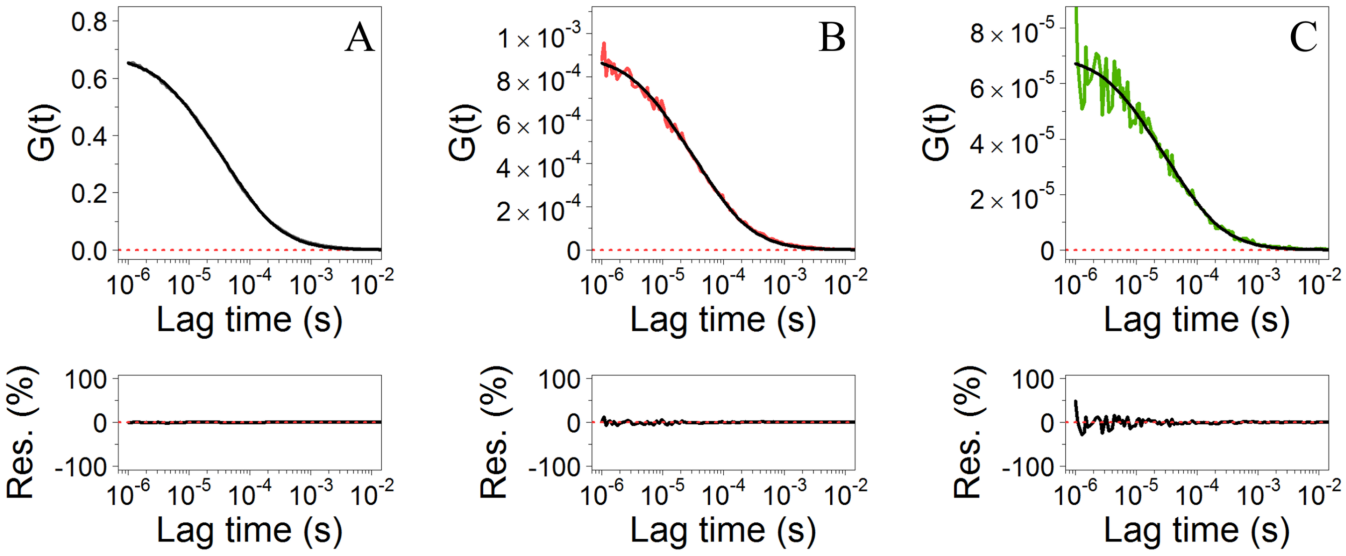

Figure S3: FCCS results at high concentrations of fluorescent molecules and using density filters. Panels A to C display the CCFs - as shown in Figure S2 - corresponding to solutions of 1 nM, 1 μM, and 10 μM of Atto 488 dye using no OD filter, OD1, and OD1.3 filters, respectively. Each CCF was obtained by averaging 72 CCFs, each one obtained from a record of 50 seconds, corresponding to an overall acquisition time of one hour. The fitting curves are also displayed in black on each panel, and the residuals are visible below each CCF (displayed in percentage). The same values of τ reported in Figure S2 are found within a 2% deviation.

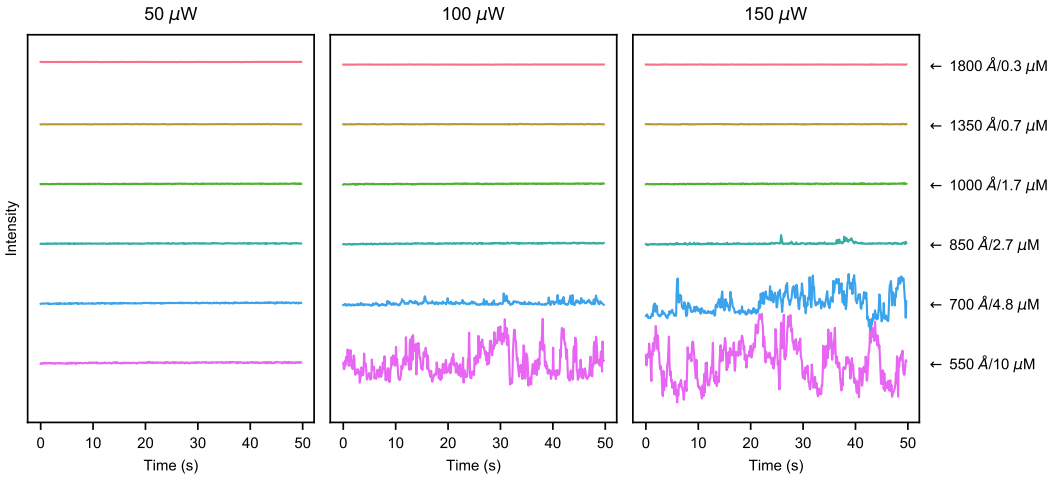

Figure S4: Traces recorded by channel 1 of the FCCS device during 60 seconds, for solutions of R-PE. The average distances among the proteins are, from top to bottom, respectively: 1800 Å, 1350 Å, 1000 Å, 850 Å, 700 Å, 550 Å. The left panel shows the results for the laser output power set at 50 μW, middle panel for the laser power set at 100μW, and right panel for the laser power set at 150 μW. The sharp increase of fluorescence fluctuations at increasing laser power and protein concentration is due to the formation of clusters as shown by the on-line Videos.

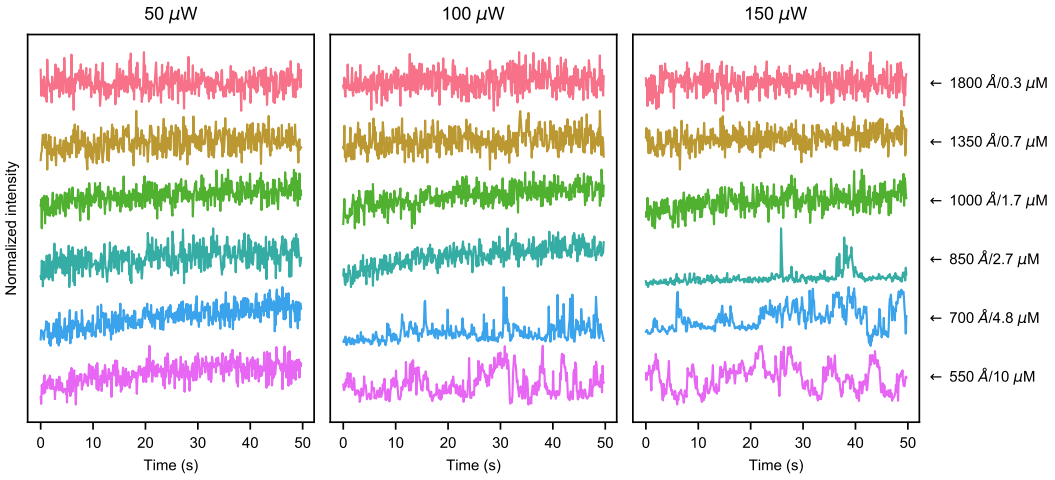

Figure S5: Normalized traces. The same outcomes of Additional Data Figure S4 are here shown by normalizing the data to the highest fluorescence value recorded for each trace. Fluorescence traces recorded by channel 1 of the FCCS device during 60 seconds, for solutions of R-PE. The traces are normalized by the highest value of each trace.

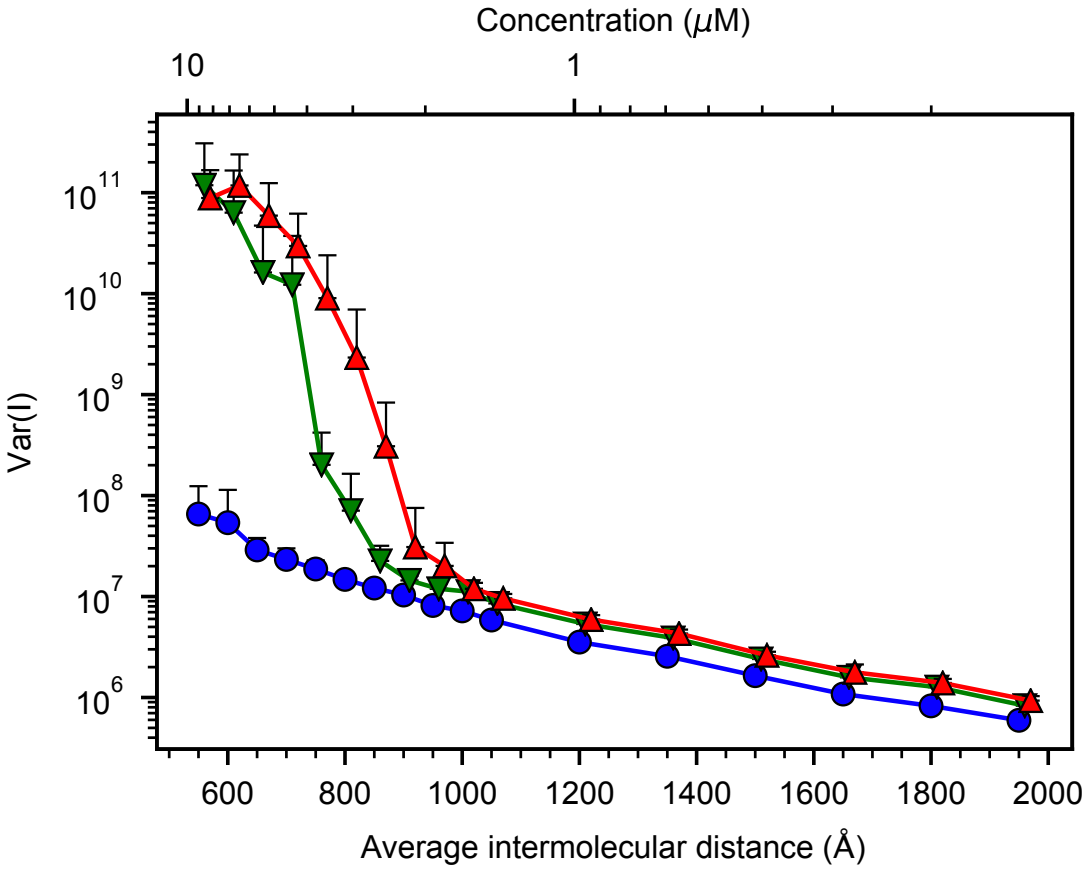

Figure S6: Variance of the fluorescence intensity. Data recorded for R-PE solutions at different laser power input: 50 μW (blue circles), 100 μW (green down-pointing triangles), and 150 μW (red up-pointing triangles). Only upper error bars are shown on this plot for convenience, as the y axis is on a logarithmic scale.

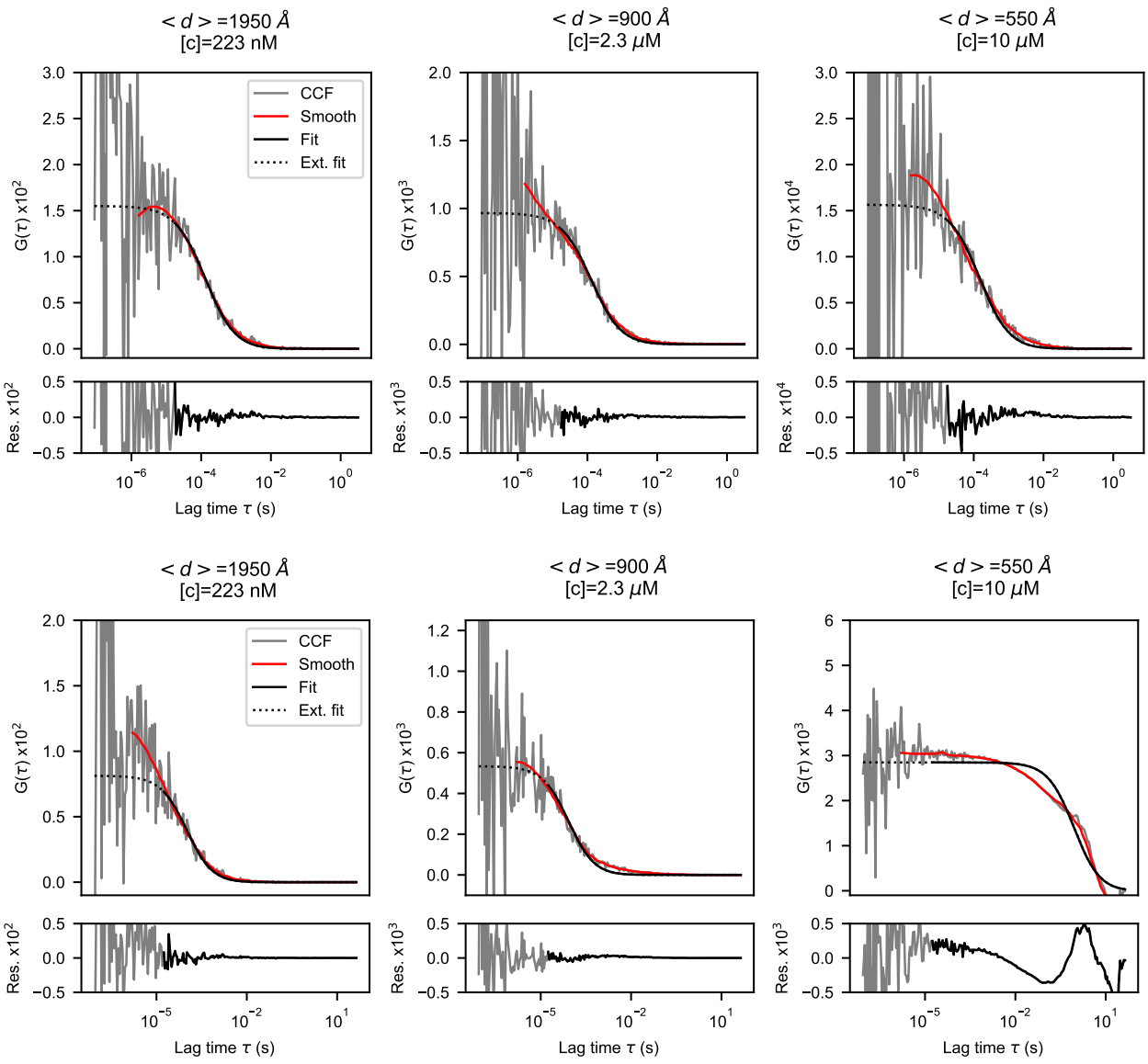

Figure S7: Cross-Correlation Functions (CCFs) for R-PE solutions at different concentrations. From left to right the average intermolecular distances are: $\langle d \rangle = 1950$ Å, $\langle d \rangle = 900$ Å, $\langle d \rangle = 550$ Å, respectively. The upper panels display the results obtained with a laser power input of 50 μW (grey lines). The lower panels display the results obtained with a laser power input of 150 μW (grey lines). A Savitzky-Golay filter is used to smooth the CCFs (red lines), and to consequently obtain the diffusion time τ_D at Half Height (THH). Fits of the CCF are also shown (black lines), and the extended fits (outside of the original fitting range) are represented with dashed lines. Residuals of the fits are also plotted: the black lines correspond to the standard residuals, and the grey lines to the extended residuals. At the highest concentration, for $\langle d \rangle = 550$Å, and laser power input of 150 μW, the correlation time is increased by several orders of magnitude, entailing a drop of the diffusion coefficient by the same amount.

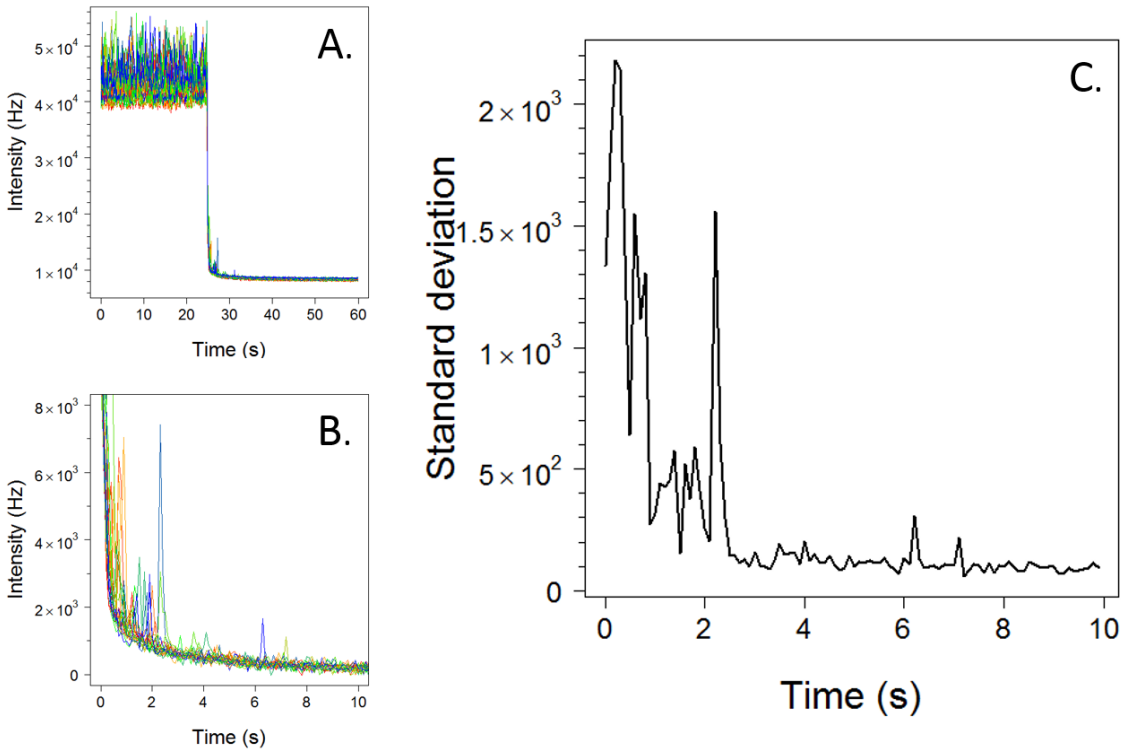

Figure S8: Decay of the clusters of R-PE induced by a sudden lowering of the laser power density. The results reported here have been obtained using ten samples corresponding to an average intermolecular distance of 650 $\mathring{A}$. On panel A, the laser input has been lowered from 150 μW to 50 μW at $t = 25\ s$. Panel B is a zoom of panel A, starting at $t = 25\ s$. Panel C is the standard deviation of the intensity of the different samples recorded and displayed by panels A and B. Here the thermal fluctuations overcome the attractive electrodynamic forces weakened by the lowering of the energy input rate, thus destroying the clusters. The initial time $t = 0$ has been arbitrarily chosen in the already clustered phase.

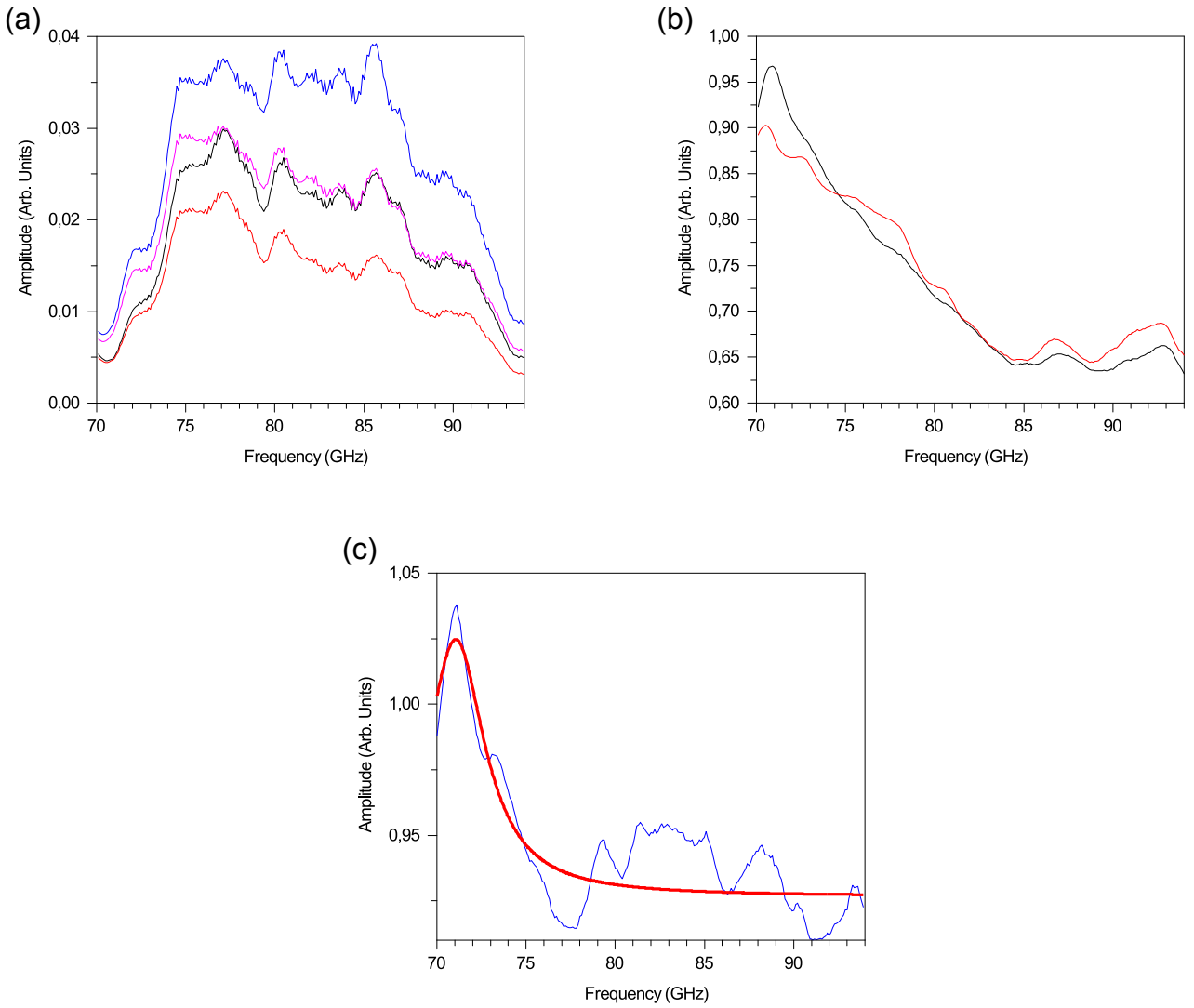

Figure S9: Data analysis procedure for rectenna-based THz-spectroscopy. (a) Spectra for NaCl-water solution and R-PE-NaCl-water solution when the laser is switched OFF (black and blue lines, respectively) and when the laser is switched ON (red and pink lines, respectively). (b) Normalization procedure, according to the Beer-Lambert law (see Main text Eq. (5)). In red (resp. black) the result is plotted which is obtained by dividing the spectra of proteins (resp. NaCl) in solution under illumination by the spectra of proteins (resp. NaCl) in solution without illumination. (c) Ratio of two previous spectra of (b) and Lorentzian fit. Only the fundamental mode at 0.071 THz is observed by operating with laser power at 50 mW, and an overall exposure time of 96 seconds (240 points are here reported, each one obtained with an exposure time of 400 ms).

Captions for Movies S1 - S4

Movie S1 : R-PE protein concentration 0.5 micro Moles (absence of protein clusters). No optical filter. Each 128 x 128 pixel confocal image corresponds to a 30 x 30 μm field of view and recorded at a pixel dwell time of 0.1 ms at a rate of 2.8 s/frame. Each stack of images was made of 375 images where the laser power has been adjusted every 75 images to 50, 150, 50, 150 and then 50 μW. When indicated, an OD1 filter (Thorlabs ND-10B) has been added before the detector. Then, the video made from these images were compressed at a rate of 7 frame/s. Thus, each second of video corresponds to about 21 seconds of measurements.

Movie S2 : R-PE protein concentration 0.5 micro Moles (absence of protein clusters). Each 128 x 128 pixel confocal image correspond to a 30 x 30 μm field of view and recorded at a pixel dwell time of 0.1 ms at a rate of 2.8 s/frame. Each stack of images was made of 375 images where the laser power has been adjusted every 75 images to 50, 150, 50, 150 and then 50 μW. An OD1 filter (Thorlabs ND-10B) has been added before the detector. Then, the video made from these images were compressed at a rate of 7 frame/s. Thus, each second of video corresponds to about 21 seconds of measurements.

Movie S3 : R-PE protein concentration of 2.9 micro Moles (protein clusters are observed). Each 128 x 128 pixel confocal image correspond to a 30 x 30 μm field of view and recorded at a pixel dwell time of 0.1 ms at a rate of 2.8 s/frame. Each stack of images was made of 375 images where the laser power has been adjusted every 75 images to 50, 150, 50, 150 and then 50 μW. An OD1 filter (Thorlabs ND-10B) has been added before the detector. Then, the video made from these images were compressed at a rate of 7 frame/s. Thus, each second of video corresponds to about 21 seconds of measurements.

Movie S4 : R-PE protein concentration 6.5 micro Moles (protein clusters are observed). Each 128 x 128 pixel confocal image correspond to a 30 x 30 μm field of view and recorded at a pixel dwell time of 0.1 ms at a rate of 2.8 s/frame. Each stack of images was made of 375

images where the laser power has been adjusted every 75 images to 50, 150, 50, 150 and then 50 μW. An OD1 filter (Thorlabs ND-10B) has been added before the detector. Then, the video made from these images were compressed at a rate of 7 frame/s. Thus, each second of video corresponds to about 21 seconds of measurements.

1 Theoretical interpretation of the experimental results

In this part we present the theoretical and numerical models that provide a conceptual framework for the interpretation of the experimental results. Throughout the following sections, it is implicitly assumed that the biomolecules of interest have undergone the condensation phenomenon of polar vibrations into the lowest frequency mode under external energy supply as it has been observed in Ref.[19] of Main text. This phonon condensation results in a collective intramolecular vibration bringing about a large coherent oscillating dipole moment that activates long-range electrodynamic intermolecular forces (Ref.[12] of Main text).

2 Effects of electrodynamic interactions on molecular diffusion

In this section we introduce three different theoretical and numerical models in order to interpret the FCCS experiment on the R-PE protein that undergoes a clustering transition. We will show how the experimental outcomes can be explained by the presence of long range dipole-dipole electrodynamic intermolecular interactions.

Some simplifications are assumed to define a model that can be studied at least semi-analytically. We are interested in the behaviour determined by long-range interactions for values of the R-PE concentration in solution such that the intermolecular average distance is much larger than the characteristic size of the molecule. For this reason the particles are considered as spherical in our models. All the spheres are assumed to have the same radius. The radius of the spherical particles has been chosen such that their volume is equivalent to the volume $\mathcal{V}_{RPE}$ of the R-PE estimated from the mass $M_{RPE} \approx 2.5\times10^2$D,

$$a = \left(\frac{3}{4\pi}\frac{M_{RPE}}{\rho_{BM}}\right)^{1/3} \simeq 42\mathring{A} \tag{1}$$

where the density of the biomolecule has been set to $\rho_{BM} = 1.27\ g\,cm^{-3}$ (Refs.[48] and [49] of Main text).

2.1 The effective potential

The interaction potential among resonant oscillating electric dipoles is supposed to take the effective form

$$U_{\text{eff}}(r_{ij}) = -\frac{3c_{\text{eff}}k_BTR_0^3}{r_{ij}^3 + 2R_0^3} \tag{2}$$

where $r_{ij} = \|\boldsymbol{r}_j - \boldsymbol{r}_i\|$ is the distance between the i-th and j-th molecules, k_B is the Boltzmann constant, T is the temperature and R_0 is a length scale. The form of the potential in Eq.(2) has been chosen such that

$$U_{\text{eff}}(R_0) = c_{\text{eff}}k_BT \quad \text{and} \quad U''_{\text{eff}}(R_0) = 0\ . \tag{3}$$

The regularization in the denominator has been introduced in order to avoid the divergence in $\|\mathbf{r}_i - \mathbf{r}_j\| = 0$. The parameter R_0 has been chosen to be the sum of two molecular radii $R_0 = 2a = 84\mathring{A}$. The strength of the potential in k_BT units at the distance R_0 between the particle centers is given by the parameter c_{eff} which can hardly be assessed *a priori*. The dynamical electric dipole moment can be estimated by equating the effective potential with the quasi-static dipole-dipole interaction energy

$$U_{dd}(r_{ij}) \simeq -\frac{p^2}{4\pi\varepsilon_0\varepsilon_W(\omega)r_{ij}^3} \tag{4}$$

at the distance $r_{ij} = r_*$, i.e. $U_{dd}(r_*) = U_{\text{eff}}(r_*)$. The relative dielectric constant of the medium is assumed to be $\varepsilon_W(\omega) = |\varepsilon(\omega_{CVM})|$, where the suffix CVM stands for Collective Vibrational Motion, the relation between the strength of the potential and the electric dipole moment of the

biomolecules is given by $p = \alpha(r_*, T)c_{\text{eff}}^{1/2}$, where the calibration constant α depends on the distance r_* at which the two potentials are set equal and T is the temperature. The results for different choices of T and r_* are reported in Table S1 for R-PE ($\nu_{CVM} = 71\text{GHz}$).

$r_*[\mathring{A}]$	$T[K]$	$\alpha[D]$
82	293	7.10×10^2
950	293	1.23×10^3
82	303	1.25×10^3
950	303	1.30×10^3

Table S1: Different values of the factor α.

2.2 Semi-Analytical Model

The discussion of this semi-analytical model is adapted from an analogous calculation for the clustering transition in a self-gravitating system (Ref.[50] of Main text) where the long-range attractive interaction potential among particles scales with the distance r as r^{-1}. The model developed in this section aims at verifying the existence of a clustering transition in the FCCS experiments when the translation degrees of freedom of the molecules are at thermal equilibrium and the interparticle electrodynamic interaction potential scales as r^{-3}. The experimental setting is represented in a simplified way as a system of N spherical particles of radius $a = 42\mathring{A}$ confined in a sphere of radius R of volume equal to the FCCS confocal effective volume as defined by

$$V_{\text{eff}} = (2)^{3/2} V_{\text{conf}} = \pi^{3/2} w_0^2 z_0 \ . \tag{5}$$

In real experiments the two sizes w_0 and z_0 have been estimated to be $w_0 = 2.8{\times}10^3\mathring{A}$ and $z_0 = 5w_0 = 1.4{\times}10^4\mathring{A}$, yielding a volume-equivalent spherical system of radius

$$R = \left(\frac{3\pi^{1/2}}{4} w_0^2 z_0\right)^{1/3} \simeq 5.265{\times}10^3\mathring{A} \tag{6}$$

The configuration of a system of N particles is in principle described by the probability density function in configuration space $\rho_N(\mathbf{r}_1, ..., \mathbf{r}_N)$. In the simplified model here considered, the volume, the temperature and the number of molecules inside the volume are fixed so that the thermodynamic equilibrium is defined as the probability distribution that minimizes the free energy functional $F[\rho_N](N, V, T)$

$$F[\rho_N](N, V, T) = U[\rho_N] - TS[\rho_N] \ . \tag{7}$$

where $U[\rho_N]$ is the average potential energy

$$U[\rho_N] = \int_{\mathcal{S}^3(R)} \mathrm{d}^3\mathbf{r}_1 ... \int_{\mathcal{S}^3(R)} \mathrm{d}^3\mathbf{r}_N \ \ U(\mathbf{r}_1, ...\mathbf{r}_N)\rho_N(\mathbf{r}_1, ..., \mathbf{r}_N) \tag{8}$$

while the entropy $S[\rho_N]$ is defined as

$$S[\rho_N] = -\int_{\mathcal{S}^3(R)} \mathrm{d}^3\mathbf{r}_1 ... \int_{\mathcal{S}^3(R)} \mathrm{d}^3\mathbf{r}_N \ \ \rho_N(\mathbf{r}_1, ..., \mathbf{r}_N) \log\left[\rho_N(\mathbf{r}_1, ..., \mathbf{r}_N)\right] \ . \tag{9}$$

The distribution $\rho_{N,\mathrm{eq}}$ such that $\delta F[\rho_{N,\mathrm{eq}}]/\delta\rho_N = 0$ and $\delta^2 F[\rho_{N,\mathrm{eq}}]/\delta\rho_N^2 > 0$ corresponds to the probability distribution at thermodynamic equilibrium in the canonical ensemble. We point out that such a description does not coincide with the real experimental setting because in FCCS experiments the particles can freely enter into and exit out of the confocal volume, so that the hypothesis of a fixed number of particles holds only in the average. The control parameter in real experiments is the concentration of biomolecules, or, equivalently, the *intermolecular average distance* $\langle d \rangle$: in our model this corresponds to a different choice of the number of molecules in the total volume

$$N(\langle d \rangle) = \left\lfloor \frac{4\pi}{3} \left(\frac{R}{\langle d \rangle} \right)^3 \right\rfloor \tag{10}$$

where $\lfloor \cdot \rfloor$ is the floor operator. It is convenient to introduce adimensionalized quantities, choosing the effective radius as the length scale of the model so that $\widetilde{R} = R/a \simeq 1.25 \times 10^2$.

We consider a *mean field approximation* for the probability density distribution of the biomolecules in the volume, i.e.

$$\rho_N(\mathbf{r}_1, ..., \mathbf{r}_N) = \prod_{i=1}^{N} \rho_1(\mathbf{r}_i) \tag{11}$$

We make the following ansatz about the functional form of $\rho_1(\mathbf{r})$, i.e.

$$\rho_1(\widetilde{\mathbf{r}}, \eta) = \eta\Theta(\widetilde{R}_*(\eta) - \|\widetilde{\mathbf{r}}\|)\rho_c(\mathbf{r}) + (1-\eta)\Theta(\|\widetilde{\mathbf{r}}\| - \widetilde{R}_*(\eta))\Theta(\widetilde{R} - \|\widetilde{\mathbf{r}}\|)\rho_h(\mathbf{r}) =$$
$$= \eta\Theta(\widetilde{R}_*(\eta) - \|\widetilde{\mathbf{r}}\|)\frac{3}{4\pi\widetilde{R}_*^3(\eta)} + (1-\eta)\Theta(\|\widetilde{\mathbf{r}}\| - \widetilde{R}_*(\eta))\Theta(\widetilde{R} - \|\widetilde{\mathbf{r}}\|)\frac{3}{4\pi\left[\widetilde{R}^3 - \widetilde{R}_*^3(\eta)\right]}$$

which means that a particle is found with probability η in a spherical cluster of radius R_* located at the center of the ambient spherical space. The radius of the cluster is determined by considering the random close packing for spheres $g = 0.637$ (maximum volume fraction for a randomly packed 3D system) such that

$$\widetilde{R}_*(\eta, \langle d \rangle) = R_*/a = \sqrt[3]{g^{-1}\eta N(\langle d \rangle)} \tag{12}$$

where a is the effective radius of the considered molecules in the cluster, taking into account all short range interactions (dispersive forces, hydrodynamic forces, etc.). In what follows, we denote by a tilde superscript the adimensionalized length expressed in units of the particles radius a.

Remark. *The mean field approximation and the form of the one-particle distribution in Eq.(12) is a sort of "fluid approximation": the more the volume of the particles is negligible with respect to the total volume, the better the approximation.*

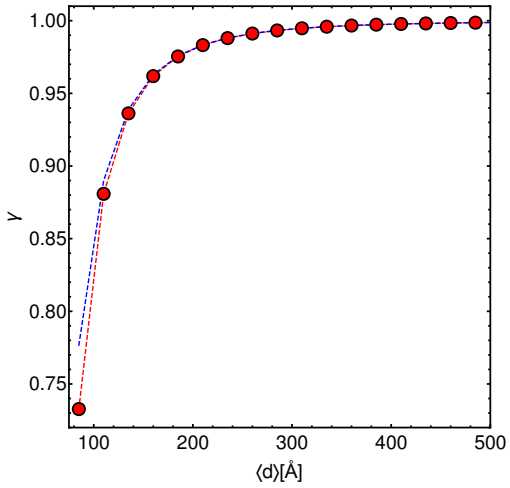

Figure S10: Dependence of effective accessible volume. Effective accessible volume γ, for intermolecular distances $\langle d \rangle$ in the range $90 \div 600 \mathring{A}$. In red circles the numerical values of γ calculated using the exact expression in Eq.(13), the blue line representing the data obtained using the approximated function. In this case, the effective radius of the molecule is $a = 42 \mathring{A}$ and the radius relative to the total volume $R = 5246 \mathring{A}$.

An estimation of the excluded volume effects is provided by the following parameter

$$\gamma(\langle \tilde{d} \rangle, \widetilde{R}) = \left\{ \frac{\prod_{i=1}^{N(\langle \tilde{d} \rangle, \tilde{R})} [V - (i-1)4\pi a^3/3]}{V^{N(\langle \tilde{d} \rangle, \tilde{R})}} \right\}^{\frac{1}{N(\langle \tilde{d} \rangle, \tilde{R})}} = \sqrt[N(\langle \tilde{d} \rangle, \tilde{R})]{\prod_{i=1}^{N(\langle \tilde{d} \rangle, \tilde{R})} \left[1 - (i-1)\widetilde{R}^{-3}\right]} \tag{13}$$

In the limit of large $\widetilde{R}, \gamma$ is a function only of the (adimensionalized) intermolecular average distance, in fact

$$\begin{aligned} \log \gamma = & \quad \frac{1}{N(\langle \tilde{d} \rangle, \tilde{R})} \sum_{i=1}^{N(\langle \tilde{d} \rangle, \tilde{R})} \log \left[1 - (i-1)\widetilde{R}^{-3}\right] \approx -\frac{\widetilde{R}^{-3}}{N(\langle \tilde{d} \rangle, \tilde{R})} \sum_{i=1}^{N(\langle \tilde{d} \rangle, \tilde{R})} (i-1) \\ = & \quad \frac{1}{2} \frac{\left[N(\langle \tilde{d} \rangle, \tilde{R}) - 1\right]}{\widetilde{R}^3} \approx -\frac{2\pi}{3\langle \tilde{d} \rangle^3} + \frac{1}{2\tilde{R}^3} \approx -\frac{2\pi}{3\langle \tilde{d} \rangle^3} = \log \gamma_{app} \ . \end{aligned} \tag{14}$$

As a first approximation, we can assume that in the FCCS experiments - reported in the main

text - the degrees of freedom relative to the centers of the biomolecules are at thermal equilibrium, and that the number of particles in the confocal volume is constant. These assumptions allow to tackle the system of interacting biomolecules in the canonical ensemble framework. The adimensional specific free energy functional is defined as

$$\frac{F[\rho_N]}{Nk_BT} = \frac{U[\rho_N]}{Nk_BT} - \frac{S[\rho_N]}{k_BN} = u[\rho_N] - s[\rho_N] \ . \tag{15}$$

The (adimensional) specific entropy $s[\rho_N]$ is given by

$$s[\rho_N] = \frac{S[\rho_N]}{Nk_B} = -\frac{1}{N}\int_{\mathbb{S}^3(\widetilde{R})} \rho_N \ \log\rho_N \ \prod_{i=1}^{N} \mathrm{d}r_i^3 \ . \tag{16}$$

and substituting Eq.(12) in the previous expression we obtain

$$s(\eta, \langle d\rangle) = s[\rho_N] = -\left[\eta \log\left(\frac{3\eta}{4\pi\widetilde{R}_*^3(\eta,\langle d\rangle)}\right) + (1-\eta)\log\left(\frac{3(1-\eta)}{4\pi(\widetilde{R}^3 - \widetilde{R}_*^3(\eta,\langle d\rangle))}\right)\right] \ . \tag{17}$$

The expected specific potential energy $u[\rho_N]$ is given by three separate contributions:

$$u[\rho_N] = u_{\mathrm{c-c}}[\rho_N] + u_{\mathrm{c-h}}[\rho_N] + u_{\mathrm{h-h}}[\rho_N] \tag{18}$$

the self-interaction of the molecules contained in the central cluster

$$u_{c-c}(\eta,\langle d\rangle) = u_{c-c}[\rho_N] = -\eta^2\frac{N(\langle d\rangle)-1}{2}\int_{\mathcal{S}(\widetilde{R}_*)}\mathrm{d}^3\widetilde{r}_i\int_{\mathcal{S}(\widetilde{R}_*)}\mathrm{d}^3\widetilde{r}_j \ \rho_c(\widetilde{\mathbf{r}}_i)\rho_c(\widetilde{\mathbf{r}}_j)\frac{24c_{\mathrm{eff}}}{\|\widetilde{\mathbf{r}}_i - \widetilde{\mathbf{r}}_j\|^3 + 16}, \tag{19}$$

the interaction among the halo and the cluster

$$u_{c-h}(\eta, \langle d \rangle) = u_{c-h}[\rho_N] = \tag{20}$$
$$= -\eta(1-\eta)\frac{N(\langle d \rangle)-1}{2} \times \int_{\mathcal{S}(\widetilde{R}_*)} \mathrm{d}^3\widetilde{r}_i \int_{\mathcal{C}(\widetilde{R},\widetilde{R}_*)} \mathrm{d}^3\widetilde{r}_j \ \rho_c(\widetilde{\mathbf{r}}_i)\rho_h(\widetilde{\mathbf{r}}_j)\frac{24c_{\text{eff}}}{\|\widetilde{\mathbf{r}}_i - \widetilde{\mathbf{r}}_j\|^3 + 16},$$

and the self-interaction of the halo

$$u_{h-h}(\eta, \langle d \rangle) = u_{h-h}[\rho_N] = \tag{21}$$
$$= -(1-\eta)^2\frac{N(\langle d \rangle)-1}{2} \times \int_{\mathcal{C}(\widetilde{R},\widetilde{R}_*)} \mathrm{d}^3\widetilde{r}_i \int_{\mathcal{C}(\widetilde{R},\widetilde{R}_*)} \mathrm{d}^3\widetilde{r}_j \ \rho_h(\widetilde{\mathbf{r}}_i)\rho_h(\widetilde{\mathbf{r}}_j)\frac{24c_{\text{eff}}}{\|\widetilde{\mathbf{r}}_i - \widetilde{\mathbf{r}}_j\|^3 + 16} \ .$$

where $\mathcal{S}(x)$ is a sphere with radius x centered at the origin and $\mathcal{C}(x, x')$ is a spherical shell centered at the origin and with external and internal radii x and x' respectively. The fraction of clustered molecules at equilibrium is obtained by minimizing the specific free energy with respect to η at fixed $\langle d \rangle$, i.e.

$$\eta_{min}(\langle d \rangle) = \left\{ \eta \in [0,1] \,|\, F(\eta_{min}) = \min_{\eta \in [0,1]} \frac{F(\eta, \langle d \rangle)}{N(\langle d \rangle) k_B T} \quad \text{for any fixed } \langle \mathrm{d} \rangle \right\} \ . \tag{22}$$

As it can be observed in Figure S11, there exists a range of values of the effective dipole-dipole potential strength c_{eff} such that the specific free energy $F/(Nk_BT)$ has a relative minimum for a certain value of the clustered fraction $\eta' \in [0,1]$, for any fixed value of the intermolecular average distance $\langle d \rangle$.

The thermodynamic equilibrium configurations are attained at the global minimum of the free energy. Being the clustering fraction η an order parameter defined on a compact domain, such a global minimum of the free energy can be attained both at $\eta = 0$ or inside the domain if a relative minimum of the free energy exists at η_c such that $F(\eta_c) < F(0)$.

In the considered system, the value of the relative minimum of the specific free energy decreases

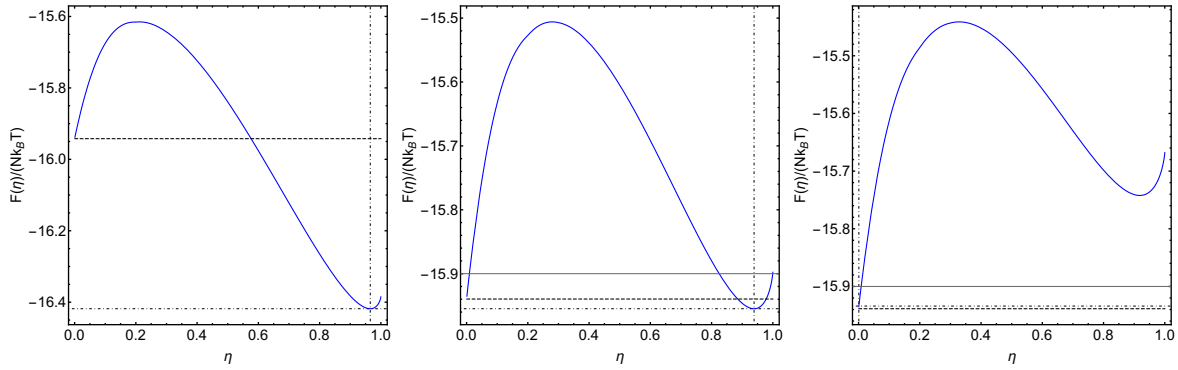

Figure S11: Specific free energy dependence with cluster fraction. Specific free energy $F/(k_BTN)$ vs. clustered fraction η for different values of the intermolecular average distance $\langle d\rangle$ at $c_{\text{eff}} = 0.325$. Left panel : $\langle d\rangle = 900\mathring{A}$; center panel: $\langle d\rangle = 950\mathring{A}$; right panel: $\langle d\rangle = 975\mathring{A}$.

with the intermolecular average distance $\langle d\rangle$. This means that it is possible to find a value of the parameter c_{eff} such that $F(\eta') < F(0)$ for $\langle d\rangle < 950$ Å.

In our case, we have empirically found that by setting $c_{\text{eff}} = 0.325$, it is $F(\eta') \lesssim F(0)$ for $\langle d\rangle \lesssim 950\mathring{A}$ (see Figure S11). Moreover, almost all the molecules are in the clustered phase as $\eta' = \eta_{min} > 0.93$ for $c_{\text{eff}} = 0.325$ and $\langle d\rangle \lesssim 950\mathring{A}$ (see Figure S11), while for $\langle d\rangle \gtrsim 975\mathring{A}$ the equilibrium of the system is in the disperse phase. We conclude that long range attractive interactions scaling as r^{-3} with the intermolecular distance r can induce a clustering phase transition at least in qualitative agreement with the experimentally observed transition. According to Table S1 the estimated value of the dynamic dielectric dipole is $|\mathbf{p}| = 400 \div 740$D. We remark that such a value constitutes an underestimation of the real dynamical dielectric dipole because the electrodynamic interactions can be either attractive or repulsive, and this depends on the mutual oscillation phase and orientation of the dipoles, whereas in the mean field model considered in this section, the interactions are always attractive.

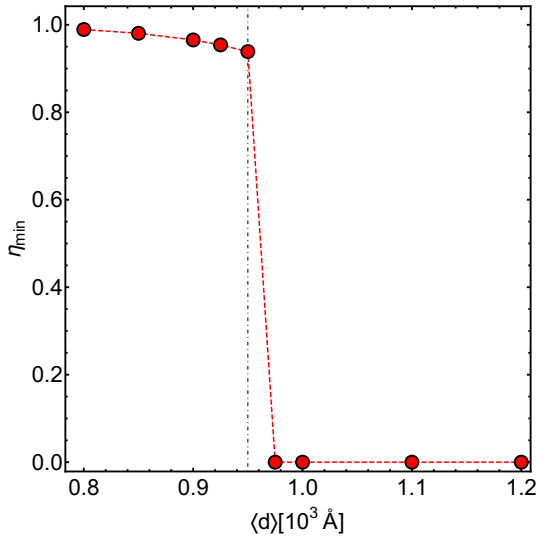

Figure S12: Specific free energy dependence with intermolecular average distance. Clustered fraction η as a function of the intermolecular average distance $\langle d \rangle$ for a system of particles interacting through the potential in Eq.(2) with $c_{\text{eff}} = 0.325$. $\eta_{min} \simeq 1$ corresponds to the clustered phase whereas $\eta_{min} \simeq 0$ corresponds to the dispersed (gaseous) phase.

2.3 Molecular Dynamics Study of the clustering transition

Molecular dynamics simulations have been done in order to estimate the effect of long-range electrodynamic interactions, described by the effective potential in Eq.(2), on the self-diffusion coefficient D of a system of interacting molecules defined by

$$D = \lim_{t \to +\infty} \frac{\langle \|\Delta \mathbf{r}_i(t)\|^2 \rangle_i}{6t} \tag{23}$$

where $\Delta \mathbf{r}_i(t)$ is the displacement at time t of the i-th molecule with respect to its initial position, and $\langle \cdot \rangle_i$ is the average over all the particles in the system. This is the physical quantity measured by means of FCCS experiments thus allowing a direct comparison between the outcomes of numerical simulations and the outcomes of lab experiments.

We considered a system made of a fixed cubic box of volume equal to the effective volume in Eq.(5), hence of side $L = (V_{\text{eff}})^{1/3} = 8486 \mathring{A}$. Given an intermolecular distance $\langle d \rangle$, the

corresponding number of particles in the box is determined through Eq.(10).

The dynamics is given by the Langevin equations in the overdamped limit (without inertial terms):

$$\frac{\mathrm{d}\mathbf{r}_i}{\mathrm{d}t} = -\frac{1}{\gamma}\nabla_{\mathbf{r}_i}\sum_{j\neq i} U(\|\mathbf{r}_i - \mathbf{r}_j\|) + \sqrt{\frac{2k_BT}{\gamma}}\boldsymbol{\xi}_i(t) \qquad \forall i = 1, .., N \tag{24}$$

where γ is the viscous friction constant, T is the temperature of the solution and $\boldsymbol{\xi}_i(s)$ is a noise term, s.t.

$$\langle \xi_{A,i}(t)\rangle_t = 0 \quad \langle \xi_{A,i}(t)\xi_{B,j}(t')\rangle_t = \delta(t-t')\delta_{AB}\delta_{i,j} \qquad \forall i,j = 1, ..., N \quad \forall A, B = 1, ..., 3\ . \tag{25}$$

As we have assumed that the molecules are represented by spherical particles, we use Stokes' formula for the viscous friction constant of a sphere in a viscous fluid, i.e.

$$\gamma = \frac{1}{6\pi R_H \eta_W(T)} \tag{26}$$

where R_H is the hydrodynamic radius in water and $\eta_W(T)$ is the dynamical viscosity of water at the temperature T. As we are more interested in a qualitative rather than a quantitative description of the diffusive dynamics and of the clustering transition, we can reasonably assume that $R_H = a = 42\mathring{A}$ while the water dynamical viscosity as a function of temperature at the atmospheric pressure is estimated using the following formula (Ref.[51] of Main text)

$$\eta_W(T) = A\exp[B/(T-C)] \tag{27}$$

where $A = 2.407{\times}10^{-5} Pa\cdot s$, $B = 571.5$K and $C = 139.7$K. The water dynamical viscosity is $\eta(T_{exp}) = 0.7915{\times}10^{-3}\ Pa\cdot s$ at the experimental value $T_{exp} = 30^oC = 303.15$K, according

to Eq.(27). The potential energy used for these molecular dynamics simulations is given by

$$U(r) = \begin{cases} U_{\text{eff}}(r) + U_{\text{Debye}}(r) & r > 2(1.01)a \\ U_{SC}(r) & r \leq 2(1.01)a \end{cases} \tag{28}$$

where U_{eff} is the effective potential energy of Eq.(2) due to electrodynamic interactions. Short range electrostatic repulsion is described through the Debye potential U_{Debye} due to electrostatic interactions

$$U_{\text{Debye}}(r) = \frac{Z^2 e^2 \exp\left[-\frac{r}{\lambda_D}\right]}{4\pi\varepsilon_0\varepsilon_W(1 + R/\lambda_D)r} \,. \tag{29}$$

Here Z is the net charge of the molecule, e is the electric charge of the electron, ε_0 is the vacuum permittivity, ε_W is the water relative dielectric constant and λ_D is the Debye length

$$\lambda_D = \left(\frac{\sum_i \rho_{\infty,i} e^2 z_i^2}{\varepsilon_0\varepsilon_W k_B T}\right)^{-1/2} , \tag{30}$$

where $\rho_{\infty,i}$ is the concentration of the i-th electrolyte species and T is the temperature of the solution. In our case, this results in a Debye length $\lambda_D \approx 9.74\mathring{A}$. The net charge for R-PE molecules has been fixed to $Z = +10$.

The effect of Pauli's repulsion among the electronic clouds of molecules is described through a Buckingham-like soft core potential U_{SC}

$$U_{SC}(R) = A_{SC} \exp\left(-\frac{r}{\lambda_{\text{B}}}\right) \tag{31}$$

where $\lambda_{\text{B}} = 2a$. The parameter A_{SC} fixing the strength of the potential has been chosen such that if the molecules overlap for the 10% of their radii they are brought back to be tangent to one another, i.e.

$$\|\Delta_{SC}\mathbf{x}_1 + \Delta_{SC}\mathbf{x}_2| = \frac{2\|\mathbf{F}_{sc}(1.9\,a)\|}{\gamma}\Delta t = 0.1\,a \tag{32}$$

yielding

$$A_{SC} = 0.05\gamma a \exp\left[-0.95\right] \approx 1.93 \times 10^{-3} \frac{\gamma a}{\Delta t} \,. \tag{33}$$

A soft-core potential has been preferred with respect to an hard-core potential because the latter would require a very small integration time step (i.e. of the same order as in all-atoms simulations, $10^{-12}s$) to avoid nonphysical large displacements. Moreover, using a very small time step would be at odds with the hypothesis of overdamped Brownian dynamics. And the dynamics would require a prohibitively large number of time steps. Periodic boundary conditions have been assumed for the positions of the particles but not for the long-range interactions that are computed without taking into account the images of the particles (due to the periodic boundary conditions). This latter choice is due to the fact that the long range electrodynamic interactions are supposed to be active only for particles in the effective volume of FCCS, simulated by the cubic box.

The numerical simulations have been performed using the Heun predictor-corrector algorithm with a time step $\Delta t = 5 \times 10^{-4} \mu$s and for a total number of steps $N_{\rm steps} = 2 \times 10^7$, corresponding to a total simulation time $T_{\rm tot} = 10^4 \mu$s. In order to assess the adequacy of this integration time, let us first notice that with the present choice of parameters the Brownian diffusion coefficient of the spherical molecules in simulations is $D = k_B T \gamma^{-1} = k_B T (6\pi a \eta_W)^{-1} \simeq 6.68 \times 10^3 \text{Å}^2 \mu\text{s}^{-1}$. Then considering a sphere circumscribing the box, thus of radius $R_{\rm circ} = \sqrt{3} l/2 \simeq 7.35 \times 10^3 \mathring{A}$, the characteristic time scale expected for a molecule to explore all the volume is $t_{\rm Vol} = R_{\rm circ}^2/(6D) \simeq 1.35 \times 10^3 \mu s$, largely contained in the total integration time of $10^4 \mu s$.

Molecular Dynamics simulations of equations (24) yield the results for the diffusion coefficient D, normalized by the Brownian value D_0, reported in Figure 5 of Main text, as a function of the average intermolecular distance $\langle d \rangle$.

The sudden drop of D/D_0 has been observed in presence of the attractive potential repre-

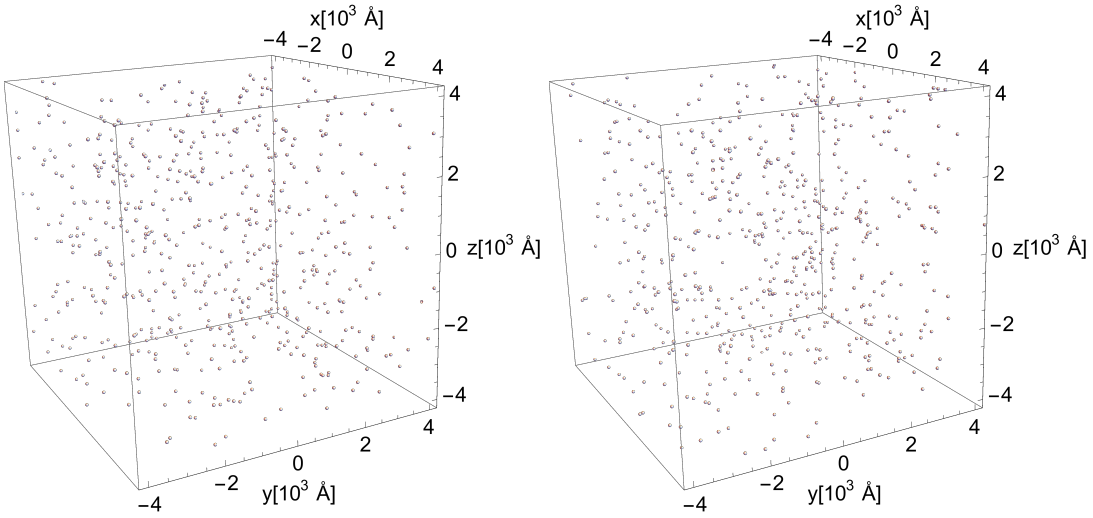

Figure S 13: Molecular dynamics simulations. Snapshots of spatial distributions of 500 molecules corresponding to an intermolecular average distance of $\langle d \rangle = 1000 \mathring{A}$. Left box: the initial condition. Right box: the final configuration after $10^4 \mu s$.

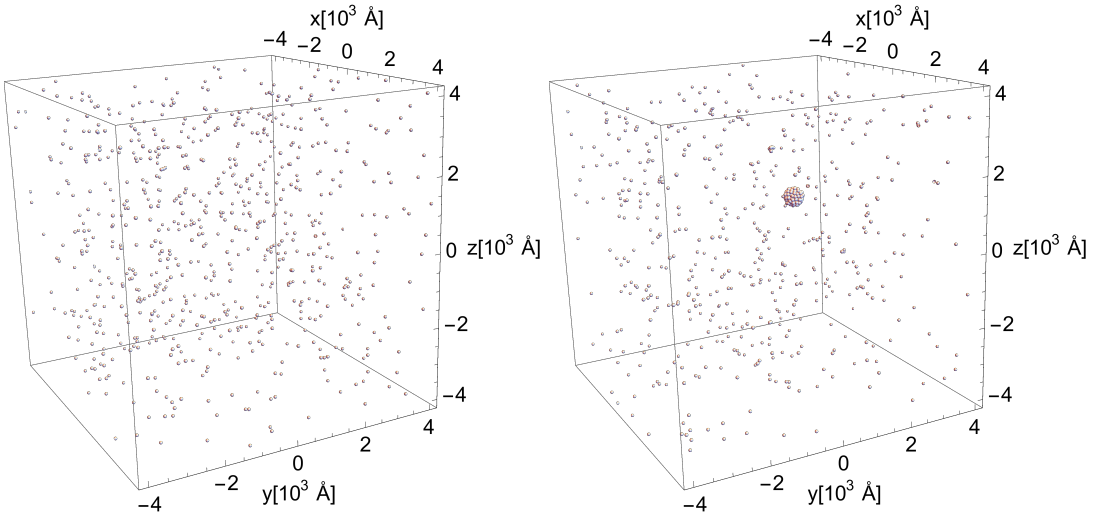

Figure S 14: Molecular dynamics simulations. Snapshots of spatial distributions of 700 molecules corresponding to an intermolecular average distance of $\langle d \rangle = 950 \mathring{A}$. Left box: the initial condition. Right box: the final configuration after $10^4 \mu$s.

senting the electrodynamic dipole-dipole potential. Such a drop of the relative diffusion coefficient is the observable manifestation of the clustering phase transition. The same pattern of D/D_0 versus $\langle d \rangle$ is displayed by the experimental outcomes of FCCS experiments reported in

Figure 4 of the Main text.

When the electrodynamic dipole-dipole interactions are not included in numerical simulations, the relative diffusion coefficient is always the Brownian one.

The parameters entering U_{eff} have been adjusted in order to observe the clustering transition at an intermolecular average distance of $\sim 1000\,\text{Å}$, this led to set $R_0 = 2a = 84\,\text{Å}$ and $c_{\text{eff}} = 2.25$ that correspond to a dynamical effective dipole of $|\mathbf{p}| \sim 1850\,\text{D}$. This dipole moment is larger than the one found in the semi-analytical model even if it is of the same order of magnitude. This can be explained by considering that the semi-analytical model represents a mean-field approximation of the particle model, thus enhancing the effect of long-range correlations. Moreover, in the semi-analytical model only the global minimum of free energy has been considered, while in molecular dynamics simulations also the relaxation from the initial disperse phase to the final equilibrium state has been followed. A-priori we cannot exclude that with $|\mathbf{p}| \sim 1850\,\text{D}$ the disperse state could be a metastable state with a very long lifetime, that is certainly not the case for the semi-analytical model which thus provides a lower bound for the estimated dynamic electric dipole modulus.

2.4 MonteCarlo simulations for the clustering transition

Equation 2 provides a first rough approximation of the electrodynamic potential derived in Ref.[12] of Main text for the interaction among two oscillating electric dipoles, as it does not take into account their mutual orientation nor their relative oscillation phase. At first, taking into account these degrees of freedom results in an interaction potential of indefinite sign, i.e. the long-range electrodynamic force between oscillating biomolecules can be attractive or repulsive as well. Therefore, it is important to check whether or not the clustering transition takes place also in presence of an orientation-dependent dipole-dipole interaction. If this is the case, we need to determine the value of the oscillating electric dipole moment that is necessary to explain

the experimental observations, that is, the value of $\langle d \rangle$ at which the clustering transition is found.

A more refined study of the clustering transition requires considering a more accurate model of the out-of-equilibrium electrodynamic interactions including information on mutual orientations and phase differences among the oscillating dipoles. For this reason a different electrodynamic dipole-dipole potential will be considered in the next section. Moreover, Molecular Dynamics simulations taking into account the relative orientations of the dipoles would require to take into account the random rotational dynamics of the dipoles. However, simulating roto-translational diffusion dynamics is a hard task to be implemented for biomolecules in our context. In fact, the rotational random dynamics is expected to introduce a faster process with respect to the spatial diffusion thus requiring integration time steps smaller than those already used. This would make the numerical simulations much heavier than those performed in the present work, simulations that already required many thousands of CPU hours. Thus an equivalent but more efficient computational approach consists in resorting to Monte Carlo simulations to study the clustering transition under the assumption of thermodynamic equilibrium of the roto-traslational degrees of freedom.

2.4.1 Potential Energy

A different model of intermolecular interactions with respect to the one used in previous sections is introduced in Monte Carlo simulations to achieve a more accurate description of the clustering transitions. The following form for the potential energy has been considered in Monte Carlo simulations:

$$
\begin{aligned}
U_{\mathrm{Tot}}(\mathbf{p}_i, \mathbf{p}_j, \mathbf{r}_{ij}) = & \qquad +\infty & r_{ij} < 2a \\
U_{\mathrm{Tot}}(\mathbf{p}_i, \mathbf{p}_j, \mathbf{r}_{ij}) = & \quad U_{\mathrm{Hmk}}(r_{ij}) + U_{\mathrm{Dby}}(r_{ij}) + U_{\mathrm{EDdip}}(\mathbf{p}_i, \mathbf{p}_j, \mathbf{r}_{ij}) & r_{ij} \geq 2a
\end{aligned}
\tag{34}
$$

where $\mathbf{r}_{ij}$ is the vector joining the centers of the i-th and the j-th molecule, $U_{\rm Dby}$ is the screened electrostatic potential of the force exerted between net charges of biomolecules, $U_{\rm EDdip}$ is the electrodynamic potential between resonant oscillating giant dipoles and $U_{\rm Hmk}$ is the pairwise Hamaker dispersion interaction energy. More in details:

- the electrostatic screened potential due to the presence of counterions has the form of the Debye-Hückel potential $U_{\rm Debye}(r)$ in Eq.(29) with the same choice of parameters;

- a more refined description of the clustering transition requires to take into account dispersion forces. At the biomolecular scale, dispersion forces can be effectively described by the Hamaker interactions (Ref.[52] of Main text)

$$U_{\rm Hmk}(r) = -\frac{A}{6}\left[\frac{2R_1R_2}{(2R_1+2R_2+r)r} + \frac{2R_1R_2}{(2R_1+r)(2R_2+r)} + \right. \\ \left. + \ln\frac{(2R_1+2R_2+r)r}{(2R_1+r)(2R_2+r)}\right] \tag{35}$$

  where $A = 3{-}10\ k_BT$ is the typical value of the *Hamaker constant* for proteins (Refs.[53] and [54] of Main text), and R_i is the radius of the i-th particle. In our simulations the Hamaker constant has been set to $A = 10\,k_BT$ with the radii $R_1 = R_2 = a = 42\mathring{A}$.

- The electrodynamic interaction is assumed to be pairwise and of the form

$$U_{\rm EDdip}(\mathbf{p}_i, \mathbf{p}_j, \mathbf{r}_{ij}) = f_{reg}(r_{ij})\frac{\mathbf{p}_i\cdot\mathbf{p}_j - 3(\mathbf{p}_i\cdot\hat{\mathbf{r}}_{ij})(\mathbf{p}_j\cdot\hat{\mathbf{r}}_{ij})}{4\pi\varepsilon_0\varepsilon_W(\omega_{CVM})r_{ij}^3} \tag{36}$$

  where $\mathbf{p}_i$ is the *dynamical* electric dipole moment of the i-th molecule, $\hat{\mathbf{r}}_{ij} = \mathbf{r}_{ij}/r_{ij}$ is the unit vector directed from the i-th particle to j-th one. The details of the derivation of this potential will be provided in Section 3.1.

  The interaction energy of a system of oscillating dipoles is generally speaking a function of time. Nevertheless, following the derivation in Ref.[12] of Main text, the interac-

tion energy has been averaged *over a time scale much larger than the typical period of dipole oscillation.* In so doing, the interaction energy depends only on the position of the dipoles, their orientations and their relative phase of oscillation $\Delta\alpha_{ij} = \alpha_j - \alpha_i$. The effect of this relative phase can be included in the relative orientation of the dipoles, so that the system of interacting oscillating resonant dipoles is mapped into a system of static dipoles. Adopting the above given form of the electrodynamic dipole-dipole potential requires to adapt accordingly its short distance regularization. In fact, as the potential depends on both the distance and the mutual orientation of the dipoles, the previously adopted algebraic regularization [in Eq.(2)] is not well suited because, at short distances, minor angular variations can determine a substantial nonphysical change of energy (very large torque). Moreover, a suited regularization should take into account the repulsion of molecular electron clouds due to the Pauli exclusion principle. For these reasons a regularization provided by the function $f_{reg}(r_{ij})$ has been introduced to soften the electrodynamic potential for $r \longrightarrow R_0 = 2a$ and to make it equal to the non regularized potential at infinity, i.e. $\lim_{r\to+\infty} f_{reg}(r) = 1$. To comply with these requirements, the regularizing function is chosen of the form

$$f_{reg}(r) = -\tanh\left(\frac{R_0 - r}{R_0\sigma_{reg}}\right) . \tag{37}$$

where the parameter $\sigma_{reg} = 0.679$ is determined after the assumption that $f_{reg}(2R_0) = 0.9$ (see Figure S15 for the effect of the regularization over a generic potential $U(r) = -r^{-3}$).

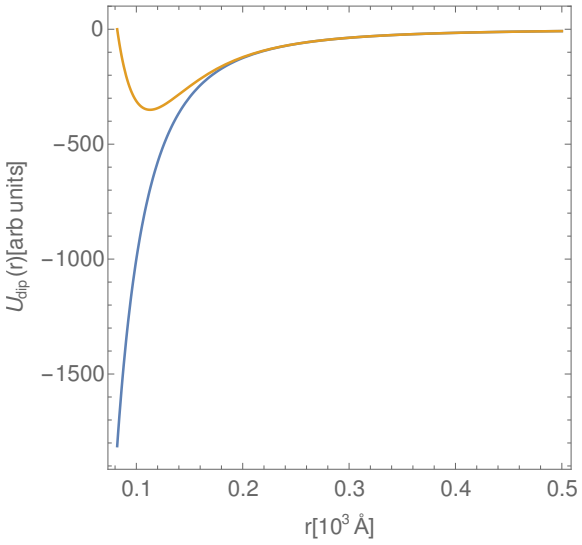

Figure S15: Effects of the damping function $f_{reg}(r)$ on a potential $U = -r^{-3}$. In blue the non regularized potential $-r^{-3}$, in orange the regularized potential $U_{reg}(r) = f_{reg}(r)U(r)$.

2.4.2 Details of the simulations

The simulations have been performed by considering a cubic box of side $L = 8486\mathring{A}$, with spherical particles of radius $a = 42\mathring{A}$. As for the molecular dynamics simulations described in Section 2.3, the dimensions of the fixed simulation box are chosen so that $V_{\text{eff}} = L^3$, and the number of particles in the box is changed in order to vary the intermolecular average distance $\langle d \rangle$, i.e.

$$N = \left\lfloor \frac{L^3}{V_{eff}} \right\rfloor \tag{38}$$

The degrees of freedom updated by the Monte Carlo Dynamics are the three Cartesian coordinates (x_i, y_i, z_i) and the polar angles (ϕ_i, θ_i) defining the position of the center of mass and the orientation of the dynamical electric dipole of each molecule, respectively. The domain for the coordinates of the center of mass is defined by $x_i, y_i, z_i \in [0, L]$ and $\phi \in [0, 2\pi)$, $\theta \in [0, \pi]$ for the angular coordinates. The m-th Monte Carlo step is performed according to the Metropolis algorithm prescriptions. In a system consisting of N particles, N random ex-

tractions of a particle are performed. At the k-th extraction of the m-th step, a test configuration $\{(\tilde{\mathbf{x}}_{i_k}(m), \tilde{\phi}_{i_k}(m), \tilde{\theta}_{i_k}(m)\}$ is created for the chosen particle labeled by i_k through a random displacement in configuration space with respect to a reference configuration, i.e.

$$\begin{cases} \tilde{\mathbf{x}}_{i_k}(m) = \mathbf{x}_{i_k}(m-1) + \Delta x \ \ \boldsymbol{\xi}_{i_k}(m) \\ \tilde{\phi}_{i_k}(m) = \phi_{i_k}(m-1) + \Delta\phi \ \ \xi_{\phi,i_k}(m) \\ \tilde{\theta}_{i_k}(m) = \theta_{i_k}(m-1) + \Delta\theta \ \ \xi_{\theta,i_k}(m) \end{cases} \tag{39}$$

where each $\xi_{\alpha,i}(m)$ is a Gaussian-distributed random variable with zero mean and unit variance:

$$\langle \xi_{\alpha,i}(m)\xi_{\beta,j}(m+m')\rangle = \delta_{m',0}\delta_{i,j}\delta_{\alpha,\beta} \tag{40}$$

where the average $\langle\cdot\rangle$ is intended over many realizations of the random process. The probability to accept a trial configuration is given by

$$T_{x\to\tilde{x}} = \min\left\{\frac{\exp[-\beta V(\tilde{x})]}{\exp[-\beta V(x)]}, 1\right\} . \tag{41}$$

If the trial configuration is accepted and a particle exits from the box, it is reinjected into the volume using the following prescription: the distance of the center of the particle from the box is calculated as:

$$d_i = \sqrt{\sum_{k=1}^{3}\left[x_{k,i}^2\Theta(-x_{k,i}) + (x_{k,i}-l)^2\Theta(x_{k,i}-l)\right]}, \tag{42}$$

then the particle is reinjected into the box at a distance d_i from a randomly chosen side of the box. More precisely, two random numbers η_1, η_2 are chosen in the interval $[0, L]$ and a third random number η_{sel} is extracted in order to determine the reinjection side, i.e. if $\mathbf{x}_i$ are the

coordinates of the reinjected particle:

$$\mathbf{x}_i = \begin{cases} (d_i, \eta_1, \eta_2) & \text{for} \quad 0 \geq \eta_{sel} > 1/6 \\ (L - d_i, \eta_1, \eta_2) & \text{for} \quad 1/6 \geq \eta_{sel} > 1/3 \\ (\eta_1, d_i, \eta_2) & \text{for} \quad 1/3 \geq \eta_{sel} > 1/2 \\ (\eta_1, L - d_i, \eta_2) & \text{for} \quad 1/2 \geq \eta_{sel} > 2/3 \\ (\eta_1, \eta_2, d_i) & \text{for} \quad 2/3 \geq \eta_{sel} > 5/6 \\ (\eta_1, \eta_2, L - d_i) & \text{for} \quad 5/6 \geq \eta_{sel} > 1\,. \end{cases}$$

This particular choice for the boundary conditions has been made in order to mimic the continuous flow of particles in and out of the confocal volume in FCCS experiments. As long-range interactions are assumed to be activated by the blue light only in the confocal volume and to be absent outside it, no correction of the long-range electrodynamic potential is needed (this would be in general the case with long range interactions requiring corrections to account for the contributions of the images of the system due to periodic boundary conditions).

2.4.3 Results of the simulations

The simulations have been performed choosing random positions and orientations for the particles in the box which represents the confocal volume. The number of Monte Carlo steps has been fixed to $N_{MCsteps} = 2 \times 10^6$ with $\delta_x = \Delta x/a = 2$ and $\delta_{ang} = \Delta\theta/\pi = \Delta\phi/(2\pi) = 0.1$.

This makes the square root of the mean square displacement of each particle $\sqrt{\mathrm{MSD}} \approx \sqrt{3}\Delta x\sqrt{2 \times 10^6} \approx 2.05{\times}10^5 \mathring{A} \approx 24L$, that is, large enough. In order to optimize the convergence rate of the dynamics to the stationary state of the system, an adaptive method has been introduced. Each 10^2 Monte Carlo steps the running acceptance ratio η_{AR} is calculated: if $\eta_{AR} < 0.33$ the relative displacements δ_x, δ_{ang} are halved (up to 5 times) while if $\eta_{AR} > 0.85$ the relative displacements are doubled one time. The results reported in Figures S16 and S17

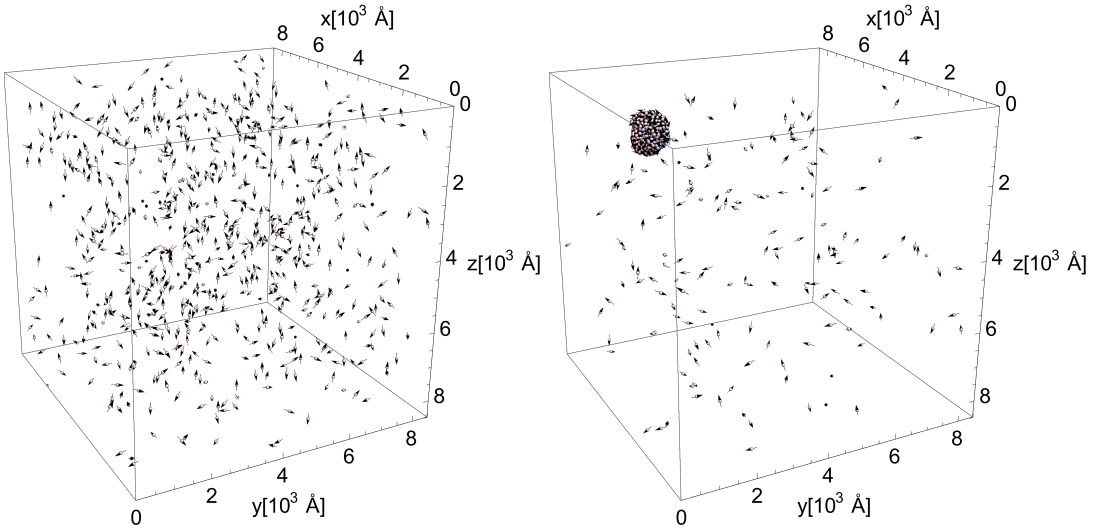

Figure S16: Clustering transition. Snapshots corresponding to the clustered phase in a cubic system of side $L = 8486\text{Å}$, with intermolecular average distance $\langle d \rangle = 950\text{Å}$, obtained by setting the number of particles to $N = 713$ and the dynamic electric dipole to $|\mathbf{p}| = 2900D$ at $T = 303.15K = 30^oC$. Left box: the initial condition. Right box: the final configuration after 2×10^6 Montecarlo steps.

correspond to a value $|\mathbf{p}| = 2900\text{D}$ of the modulus of the dynamic electric dipolar momentum. This value of $|\mathbf{p}|$ has been found to produce a clustering transition for $\langle d \rangle \lesssim 950\mathring{A}$ (see Figure S16) that disappears for $\langle d \rangle \gtrsim 1000\mathring{A}$ (see Figure S17).

The value $|\mathbf{p}| \sim 2900D$ is significantly larger than the value derived from the semi-analytical model and from molecular dynamics simulations for a clustering transition in the same conditions of temperature and concentration. This is due to the fact that in MC simulations the long-range electrodynamic interactions are both attractive and repulsive depending on the mutual dipole orientations (and mutual oscillation phases) while in the simulations performed with molecular dynamics the electrodynamic interactions were represented by an only attractive potential.

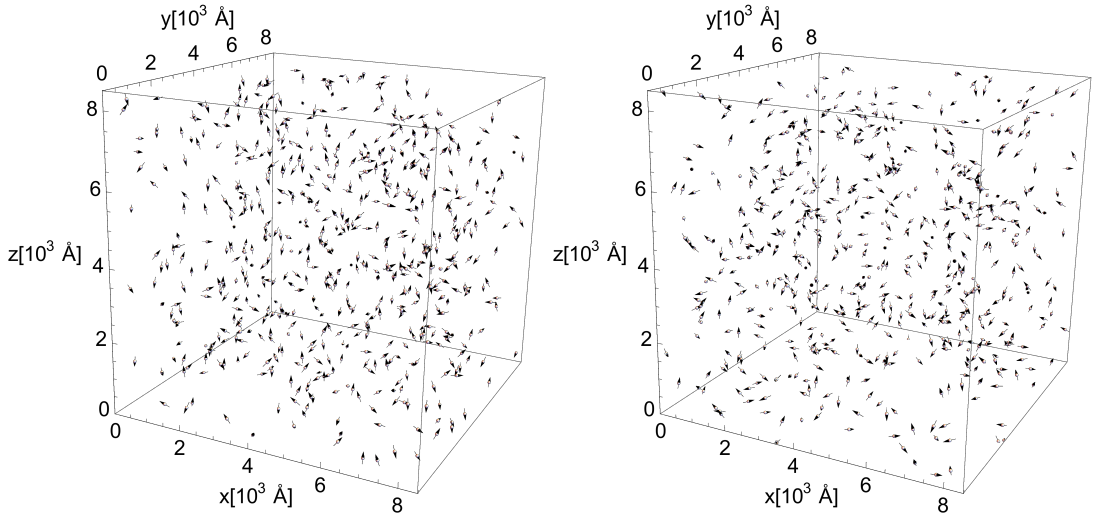

Figure S17: Absence of clustering. Snapshots of the disperse phase in a cubic system of side $L = 8486\text{Å}$, with intermolecular average distance $\langle d \rangle = 1000\text{Å}$, obtained by setting the number of particles to $N = 611$ and the dynamic electric dipole to $|\mathbf{p}| = 2900D$ at $T = 303.15K = 30^oC$. Left box: the initial condition. Right box: the final configuration after 2×10^6 Monte Carlo steps.

2.5 Further Remarks on the interpretation of FCCS experiments

The experimental values of the Brownian diffusion coefficients obtained at highly diluted solutions of R-PE are found to increase with the laser power.

On the basis of the standard equations for the Brownian diffusion coefficient

$$D = \frac{k_B T}{6\pi\eta(T) R_H} \ , \tag{43}$$

and for the temperature dependence of water viscosity in the interval between $0°$C and $370°$C

$$\eta(T) = A \times 10^{B/(T-C)} \ , \tag{44}$$

by using $\eta(20°C) = 10^{-3}$ Pa sec, with standard values A, B, and C equal to 2.414×10^{-5} Pa sec, 247.8 K, and 140 K, respectively, a straightforward computation, with $R_H = 42\mathring{A}$

for R-PE, yields the experimentally measured values: $D_0(50\mu W) = 125\mu m^2 sec^{-1}$ for $T = 58.5^oC$, $D_0(100\mu W) = 200\mu m^2 sec^{-1}$ for $T = 87.2^oC$, and $D_0(150\mu W) = 260\mu m^2 sec^{-1}$ for $T = 106.6^oC$. Such an increase of temperature within the confocal volume of the FCCS apparatus would be in-principle possible but ruled out by the measurement of the diffusion coefficients obtained with dye molecules (Atto488 or Alexa488) at the same laser powers, in fact no increase of the diffusion coefficient with increasing laser power is observed in this case. We might venture the guess that, instead of heating all the water in the confocal volume, a pseudo-heating effect could concern only the first hydration layers of each protein. In fact, considering an amplitude of some Angstroms of the collective oscillation of each molecule at the frequency of 96 GHz we obtain a velocity of expansion of roughly 100 meters s^{-1} of each protein molecule. This value is comparable with the velocity of water molecules at $T = 30^oC$ which would hit the proteins at a higher relative velocity and thus higher effective temperature.

A thorough understanding of this phenomenon is beyond the aim of the present work. The clustering phase transition with sudden drop of D/D_0 is a phenomenon which has nothing to do with the increase of D_0, whatever physical mechanism might cause this increment.

3 Effects of electrodynamic interactions on the frequency of molecular collective vibrations

3.1 Preliminaries

In THz spectroscopy experiments, the absorption peaks corresponding to collective intramolecular vibrations of both R-PE and BSA proteins have been observed to undergo a frequency shift proportional to molecular concentration.

In Ref.[19] of Main text it has been argued that these absorption peaks are the spectroscopic signature of a classical Fröhlich condensation-like phenomenon, i.e. when the energy injection

rate exceeds a threshold value the energy pumped into a molecule is almost entirely channelled into its lowest frequency mode, and this entails a collective vibration of the whole molecule. Recently, it has been suggested that a full and deeper understanding of the experimental results reported in Ref.[19] of Main text requires a quantum treatment of the Fröhlich condensation mechanism (see Ref.[22] of Main text). Even if this proposal is very interesting and robust, for the purposes of the present work we can proceed in the conceptual framework of a classical description.

In what follows, a biomolecule of mass M is represented by two spheres, each one of mass $M_{\rm crg} = M/2$ (where crg stands for center of charge) and effective charge $Z_{\rm eff}$, connected by a spring of elastic constant $k_{\rm CVM} = M\omega_{CVM}^2/4$, where $\omega_{CVM} = 2\pi\nu_{CVM}$ is the characteristic angular frequency of a collective vibrational mode (CVM).

The spring is actually a doughnut-like extended elastic object (the R-PE protein) mainly made of α-helices each one behaving like a spring. In its extension modes, because of geometric constraints, we can reasonably assume that the velocity of each mass element of the molecule is larger at the external parts and lower at the interior. A common assumption is a linear dependence of the velocity of a mass element passing from the slower part to the faster part. This means that the velocity u of a mass element dM is taken as $u(r) = vr/R$, where R is the larger radius of the molecule and r is the radial coordinate from the center. Assuming a uniform mass distribution we have $dM = M\ dr/R$. The kinetic energy of the molecule then reads

$$K = \int_{molecule} \frac{1}{2}u^2(r)dM(r) = \frac{Mv^2}{2R^3}\int_0^R r^2 dr = \frac{1}{2}\frac{M}{3}v^2$$

where M is the rest mass of the molecule and $M/3$ (known as the Rayleigh limit) is the effective mass contributing to the kinetic energy of the spring consisting of an extended elastic object.

Due to the Equipartition Theorem, at thermodynamic equilibrium the amplitude Δx of the

oscillations of the charge barycenters around the equilibrium position is given by

$$\begin{aligned} k_B T &= M_{eff}\ \omega_{CVM}^2 \langle (\Delta x)^2 \rangle \Rightarrow \\ \sigma_{\Delta x,Th} &= \sqrt{\langle (\Delta x)^2 \rangle} = \frac{1}{2\pi\nu_{CVM}} \sqrt{\frac{k_B T}{M_{eff}}} = \frac{7.99{\times}10^{-2}\mathring{A}}{[\nu_{CVM}]\sqrt{[M_{eff}]}} \end{aligned} \tag{45}$$

where $M_{eff} = M_{crg}/6 = M/12$, the experimental temperature is assumed $T = 30^o C = 303.15K$, the frequency $[\nu_{CVM}]$ is expressed in THz and the reduced mass $[M_{eff}]$ is expressed in KDa. For R-PE the observed peak of lowest frequency is at $[\nu_{CVM}] = 0.71 \times 10^{-1}$ and $[M_{eff}] \simeq 20$, so that the amplitude of the oscillation is $\sigma_{\Delta x} \approx 2.5{\times}10^{-1}\mathring{A}$. For BSA, the observed collective vibrational mode is $[\nu_{CVM}] = 3.14 \times 10^{-1}$ and the characteristic mass $[M_{eff}] \simeq 5.5$ so that the expected relative distance between the charge barycenters is $\sigma_{\Delta x} \approx 1.85{\times}10^{-2}\mathring{A}$.

The R-PE has a strong absorption at $\lambda_{AB,RPE} \simeq 488\text{n m}$ and the emission peak is around $\lambda_{EM,RPE} \simeq 580\text{n m}$: it follows that for a number n_{flc} of fluorochromes the quantity of energy absorbed is

$$\begin{aligned} &\Delta E_{RPE} < \eta\ n_{flc} hc(\lambda_{AB,RPE}^{-1} - \lambda_{EM,RPE}^{-1}) \simeq \\ &\simeq \eta\ n_{flc} \times 15.4 \times (k_B\, 303.15\text{K}) \simeq \eta\ n_{flc} \times 6.457{\times}10^{-13}\text{erg}\ , \end{aligned} \tag{46}$$

where η is the efficiency of the energy transfer from the fluorochrome to mechanical vibrational (phonon) modes of the biomolecule. The description of the details of such a process goes far beyond the purpose of this paper, but an efficiency around 10% can be expected (for further details on a possible mechanism describing the conversion of electronic excitation into mechanical excitation see Ref.[55] of Main text). It follows that the maximum extension amplitude for the

oscillating molecules is estimated to be:

$$\sigma_{\Delta x,Fc} = \sqrt{\langle(\Delta x)^2\rangle} < \frac{1}{2\pi\nu_{CVM}}\sqrt{\frac{\eta(n_{flc}\times 15.4)\times k_B T}{M_{eff}}} \simeq \sqrt{\eta\ n_{flc}}\times 5.6{\times}10^{-1}\mathring{A} \quad (47)$$

assuming for R-PA $n_{flc} = 38$ and $\eta = 0.1$ we obtain $\sigma_{\Delta x,Fc} \approx 1.08\mathring{A}$. The activation of a coherent excitation due to Fröhlich-like condensation is represented in the simple model here considered as a phase-defined oscillation maintained for a coherence time $\tau_{cohr} \gg \nu_{CVM}^{-1}$. So we can assume that the time evolution of the distance between the charge barycenters is given by

$$x_i(t) = x_{0,i} + \frac{1}{2}\left\{x_{\omega,i}\exp[-\mathrm{i}(\omega_{CVM}t + \phi_i)] + c.c\right\}\ , \quad (48)$$

where $x_{\omega_0,i} = \sqrt{2}\sigma_{\Delta x}$ and ω_{CVM} is the angular frequency of oscillation of the single molecule in the limit of high dilution. In real experiments with R-PE, the Q-factor of the absorption peak at the resonant frequency has been measured to be in the range $50 \div 90$, therefore it is not out of place to comment about such large values which are at odds with the much smaller values commonly reported in the literature for low frequency vibrational modes, but importantly, at thermal equilibrium and for non-interacting molecules. In the present work we deal with systems composed of molecules kept far from thermal equilibrium and interacting through electrodynamic forces. This last fact is crucial, in fact these mutual interactions affect the frequency of the intramolecular collective vibrations (as is shown in what follows), consequently we can hypothesize that these mutual interactions can also affect the oscillation phases of the molecules producing a synchronized oscillation of many molecules, thus making all of them - loosely speaking - more resistant against viscous damping due to the surrounding water molecules. This is a working hypothesis based on the results on synchronization phenomena reported in Ref.[26] of Main text. With the above given Q-values the coherence time can be estimated

through:

$$\tau_{cohr}\nu_{CVM} \approx \frac{Q}{\pi} \approx 16 \div 30 \Rightarrow \tau_{cohr} \approx \frac{Q}{\pi\nu_{CVM}} \approx 2.2 \div 4.0{\times}10^{-1} nsec \tag{49}$$

Considering the largest experimentally measured diffusion coefficient of R-PE $D_0 \approx 250\mu m^2 s^{-1}$ the displacement of the particle during the interval of coherent oscillations is of the order

$$\sqrt{\langle|\Delta\mathbf{r}_i|^2\rangle} \approx \sqrt{6D_0\tau_{cohr}} \approx 5.8 \div 8 \ \mathring{A} \tag{50}$$

which is much less than the intermolecular average distance and the characteristic dimension of the protein.

In this limit we can consider the positions and the orientations of the molecules as fixed for a time t_{int} such that $\tau_{cohr} \geq t_{int} \gg \nu_{CVM}^{-1}$. We can assume that the electric dipole of a single molecule in a reference frame attached to the molecule takes the form

$$\begin{aligned} \mathbf{p}_i(t) &\simeq \mathbf{p}_{stat,i} + \mathbf{p}_{dyn,i}(t) = \left\{ p_0 + \frac{p_\omega}{2} \exp\left[-\mathrm{i}(\omega t + \phi_i)\right] + c.c. \right\} \hat{\mathbf{p}}_i \\ &= Z_{eff} e \left\{ x_0 + \frac{x_\omega}{2} \exp\left[-\mathrm{i}(\omega t + \phi_i)\right] + c.c. \right\} \hat{\mathbf{p}} \qquad x_0, x_\omega \in \mathbb{R}^+ \end{aligned} \tag{51}$$

where $\hat{\mathbf{p}}$ is the normal vector that indicates the direction of the dipole, Z_{eff} can be defined as the equivalent charge of a symmetric dipole, i.e.

$$Z_{eff} = Z_+ - \frac{Z_{Tot}}{2} = Z_- + \frac{Z_{Tot}}{2} \tag{52}$$

and the relative position $\mathbf{x}$ of the charge barycenters is defined by

$$\mathbf{p}(t) = Z_+\mathbf{r}_+(t) + Z_-\mathbf{r}_-(t) = Z_{eff}\mathbf{x}(t)\,. \tag{53}$$

If the value of the effective separated charge is $Z_{eff}\, e$ then the electric dynamic dipole has the

following maximum

$$p_{\omega_{CVM}} = Z_{eff}\, e\, \sigma_{\Delta x} < Z_{eff} \times \sqrt{\eta n_{flc}} \times 2.69\,\mathrm{D}\ . \tag{54}$$

It is known that the charge barycenters of biomolecules are separated by a distance of $3 \div 10\ \mathring{A}$ (Ref.[56] of Main text), one order of magnitude larger than the estimated $x_{\omega_{CVM}}$. However, we can argue that in an electrolytic solution with a Debye length comparable to that of living cells, that is $\lambda_D \approx 10\mathring{A}$, the interaction due to the electrostatic charge distribution is negligible with respect to the electrodynamic interactions between the dynamical parts $\mathbf{p}_{dyn,i}(t)$ of the oscillating dipoles of two distinct molecules.

3.1.1 The electrodynamic interactions among the oscillating dipoles

According to Ref.[12] of Main text, the large dipole oscillations induced by energy injection and subsequent Fröhlich condensation are responsible for long-range interactions between biomolecules. Let us consider a system of N identical molecules, such that each molecule has a total mass M and a net charge Z_{net} and can be represented as a system of two lumps of mass $M_{eff} = M/2$ and connected by an harmonic spring of elastic constant $k = M_{eff}\,\omega_{CVM}^2$. The dynamical variables are the distances among the charge barycenters, $x_i(t) = x_{0,i} + x_{dyn,i}(t)$ so that the dipole moment can be written $\mathbf{p}_i(t) = Z_{eff} x(t) e \hat{\mathbf{p}}_i$. As the measured frequency shift due to the interactions is expected to be a perturbation we can assume that the vibrational mode with an angular frequency very close to ω_{CVM} has been activated, i.e.

$$x_i(\omega) \neq 0 \qquad \text{only if } |\omega \pm \omega_{\mathrm{CVM}}| < \epsilon\omega_{\mathrm{CVM}} \quad \text{for } \epsilon \ll 1 \tag{55}$$

If we consider the systems on a characteristic time scale τ_{int} such that $\tau_{cohr} \geq \tau_{int} \gg \nu_{CVM}^{-1}$, we can ignore the dissipation in the oscillation of the dipoles, and the total energy of the system

for the vibrational degrees of freedom can be written as

$$H_{dipOs}(\boldsymbol{P}_i(t), \mathbf{x}_i(t)) = \sum_{i=1}^{N} \Bigg[\frac{P_i^2(t)}{2M_{eff}} + \frac{1}{2} M_{eff}\ \omega_{CVM}^2 (x_i(t) - x_{0,i})^2 + \\ - \frac{1}{2} \sum_{\substack{j=1 \\ j \neq i}}^{N} Z_{eff,i} e x_i(t) \hat{\mathbf{p}}_i \cdot \mathbf{E}_j(\mathbf{r}_i, t) \Bigg] \tag{56}$$

where $x_{0,i}$ is the distance between charge barycenters, $P_i = M_{eff}\, \dot{x}_i(t)$ is the conjugate variable of the charge barycenter distance x_i and $\mathbf{E}_j(\mathbf{r}_i, t)$ is the electric field generated by the j-th particle at the point $\mathbf{r}_i$ at the time t. In order to derive the equations of motion for the system of coupled dipoles we need to express the electromagnetic field generated by the j-th molecule in terms of the charge barycenter separation x_j. According to Eq.(50) we can consider as fixed the positions and the orientations of the dipoles.

The electric field $\mathbf{E}_i(\mathbf{r}, t)$ splits into a static and a dynamic component, i.e.

$$\mathbf{E}_i(\mathbf{r}, t) = \mathbf{E}_{stat,i}(\mathbf{r}) + \mathbf{E}_{dyn,i}(\mathbf{r}, t) = \mathbf{E}_{stat,i}(\mathbf{r}) + \int_{\mathbb{R}\setminus\{0\}} \mathbf{E}_{\mathrm{dyn},i}(\mathbf{r}, \omega) \exp(-\mathrm{i}\,\omega t)\ \mathrm{d}\omega = \\ = \mathbf{E}_{stat,i}(\mathbf{r}) + \int_{\mathbb{R}^+\setminus\{0\}} 2|\mathbf{E}_{\mathrm{dyn},i}(\mathbf{r}, \omega)| \cos\left[\omega t - \theta_E(\omega)\right]\ \mathrm{d}\omega \tag{57}$$

where the polar representation of the Fourier coefficients of the electromagnetic field has been used; $\mathbf{E}_{stat,i}(\mathbf{r})$ is the static component of the electromagnetic field while $\mathbf{E}_{\mathrm{dyn},i}(\mathbf{r}, \omega)$ is the dynamical electric field generated by dipole oscillations. The electric fields generated by static charge distributions are subjected to Debye-Hückel screening due to the freely moving ions on a characteristic length scale λ_D

$$\lambda_D = \left(\frac{\sum_i \rho_{\infty,i} e^2 z_i^2}{\varepsilon_0 \varepsilon_W k_B T} \right)^{-1/2} \tag{58}$$

where $\rho_{\infty,i}$ is the concentration of the i-th electrolyte species and T is the temperature of the

solution. In the experiment mimicking biological conditions, the ionic strength of the NaCl solution is $200\mathrm{mM}$, which results in a Debye length $\lambda_D \approx 9.74\mathring{A}$. In this condition, the electrostatic field $\mathbf{E}_{stat,i}(\mathbf{r})$ generated by the i-th charge distribution located at $\mathbf{r}_i$ and characterized by a total charge Z_{Net} and electric dipole $\mathbf{p}_i = p_{0,i}\hat{\mathbf{p}}_i$ is given by (Ref.[57] of Main text)

$$\mathbf{E}_{stat,i}(\mathbf{r}) = \frac{\exp[-|\mathbf{r}-\mathbf{r}_i|/\lambda_D]}{4\pi\varepsilon_0\varepsilon_W(0)}\left\{\frac{Z_{Net,i}e}{|\mathbf{r}-\mathbf{r}_i|}\left(\frac{1}{|\mathbf{r}-\mathbf{r}_i|}+\frac{1}{\lambda_D}\right)\hat{\mathbf{n}}_{\mathbf{r}_i\mathbf{r}} + \right. \tag{59}$$
$$\left. + \;\; p_{0,i}\left[\left(\frac{3}{|\mathbf{r}-\mathbf{r}_i|^3}+\frac{3}{\lambda_D|\mathbf{r}-\mathbf{r}_i|^2}+\frac{1}{\lambda_D^2|\mathbf{r}-\mathbf{r}_i|}\right)(\hat{\mathbf{n}}_{\mathbf{r}_i\mathbf{r}}\cdot\hat{\mathbf{p}}_i)\hat{\mathbf{n}}_{\mathbf{r}_i\mathbf{r}} - \left(\frac{1}{|\mathbf{r}-\mathbf{r}_i|^3}-\frac{1}{\lambda_D|\mathbf{r}-\mathbf{r}_i|^2}\right)\hat{\mathbf{p}}_i\right]\right\}$$

where $\hat{\mathbf{n}}_{\mathbf{r}\mathbf{r}_i} = (\mathbf{r}-\mathbf{r}_i)/|\mathbf{r}-\mathbf{r}_i|$ is the unit vector along the direction joining the dipole and the point $\mathbf{r}$. The minimal intermolecular average distance among biomolecules considered in THz spectroscopy experiments was $\langle r\rangle \approx 600\mathring{A}$, so that in the electrostatic potential the leading term is

$$\mathbf{E}_{stat,i}(\mathbf{r}) \approx \frac{\exp[-|\mathbf{r}-\mathbf{r}_i|/\lambda_D]}{4\pi\varepsilon_0\varepsilon_W(0)\lambda_D|\mathbf{r}-\mathbf{r}_i|}\left[Z_{Net}e + \frac{p_{0,i}}{\lambda_D}(\hat{\mathbf{n}}_{\mathbf{r}_i\mathbf{r}}\cdot\hat{\mathbf{p}}_i)\right]\hat{\mathbf{n}}_{\mathbf{r}_i\mathbf{r}} = E_{stat,i}(\mathbf{r})\hat{\mathbf{n}}_{\mathbf{r}_i\mathbf{r}} \tag{60}$$

The dynamics of the electromagnetic field is described by D'Alembert equation in Lorenz gauge reading

$$\left[|\mathbf{k}|^2 - \frac{\omega^2}{v_c^2(\omega)}\right]\mathbf{A}_i(\mathbf{k},\omega) = \mu(\omega)\mathbf{J}_i(\mathbf{k},\omega) \tag{61}$$

where $v_c^2(\omega) = c^2/[\varepsilon_W(\omega)\mu_W(\omega)] = [\varepsilon_0\varepsilon_W(\omega)\mu_0\mu_W(\omega)]^{-1}$ represent the (complex) speed of propagation of light. In our case we can safely assume that the relative magnetic permittivity is 1, i.e. $\mu_W(\omega) = 1$ and the solution in real space is given by

$$\mathbf{A}_i(\mathbf{r},\omega) = \mu_0\int_{\mathbb{R}^3}\frac{\mathbf{J}_i(\mathbf{k},\omega)}{|\mathbf{k}|^2 - k_0^2(\omega)}\exp[+\mathrm{i}\mathbf{k}\cdot\mathbf{r}]\;\mathrm{d}^3k \tag{62}$$

where $k_0^2(\omega) = \omega^2/v_c^2(\omega) \in \mathbb{C}$. The Fourier components of the current associated to the

oscillation of the i-th dipole are

$$\mathbf{J}_i(\mathbf{k},\omega) = \frac{1}{(2\pi)^4}\int_{\mathbb{R}^3} \mathrm{d}^3k \int_{\mathbb{R}} \mathrm{d}\omega \;\; \mathbf{J}_i(\mathbf{r},t)\exp[-\mathrm{i}(\mathbf{k}\cdot\mathbf{r}-\omega t)] = \frac{p_i(\omega)\hat{\mathbf{p}}_i}{(2\pi)^3}(-\mathrm{i}\omega)\exp(-\mathrm{i}\mathbf{k}\cdot\mathbf{r}_i) \tag{63}$$

where

$$\begin{aligned}
&\mathbf{J}_i(\mathbf{r},t) = Z_{eff}e\left[\dot{r}_{i,+}\delta(\mathbf{r}-\mathbf{r}_{i,+}) - \dot{r}_{i,-}\delta(\mathbf{r}-\mathbf{r}_{i,-})\right]\hat{\mathbf{p}}_i = \frac{Z_{eff}e\;\hat{\mathbf{p}}_i}{(2\pi)^3}\int_{\mathbb{R}^3}\left[\dot{r}_{i,+}\exp\left(-\mathrm{i}\mathbf{k}\cdot\Delta\mathbf{r}_{i,+}\right) + \right. \\
&\left. -\dot{r}_{i,-}\exp\left(-\mathrm{i}\mathbf{k}\cdot\Delta\mathbf{r}_{i,-}\right)\right]\exp\left[-\mathrm{i}\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}_i)\right]\;\mathrm{d}^3k \approx \frac{Z_{eff}e\dot{x}_i\;\hat{\mathbf{p}}_i}{(2\pi)^3}\int_{\mathbb{R}^3}\exp\left[-\mathrm{i}(\mathbf{r}-\mathbf{r}_i)\right]\;\mathrm{d}^3k = \\
&= \frac{\dot{p}_i\;\hat{\mathbf{p}}_i}{(2\pi)^3}\int_{\mathbb{R}^3}\exp\left[-\mathrm{i}\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}_i)\right]\;\mathrm{d}^3k = \delta(\mathbf{r}-\mathbf{r}_i)\frac{\mathrm{d}\mathbf{p}_{\mathrm{dyn},i}}{\mathrm{d}t}
\end{aligned} \tag{64}$$

where we have introduced the distance $\Delta\mathbf{r}_{\pm,i} = \mathbf{r}_\pm - \mathbf{r}_i$ of the positive and negative charge barycenters from the center of net charges, and the dipole approximation has been considered, i.e. $\mathbf{k}\cdot\Delta\mathbf{r}_{\pm,i} \ll 1$. Moreover, it has been assumed that the orientation of the dipoles is fixed ($\dot{\hat{\mathbf{p}}}_i = 0$). Under these assumptions it follows that the only source of the electromagnetic field are the oscillating dipoles.

Substituting Eq.(63) in Eq.(62) we obtain

$$\begin{aligned}
\mathbf{A}_i(\mathbf{r},\omega) &= \frac{\mu_0(-\mathrm{i}\omega)p_i(\omega)\hat{\mathbf{p}}_i}{(2\pi)^3}\int_{\mathbb{R}^3}\frac{\exp[\mathrm{i}\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}_i)]}{|\mathbf{k}|^2-k_0^2}\mathrm{d}^3k = \\
&= \frac{\mu_0(-\mathrm{i}\omega)p_i(\omega)\hat{\mathbf{p}}_i}{4\pi|\mathbf{r}-\mathbf{r}_i|}\exp[\varsigma\,\mathrm{i}\,k_0(\omega)|\mathbf{r}-\mathbf{r}_i|] = \\
&= \frac{\mu_0(-\mathrm{i}\omega)p_i(\omega)\hat{\mathbf{p}}_i}{4\pi|\mathbf{r}-\mathbf{r}_i|}\exp[\varsigma\,\mathrm{i}\,\mathrm{Re}(\mathrm{k}_0)|\mathbf{r}-\mathbf{r}_\mathrm{i}|]\exp[-|\mathrm{Im}(\mathrm{k}_0)||\mathbf{r}-\mathbf{r}_\mathrm{i}|]
\end{aligned} \tag{65}$$

where $k_0 = \sqrt{k_0(\omega)^2} \in \mathbb{C}$ and $\varsigma = \mathrm{sgn}[\mathrm{Im}(\mathrm{k}_0)]$. According to the conventions adopted in Appendix A for the dielectric constant of water $\varsigma = \mathrm{sgn}(\omega)$. Using Maxwell equations, the

expressions of the magnetic field is

$$\mathbf{H}_i(\mathbf{r},\omega) = \frac{\mathrm{rot}\mathbf{A}_i(\mathbf{r},\omega)}{\mu_0\mu_W(\omega)} = (\hat{\mathbf{n}}_{\mathbf{rr}_i}\times\hat{\mathbf{p}}_i)\frac{\varsigma\omega k_0(\omega)p_i(\omega)}{4\pi|\mathbf{r}-\mathbf{r}_i|}\left(1-\frac{1}{\mathrm{i}\varsigma\mathrm{k}_0(\omega)|\mathbf{r}-\mathbf{r}_\mathrm{i}|}\right)\exp[\varsigma\,\mathrm{i}\,k_0(\omega)|\mathbf{r}-\mathbf{r}_i|] \tag{66}$$

and of the electric field is

$$\mathbf{E}_i(\mathbf{r},\omega) = \frac{1}{(-i\omega)\varepsilon_0\varepsilon_W(\omega)}\left[\mathrm{rot}\mathbf{H}_i(\mathbf{r},\omega)-\mu_W(\omega)\mathbf{J}_i(\mathbf{r},\omega)\right]\ . \tag{67}$$

The Fourier transform in frequency domain of the dynamical part of the electric field generated by the i-th dipole for $\mathbf{r}\neq\mathbf{r}_i$ reads

$$\begin{aligned}\mathbf{E}_{dyn,i}(\mathbf{r},\omega) &= \frac{p_i(\omega)}{4\pi\varepsilon_0\varepsilon_W(\omega)}\left\{\frac{k_0^2(\omega)}{|\mathbf{r}-\mathbf{r}_i|}(\hat{\mathbf{n}}_{\mathbf{rr}_i}\times\hat{\mathbf{p}}_i)\times\hat{\mathbf{n}}_{\mathbf{rr}_i}+\right.\\ &+ \left.\left(\frac{1}{|\mathbf{r}-\mathbf{r}_i|^3}-\frac{\mathrm{i}\varsigma k_0(\omega)}{|\mathbf{r}-\mathbf{r}_i|^2}\right)[3(\hat{\mathbf{n}}_{\mathbf{rr}_i}\cdot\hat{\mathbf{p}}_i)\hat{\mathbf{n}}_{\mathbf{rr}_i}-\hat{\mathbf{p}}_i]\right\}\exp[\mathrm{i}\varsigma k_0(\omega)|\mathbf{r}-\mathbf{r}_i|]\end{aligned} \tag{68}$$

The attenuation range (by a factor e) of the field is given by $\lambda_{att}(\omega)=\mathrm{Im}^{-1}[\mathrm{k}_0(\omega)]$ while the field wavelength is given by $\lambda(\omega)=2\pi\mathrm{Re}^{-1}[\mathrm{k}_0(\omega)]$. In the NaCl water solution condition used in experiments, and in the range of (angular) frequencies ω attributed to the collective oscillations (CVM) of the biomolecules under consideration (R-PA and BSA) it is $\lambda(\omega)\gtrsim 10\lambda_{att}(\omega)$. It follows that the retardation effects can be neglected, leading to

$$\mathbf{E}_{dyn,i}(\mathbf{r},\omega) = \frac{p_i(\omega)}{4\pi\varepsilon_0\varepsilon_W(\omega)|\mathbf{r}-\mathbf{r}_i|^3}\left[3(\hat{\mathbf{n}}_{\mathbf{rr}_i}\cdot\hat{\mathbf{p}}_i)\hat{\mathbf{n}}_{\mathbf{rr}_i}-\hat{\mathbf{p}}_i\right]\ . \tag{69}$$

So, the real electric field can be rewritten as

$$\begin{aligned}\mathbf{E}_{dyn,i}(\mathbf{r},t) &= \int_{\mathbb{R}^+/\{0\}} \frac{p_i(\omega)}{4\pi\varepsilon_0\varepsilon_W(\omega)|\mathbf{r}-\mathbf{r}_i|^3}\left[3(\hat{\mathbf{n}}_{\mathbf{r}\mathbf{r}_i}\cdot\hat{\mathbf{p}}_i)\hat{\mathbf{n}}_{\mathbf{r}\mathbf{r}_i}-\hat{\mathbf{p}}_i\right]\exp(-\mathrm{i}\omega t)\mathrm{d}\omega = \\ &= \frac{\left[3(\hat{\mathbf{n}}_{\mathbf{r}\mathbf{r}_i}\cdot\hat{\mathbf{p}}_i)\hat{\mathbf{n}}_{\mathbf{r}\mathbf{r}_i}-\hat{\mathbf{p}}_i\right]}{4\pi\varepsilon_0|\mathbf{r}-\mathbf{r}_i|^3}\int_{\mathbb{R}^+/\{0\}}\frac{p_i(\omega)\exp(-\mathrm{i}\phi_{\varepsilon_\mathrm{W}}(\omega))\exp(-\mathrm{i}\omega\mathrm{t})}{|\varepsilon_W(\omega)|}\mathrm{d}\omega \quad (70)\end{aligned}$$

where we have used $\varepsilon_W(\omega) = |\varepsilon_W(\omega)|\exp[+\mathrm{i}\phi_{\varepsilon_\mathrm{W}}(\omega)]$. Using the reality condition on the time dependent dielectric constant $\varepsilon_W(t)$, i.e. $\phi_{\varepsilon_W}(-\omega) = -\phi_{\varepsilon_W}(\omega)$, we obtain

$$\mathbf{E}_{dyn,i}(\mathbf{r},t) = \frac{\left[3(\hat{\mathbf{n}}_{\mathbf{r}\mathbf{r}_i}\cdot\hat{\mathbf{p}}_i)-\hat{\mathbf{p}}_i\right]}{4\pi\varepsilon_0|\mathbf{r}-\mathbf{r}_i|^3}\mathcal{P}_i(t)\ , \quad (71)$$

where the effective generating dipole $\mathcal{P}_i(t)$ is defined as

$$\mathcal{P}_i(t) = \int_{\mathbb{R}^+/\{0\}}\frac{2|p_i(\omega)|\cos\left[\omega t-\phi_i(\omega)-\phi_{\varepsilon_W}(\omega)\right]}{|\varepsilon_W(\omega)|^2}\mathrm{d}\omega = \int_{-\infty}^{t}\chi(t-t')p_i(t')\,\mathrm{d}t'\ , \quad (72)$$

and the response function $\chi(t-t')$ depends on the dielectric properties of the water solution. The interaction energy between the i-th dipole and the electric field generated by the j-th dipole $\mathcal{P}_j$ is given by

$$\begin{aligned}V_{dyn,ij} &= \frac{\hat{\mathbf{p}}_i\cdot\hat{\mathbf{p}}_j-3(\hat{\mathbf{p}}_i\cdot\hat{\mathbf{n}}_{\mathbf{r}_i\mathbf{r}_j})(\hat{\mathbf{p}}_i\cdot\hat{\mathbf{n}}_{\mathbf{r}_i\mathbf{r}_j})}{4\pi\varepsilon_0|\mathbf{r}_i-\mathbf{r}_j|^3}\mathcal{P}_j(t)p_i(t) = \\ &= \frac{\hat{\mathbf{p}}_i\cdot\hat{\mathbf{p}}_j-3(\hat{\mathbf{p}}_i\cdot\hat{\mathbf{n}}_{\mathbf{r}_i\mathbf{r}_j})(\hat{\mathbf{p}}_i\cdot\hat{\mathbf{n}}_{\mathbf{r}_i\mathbf{r}_j})}{4\pi\varepsilon_0|\mathbf{r}_i-\mathbf{r}_j|^3}\int_{-\infty}^{t}\mathrm{d}t'\ \chi(t-t')p_j(t')p_i(t)\,. \quad (73)\end{aligned}$$

The dispersion effects do not allow to provide a straightforward Hamiltonian formulation of the dynamics of the oscillating dipoles; for this reason we introduce an effective dipole $\tilde{\mathcal{P}}_i(t) = \mathcal{C}p_i(t)$ with no phase mismatch effects due to the dispersion properties of the aqueous solution, i.e.

$$\sqrt{\langle\mathcal{P}^2\rangle_{[0,+\infty)}} = \sqrt{\langle\tilde{\mathcal{P}}^2\rangle_{[0,+\infty)}} \quad (74)$$

where $\langle\cdot\rangle_{[0,T]}$ stands for the time average on the interval $[0,T]$. The condition in Eq.(74) can be rewritten as

$$\sqrt{\int_0^{+\infty} \frac{2|p_i(\omega)|^2}{|\varepsilon_W(\omega)|^2}\,\mathrm{d}\omega} = \mathcal{C}\sqrt{\int_0^{+\infty} 2|p_i(\omega)|^2\,\mathrm{d}\omega} \tag{75}$$

and assuming that the modulus of the complex dielectric constant is almost constant on the support of $|p_i(\omega)|$, centered around ω_{CVM}, we obtain

$$\mathcal{C} = \frac{1}{|\varepsilon_W(\omega_{CVM})|}\ . \tag{76}$$

We verify that, in the case of study, we can effectively expect that the variation of the modulus of the dielectric constant is negligible. According to what is reported in Appendix A about the dielectric properties of salty water, we can expect that in conditions analogous to the experimental ones the maximum variation of the modulus of the (complex) dielectric constant is given by

$$\Delta|\varepsilon(\omega)| \approx \left|\frac{\mathrm{d}|\varepsilon(\omega_{CVM})|}{\mathrm{d}\omega}(\Delta\omega)_{FS}\right| \tag{77}$$

where $(\Delta\omega)_{FS}$ is the maximum of the experimentally measured frequency shift of the absorption peak in the THz spectrum of a protein. For R-PE $\mathrm{d}|\varepsilon(\omega_{CVM})|/\mathrm{d}\omega = -43.610^{-12}sec$ and $(\Delta\omega)_{FS} \approx 2.6\times10^{-3}$THz, from which it follows that $(\Delta|\varepsilon(\omega)|)_{RPE} \approx 1.2\times10^{-1}$ and $(\Delta|\varepsilon(\omega)|)_{RPE}/|\varepsilon(\omega_{CVM})| \approx 5.1\times10^{-3}$. This means that at the level of accuracy expected by the current theoretical interpretation of the experimental results, the approximation of a constant absolute value of the dielectric constant is a good one. In the case of BSA we have that $\mathrm{d}|\varepsilon(\omega_{CVM})|/\mathrm{d}\omega = -1.9110^{-12}sec$ and $(\Delta\omega)_{FS} \approx 0.81\times10^{-1}$THz; according to Eq.(77) the estimated variation of the modulus of the relative dielectric constant is $(\Delta|\varepsilon(\omega)|)_{RPE} \approx 1.6\times10^{-1}$, whence $(\Delta|\varepsilon(\omega)|)_{RPE}/|\varepsilon(\omega_{CVM})| \approx 2.1\times10^{-2}$.

From the above considerations, it follows that the potential between oscillating dipoles can

be rewritten as

$$\begin{aligned} V_{dyn,ij}(t) &= \frac{\hat{\mathbf{p}}_i \cdot \hat{\mathbf{p}}_j - 3(\hat{\mathbf{p}}_i \cdot \hat{\mathbf{n}}_{\mathbf{r}_i\mathbf{r}_j})(\hat{\mathbf{p}}_i \cdot \hat{\mathbf{n}}_{\mathbf{r}_i\mathbf{r}_j})}{4\pi\varepsilon_0|\varepsilon_W(\omega_{CVM})||\mathbf{r}_i - \mathbf{r}_j|^3} p_j(t)p_i(t) = \\ &= Z_i Z_j e^2 \frac{\hat{\mathbf{p}}_i \cdot \hat{\mathbf{p}}_j - 3(\hat{\mathbf{p}}_i \cdot \hat{\mathbf{n}}_{\mathbf{r}_i\mathbf{r}_j})(\hat{\mathbf{p}}_i \cdot \hat{\mathbf{n}}_{\mathbf{r}_i\mathbf{r}_j})}{4\pi\varepsilon_0|\varepsilon_W(\omega_{CVM})||\mathbf{r}_i - \mathbf{r}_j|^3} x_i(t)x_j(t)\,. \end{aligned} \tag{78}$$

This is the electrodynamic potential between two biomolecules when collective giant dipole oscillation are activated by an external source of energy.

We are now able to describe the dynamics of the hamiltonian system of dipole oscillators. Let us introduce the variables x_{dyn} expressing the dynamical part of the separation between the barycenters of charge , i.e.

$$x_i(t) = x_{eq,i} + x_{\mathrm{dyn},i}(t) = x_{eq,i} + \int_{\mathbb{R}/\{0\}} 2x_i(\omega)\cos[\omega t + \phi_i(\omega)]\ \mathrm{d}\omega \tag{79}$$

where $x_{eq,i}$ is the static equilibrium elongation of the dipole associated with the i-th molecule, i.e.

$$\partial_{x_{\mathrm{dyn},i}} H\big|_{\boldsymbol{x}_{dyn}=0} = 0 \qquad \forall i = 1, ..., N\ . \tag{80}$$

From Eq.(79) it follows that the canonical conjugate momenta associated to the variables $x_{\mathrm{dyn},i}(t)$ remain unchanged with respect to the momenta P_i associated to the variables x_i.

According to Eqs.(60) and (71), we can rewrite the electric field generated by the i-th molecule using the variables introduced in Eq.(79)

$$\mathbf{E}_i(\mathbf{r}, t) = \boldsymbol{\chi}^{(s)}(\mathbf{r}; \mathbf{r}_i, \hat{\mathbf{p}}_i)p_{0,i} + \boldsymbol{\chi}^{(d)}(\mathbf{r}; \mathbf{r}_i, \hat{\mathbf{p}}_i)p_{dyn,i}(t) \tag{81}$$

where

$$\boldsymbol{\chi}^{(s)}(\mathbf{r};\mathbf{r}_i,\hat{\mathbf{p}}_i) = \frac{\exp[-|\mathbf{r}-\mathbf{r}_i|/\lambda_D]}{4\pi\varepsilon_0\varepsilon_W(0)\lambda_D|\mathbf{r}-\mathbf{r}_i|}\frac{\hat{\mathbf{n}}_{\mathbf{r}_i\mathbf{r}}}{\lambda_D}(\hat{\mathbf{n}}_{\mathbf{r}_i\mathbf{r}}\cdot\hat{\mathbf{p}}_i) \tag{82}$$

$$\boldsymbol{\chi}^{(d)}(\mathbf{r};\mathbf{r}_i,\hat{\mathbf{p}}_i) = \frac{[3(\hat{\mathbf{n}}_{\mathbf{r}\mathbf{r}_i}\cdot\hat{\mathbf{p}}_i)-\hat{\mathbf{p}}_i]}{4\pi\varepsilon_0\varepsilon_W(\omega_{CVM})|\mathbf{r}-\mathbf{r}_i|^3}\ . \tag{83}$$

With the notations introduced in Eqs.(79) and (82) the equations of motions become

$$\begin{aligned}\dot{x}_i &= \frac{P_i}{M_{eff}}\\ \dot{P}_i &= -M_{eff}\omega_{CVM}^2 x_{dyn,i} + \frac{Z_{eff,i}Z_{eff,j}e^2}{2}\times \\ &\times \sum_{j\neq i}^{j=1\ldots N}\left[p_{eq,j}\left(\hat{\mathbf{p}}_i\cdot\boldsymbol{\chi}^{(s)}(\mathbf{r}_i;\mathbf{r}_j,\hat{\mathbf{p}}_j)+\hat{\mathbf{p}}_j\cdot\boldsymbol{\chi}^{(d)}(\mathbf{r}_j;\mathbf{r}_i,\hat{\mathbf{p}}_i)\right)+2p_{dyn,j}\hat{\mathbf{p}}_i\cdot\boldsymbol{\chi}^{(d)}(\mathbf{r}_i;\mathbf{r}_j,\hat{\mathbf{p}}_j)\right]\end{aligned} \tag{84}$$

where we have used $\hat{\mathbf{p}}_i\cdot\boldsymbol{\chi}^{(d)}(\mathbf{r}_i;\mathbf{r}_j,\hat{\mathbf{p}}_j) = \hat{\mathbf{p}}_j\cdot\boldsymbol{\chi}^{(d)}(\mathbf{r}_j;\mathbf{r}_i,\hat{\mathbf{p}}_i)$. As we are interested in the long range behaviour when $|\mathbf{r}_i-\mathbf{r}_j| \gg \lambda_D$ we can neglect terms containing χ^s in Eq.(84). Introducing the geometric coupling parameter $\zeta_{ij} = -\hat{\mathbf{p}}_i\cdot\boldsymbol{\chi}^{(d)}(\mathbf{r}_i;\mathbf{r}_j,\hat{\mathbf{p}}_j)$ we can rewrite the equations of motion for $x_{dyn,i}$ in terms of its Fourier components yielding

$$\begin{aligned}&\int_{\mathbb{R}/\{0\}}\left\{[-\omega^2+\omega_{CVM}^2]x_{\omega,i}\cos(\omega t+\phi_{i,\omega})+\sum_{j\neq i}^{j=1\ldots N}\frac{Z_{eff,i}Z_{eff,j}\zeta_{ij}}{M_{eff}}x_{\omega,j}\cos(\omega t+\phi_{j,\omega})\right\}\ \mathrm{d}\omega +\\ &+\ \omega_{CVM}^2(x_{eq,i}-x_{0,i})+\sum_{j\neq i}^{j=1\ldots N}\frac{Z_{eff,i}Z_{eff,j}e^2\zeta_{ij}}{M_{eff}}x_{eq,j}=0\end{aligned} \tag{85}$$

The condition for $x_{eq,i}$ can be rewritten as

$$M_{ij}x_{eq,j} = x_{0,i} \qquad \text{with} \quad M_{ij} = \delta_{ij}+\frac{Z_{eff,i}Z_{eff,j}e^2\zeta_{ij}}{M_{eff}\omega_{CVM}^2}(1-\delta_{ij}) \tag{86}$$

while the condition for the solutions of the dynamics can be rewritten as a set of equations for

the normal modes of the system, i.e.

$$(M_{ij} - \tilde{\omega}^2 \delta_{ij}) x_{\omega,j} = 0 \tag{87}$$

where $\tilde{\omega} = \omega/\omega_{CVM}$. The set of eigenvalues of the matrix M defines the square of the angular frequencies of the normal modes $\{\tilde{\omega}_A\}_{A=1,...,N}$. The eigenfrequencies are well defined as $\tilde{\omega}_A^2 > 0$ for all $A = 1, ..., N$ in the perturbative regime, i.e.

$$\left| \frac{Z_{eff,i} Z_{eff,j} e^2 \zeta_{ij}}{M_{eff} \omega_{CVM}^2} \right| \ll 1 \ . \tag{88}$$

Within this theoretical framework it is possible to provide an explanation for the observed frequency shift as discussed in Subsection 3.1.2. The Hamiltonian of the system can be rewritten as

$$H_{\text{dipOs}}(\{\mathbf{R}_{ij}\}_{ij}) = H_{\text{eq}}(\{\mathbf{R}_{ij}\}_{ij}) + \sum_{A=1}^{N} \tilde{J}_A \tilde{\omega}_A(\{\mathbf{R}_{ij}\}_{ij}) \omega_{\text{CVM}} \approx \sum_{A=1}^{N} \tilde{J}_A \tilde{\omega}_A(\{\mathbf{R}_{ij}\}_{ij}) \omega_{\text{CVM}} \ . \tag{89}$$

where $H_{\text{eq}}(\{\mathbf{R}_{ij}\}_{ij})$ is the energy contribution of the static dipoles that we assume to be negligible with respect to the other term representing the contribution of the dynamics of the coupled oscillators. In order to define the interaction energy among dipoles as a function of their relative positions $\mathbf{R}_{ij}$, the energy of the oscillators in the limit of infinite mutual distances is considered. In fact, a parameter $\epsilon = \min_{ij} |\mathbf{R}_{ij}|^{-1}$ representing the inverse of a length scale can be introduced such that $\mathbf{R}_{ij} = \widetilde{\mathbf{R}}_{ij}/\epsilon$ where $\widetilde{\mathbf{R}}_{ij} \geq 1$. It follows that the Hamiltonian describing the energy of the coupled oscillators depends on ϵ and that the case of decoupled oscillators can be obtained in the limit

$$H_{\text{freeOs}} = \lim_{\epsilon \to 0^+} H_{\text{dipOs}}(\widetilde{\mathbf{R}}_{ij}/\epsilon) = \sum_{i=1} J_i \omega_{CVM} \ .$$

In this framework, the interaction energy is defined as

$$\Delta U_{\rm int} = \int_0^1 \frac{{\rm d}H_{\rm dipOs}}{{\rm d}\epsilon}{\rm d}\epsilon\,. \tag{90}$$

Assuming an adiabatic process ideally connecting the asymptotic state of non-interacting dipoles $\epsilon = 0$ and the state of interacting dipoles $\epsilon = 1$, the action J is an adiabatic invariant, i.e. $J_i \approx J_A|_{\epsilon=0}$. The interaction energy takes the form

$$\Delta U_{\rm int}(\{\mathbf{R}_{ij}\}_{ij}) = \sum_{A=1}^{N} J_A \int_1^{\tilde{\omega}_A} \omega_{\rm CVM}\ {\rm d}\eta_A = \sum_{A=1}^{N} J_A \omega_{\rm CVM}[\tilde{\omega}_A(\{\mathbf{R}_{ij}\}_{ij}) - 1]\,. \tag{91}$$

This form of the interaction energy is the generalization to the case of N oscillating dipoles of the results derived in Ref.[12] of Main text for a pair of oscillating resonant dipoles. In fact, the frequency shift in a system of two coupled identical oscillating dipoles is given by

$$\tilde{\omega}_{\pm} = 1 \pm \frac{Z_{\rm eff}^2 e^2}{8\pi\varepsilon_0 \varepsilon(\omega_{\rm CVM})|\mathbf{R}_{12}|^3 M_{\rm eff}\omega_{\rm CVM}} \tag{92}$$

and, consequently, the interaction energy in that case scales as $|\mathbf{R}_{12}|^{-3}$. In the general case of N interacting dipoles, the frequency shift of each normal mode with respect to the reference frequency $\omega_{\rm CVM}$ is a non trivial function depending on the position of all the particles in the system. For such a reason we performed numerical simulations in order to provide a theoretical explanation of the experimental outcomes.

3.1.2 Theoretical interpretation of the observed frequency shift of the absorption peak in THz spectroscopy experiments

In THz absorption spectroscopy experiments, the external monochromatic reading THz field $\mathbf{E}_{read}(t) = E_{read}(t)\hat{\mathbf{E}}_{read}$ couples with the system of mutually interacting excited oscillating

dipoles, i.e.

$$U_{read} = -\mathbf{E}_{read}(t) \cdot \mathbf{p}_{\rm Tot} = E_{read}(t) \sum_{i=1}^{N} Z_i e x_i(t) \left(-\hat{\mathbf{p}}_i \cdot \hat{\mathbf{E}}_{test}\right) = \tag{93}$$

$$E_{read}(t) \sum_{i=1}^{N} \sum_{A=1}^{N} \left(-\hat{\mathbf{p}}_i \cdot \hat{\mathbf{E}}_{test}\right) Z_i e \mathcal{O}_i^A \tilde{x}_{0,A} \cos(\theta_A + \phi_A) = E_{read}(t) \sum_{A=1}^{N} \mathcal{C}_{cpl}(\tilde{\omega}_A) \cos\left[\tilde{\theta}_A(t) + \phi_A\right]$$

where we have used $x_i = \sum_{A=1}^{N} \mathcal{O}_i^A \tilde{x}_A = \sum_{A=1}^{N} \mathcal{O}_i^A \tilde{x}_{0,A} \cos\left[\tilde{\theta}_A(t) + \phi_A\right]$.

We assume that the major contribution to THz absorption in experiments is due to the normal mode maximally coupled with the external probing field, i.e.:

$$\mathcal{C}_{Mcpl}(\{\mathbf{r}_i, \hat{\mathbf{p}}_i\}_i) = \max_A \ |\mathcal{C}_{cpl}(\tilde{\omega}_A)| = |\mathcal{C}_{cpl}(\tilde{\omega}_{Mcpl}(\{\mathbf{r}_i, \hat{\mathbf{p}}_i\}_i))| \tag{94}$$

where $\tilde{\omega}_{Mcpl}$ is the angular frequency of the collective mode with the largest absolute value of dipole-field coupling constant.

In Eq.(94) we stressed that the mode with the maximal coupling constant depends on the (fixed) positions and orientations of the molecules. In order to establish the value of the amplitude of each normal mode $\tilde{x}_{0,A}$ we assume that $\tilde{x}_{0,A} \simeq \tilde{x}_{0,B} \simeq \tilde{x}_0$. The value of the coefficients $\tilde{x}_{0,A}$ has been chosen assuming that $\langle (\Delta x_i)^2 \rangle = \frac{x_{CVM}^2}{2}$, where x_{CVM}^2 is the amplitude of the oscillation of the barycenters of electric charge and $\langle \cdot \rangle$ is the time average over a time much larger than the ν_{CVM}^{-1}. Under these hypotheses, we deduce that

$$\begin{aligned} \langle (\Delta x_i)^2 \rangle &= \sum_{A,B} \mathcal{O}_i^A \mathcal{O}_i^B \tilde{x}_{0,A} \tilde{x}_{0,B} \langle \cos(\tilde{\omega}_A t + \phi_A) \cos(\tilde{\omega}_B t + \phi_B) \rangle = \sum_{A,B} \mathcal{O}_i^A \mathcal{O}_i^B \tilde{x}_{0,A} \tilde{x}_{0,B} \frac{\delta_{A,B}}{2} = \\ &= \frac{\tilde{x}_0^2}{2} \sum_A (\mathcal{O}_i^A)^2 = \frac{\tilde{x}_0^2}{2} \end{aligned} \tag{95}$$

where we have used the properties of the orthogonal matrix $\sum_{A=1}^{N} (\mathcal{O}_i^A)^2 = 1$. From Eq.(95) it

follows that the total dipole associated to each normal mode is given by

$$\mathcal{C}_{cpl}(\tilde{\omega}_A) = \sum_i (-\hat{\mathbf{p}}_i \cdot \hat{\mathbf{E}}_{read}) Z_i e \mathcal{O}_i^A \sqrt{2}\sigma_i = \sum_i (-\mathbf{p}_{A,i} \cdot \hat{\mathbf{E}}_{read}) \tag{96}$$

where we have introduced the electric dipole moment

$$\mathbf{p}_{A,i} = Z_i e \mathcal{O}_i^A \sqrt{2}\sigma_i \hat{\mathbf{p}}_i \tag{97}$$

as the effective electric dipole amplitude at the frequency $\tilde{\omega}_A$ of the i-th molecule. The (absolute value) of the frequency shift can be defined as

$$\|\Delta\omega\| = \tilde{\omega}_{Mcpl} - \omega_{CVM} \ . \tag{98}$$

We can analytically estimate the frequency shift for a system of two oscillating dipoles representing two excited biomolecules with fixed positions and orientations

$$\Delta\omega = \frac{Z_i Z_j e^2}{2M_{eff}\omega_0} \frac{\left[\hat{\mathbf{p}}_i \cdot \hat{\mathbf{p}}_j - 3\left(\hat{\mathbf{p}}_i \cdot \hat{\mathbf{r}}_{ij}\right)\left(\hat{\mathbf{p}}_i \cdot \hat{\mathbf{r}}_{ij}\right)\right]}{4\pi|\varepsilon_W(\omega_{CVM})|\varepsilon_0 r_{ij}^3} \ . \tag{99}$$

In THz spectroscopy experiments on R-PE, a relative frequency shift $\|\Delta\omega(600\mathring{A})\|/\omega_{RPE,CVM} \approx 6\times10^{-3}$ has been measured. Assuming $|\varepsilon_W(\omega_{RPE})| \approx 22.7$, $M_{eff} \simeq M_{RPE}/12 \approx 0.2\times10^2$kD, the geometric factor $|\hat{\mathbf{p}}_i \cdot \hat{\mathbf{p}}_j - 3(\hat{\mathbf{p}}_i \cdot \hat{\mathbf{r}}_{ij})(\hat{\mathbf{p}}_j \cdot \hat{\mathbf{r}}_{ij})| \approx 1$, then the effective electric charge is estimated to be

$$Z \approx \sqrt{\frac{8M_{eff}\omega_0^2\zeta\pi\varepsilon_W(\omega_0)\varepsilon_0\left(600\mathring{A}\right)^3}{e^2}} = 1.3\times10^3 \ . \tag{100}$$

The spatial power density of energy injection in THz spectroscopy experiments is much smaller than the spatial power density of energy injection in FCCS experiments. In fact, if we assume that the strength of the dynamic dipole (and consequently of the interaction) is proportional to

the amount of the energy injection rate into the protein, we can expect that the oscillating dipole strength (and the strength of the interaction) is smaller in THz spectroscopy experiments than in FCCS experiments. This is confirmed by the experimental observation that in THz spectroscopy experiments no cluster was detected for intermolecular average distances $\langle d \rangle \geq 600\mathring{A}$ while in FCCS experiments the clustering transition was observed for $\langle d \rangle \lesssim 900\mathring{A}$. This means that in estimating the mean squared elongation of the dipole σ_i^2 we have to consider a smaller value than the one estimated in (47). Assuming that $\sigma_{\Delta x} \sim 0.5 \times \sigma_{\Delta x,Fc} \approx 0.5\mathring{A}$ as in the case of R-PE, we derive a value of $p_\omega \approx 5.3 \times 10^3 \mathrm{D}$ to explain the frequency shift observed in THz experiment.

This value of the estimated oscillating dipole strength seems to contradict the previous assumption of a smaller dipole strength than the one we have evaluated for FCCS experiments $|\mathbf{p}| \sim 2900\mathrm{D}$. However, in the previous estimation of the frequency shift we have considered only two oscillating dipoles: as the electrodynamic interactions among resonating oscillating dipoles are supposed to be long-range (scaling as r^{-3}), we are underestimating the effect of long range dipole-dipole electrodynamic interactions in a many-body system. So we expect that such a value of the dynamical electric dipole is an overestimation of the real value in experiments.

In real spectroscopic experiments the situation is different with respect to the ideal two body case because the number of R-PE molecules contained in the observation volume ($\approx 1\mu L$) is of the order $10^{15} - 10^{18}$. So, numerical simulations have been performed in order to investigate how the frequency shift of the absorption peak depends on the intermolecular average distance for a system of interacting dipoles, with fixed positions and orientations, linearly coupled through a quasi-static dipole-dipole potential.

We assume that the testing field of the experiments is linearly polarized along the $\hat{z}$ direction, so that the normal mode of the system that maximizes the coupling with the reading field is the

one having the largest polarization along the z-axis. It follows that the relative frequency shift $\Delta\omega = \Delta\omega(\{\hat{\mathbf{p}}_i, \mathbf{r}_i\}_i)$ associated to a given configuration for the position and orientations of the system of dipoles is given by

$$\Delta\omega(\{\hat{\mathbf{p}}_i, \mathbf{r}_i\}_i) = 1 - \sqrt{\omega^2_{max}} \qquad \tilde{\omega}^2_{max} = \{\tilde{\omega}^2_A \,|\, |\mathcal{C}_{cpl}(\tilde{\omega}_A)| = \max_{B=1,\dots,N}\{|\mathcal{C}_{cpl}(\tilde{\omega}_B|\}\}\,. \quad (101)$$

In order to compare the result of numerical simulations with the outcomes of THz spectroscopy experiments:

- the average over many configurations of the positions and the orientations of the dipoles has to be considered in order to take into account the thermal and statistical fluctuations of the orientations and positions;
- the average frequency shift for a fixed intermolecular average distance has to be calculated for systems of different size and *extrapolated for large* N, as finite size effects affect systems with long range interactions.

In the following section the protocol to calculate the average over the configurations and to extrapolate the frequency shift for large N is discussed.

3.2 Algorithm for numerical simulations

The aim of the simulations is to estimate the frequency shift for an ensemble of N spherical molecules of radius a in a cubic box of size L as a function of the intermolecular average distance $\langle d\rangle$. The system of oscillating dipoles consists of three different sets of degrees of freedom: the coordinates of the center of mass of each particle, the orientation of the main dipole and the coordinate which describes the vibration of the dipole. For each fixed value of the intermolecular average distance $\langle d\rangle$, $n_{CMconf} = 5 \times 10^3$ configurations of the ensemble of molecules have been randomly chosen. The position of the center of each molecule has been

randomly set with a uniform probability distribution inside the box, avoiding the overlaps. The orientation of each dipole is described by a couple of polar angular coordinates (θ, ϕ) such that $\theta_i \in [0, \pi)$ and $\phi_i = [0, 2\pi)$ for $i = 1, ..., N$. For each fixed configuration of the centers of mass of the molecules, the orientational degrees of freedom of the molecules have been thermalized using a Monte Carlo-Metropolis scheme. All the trial configurations for the orientations have been generated by adding to the angular coordinates of each particle a randomly chosen number with a uniform distribution. The intervals are $[-\eta\pi, \eta\pi]$ for the θ_i angles and $[-2\eta\pi, 2\eta\pi]$ for the ϕ_i angles. In the performed simulations the parameter that describes the width of the interval has been set to $\eta = 7.5 \times 10^{-3}$. For each configuration of the centers of mass, $n_{therm} = 2 \times 10^3$ Monte Carlo steps have been performed to thermalize the orientations of the dipoles of the system. Then, starting from the final configuration so obtained, an average value of the frequency shift has been computed over $n_{MCstep} = 50$ configurations furtherly generated and interspersed with 10 Monte Carlo steps. The relative frequency shift associated to each intermolecular average distance has been computed as the overall average on a total number of $n_{totsamp} = n_{MCstep} \times n_{CMconf} = 2.5 \times 10^5$ different configurations of the dipole orientations and of the positions of the centers of mass. The potential energy dependence among different configurations has been calculated using the time average of Eq.(78) assuming $x_i(t) = x_{\rm CVM} \cos(\omega_{\rm CVM} t)$.

3.3 Numerical results

The numerical simulations were performed for systems of different sizes and different values of the dipole moment p_i and of the effective charge Z_i. The entries of the matrix M_{ij} depend on the value of the effective charge and only indirectly on the value of the modulus of the dipole moment of each molecule. The value of the dynamic electric dipole of each molecule enters in the calculation of the total interaction energy among the dipoles required by Monte Carlo-

Metropolis algorithm. The tested values of the effective charge were empirically chosen in order to provide a relative frequency shift comparable with the one observed in the experiments. The value of the amplitude of the dynamic dipole oscillations has been chosen heuristically in order to reproduce the experimental data.

The average frequency shift for a fixed intermolecular average distance has been measured for different sizes of the system. For each fixed number of molecules the relative frequency shift $\Delta\nu_0/\nu_0$ has been plotted as a function of the intermolecular average distance $\langle d\rangle$. The data have been fitted with a power law of the form $\Delta\nu_0/\nu_0 = Ax^{-k}$ and an inverse cubic law $\Delta\nu_0/\nu_0 = Bx^{-3}$.

In conclusion, the highly remarkable result of our computations is that the frequency shift of the collective intramolecular oscillations of an ensemble of proteins interacting through a dipole-dipole electrodynamic force, scales as $1/\langle d\rangle^3$. By inversion, the experimental observation of this result proves that the molecules under investigation interact through the electrodynamic dipole-dipole forces above discussed.

N	$A[10^6\ \mathring{A}^3]$	k	$B[10^6\ \mathring{A}^3]$
50	2.04 ± 0.04	2.65 ± 0.05	1.79 ± 0.04
100	1.438 ± 0.014	2.90 ± 0.02	1.38 ± 0.10
200	1.17 ± 0.02	$k = 3.03 \pm 0.04$	1.185 ± 0.005

Table S2: Table of the fitted parameters for the relative frequency shift. The fitted parameters are reported as a function of the intermolecular average distance $\langle d\rangle$ and different system sizes (the number of molecules in the box is N). The parameters $\{A_N, k_N\}_N$ correspond to the fit $\Delta\nu_0/\nu_0 = A_N x^{-k_N}$ and $\{B_N\}_N$ are the parameters characterizing the inverse cubic distribution $\Delta\nu_0/\nu_0 = B_N x^{-3}$. The modulus of each dipole was chosen to be $p_i = 2100$ D while the effective charge of the dipole is $Z_i = 850$. With these choices the amplitude of dipole oscillation is $x_{\omega_0} \simeq 0.51\mathring{A}$.

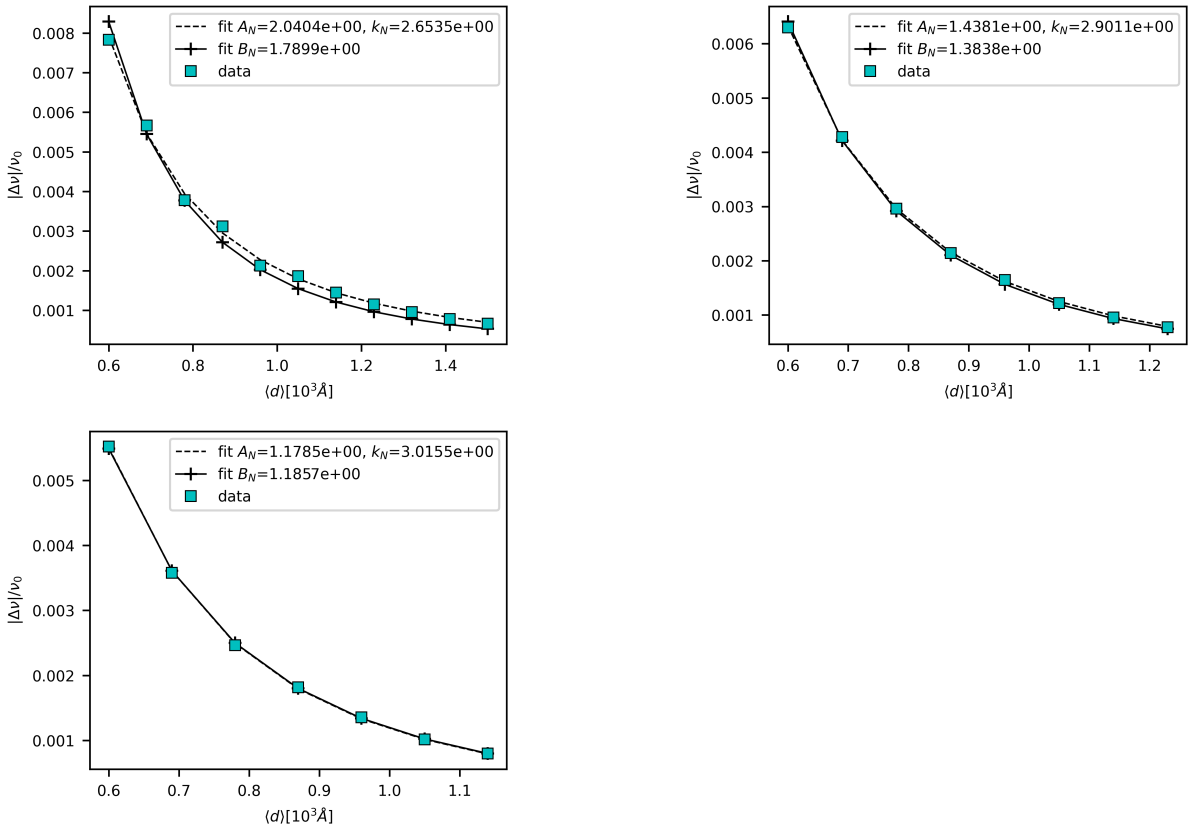

Figure S18: Relative frequency shift as a function of the intermolecular average distance. N = 50 (top), N = 100 (center) N = 200 (bottom). The coefficients A_N and B_N are expressed in $10^6\mathring{A}^3$. The modulus of each dipole was chosen to be $p_i = 2100\ D$ while the effective charge of the dipole is $Z_i = 850$. With these choices the amplitude of dipole oscillation is $x_{\omega_0} \simeq 0.51\mathring{A}$.

N	$A[10^6\ \mathring{A}^3]$	k	$B[10^6\ \mathring{A}^3]$
50	1.55 ± 0.02	2.67 ± 0.04	1.37 ± 0.03
100	1.10 ± 0.010	2.95 ± 0.02	1.081 ± 0.005
200	0.907 ± 0.009	3.05 ± 0.02	9.27 ± 0.04

Table S3: Table of the fitted parameters for the relative frequency shift. The fitted parameters are reported as a function of the intermolecular average distance $\langle d\rangle$ and different system sizes (the number of molecules in the box is N). The parameters $\{A_N, k_N\}_N$ correspond to the fit $\Delta\nu_0/\nu_0 = A_N x^{-k_N}$ and $\{B_N\}_N$ are the parameters characterizing the inverse cubic distribution $\Delta\nu_0/\nu_0 = B_N x^{-3}$. The modulus of each dipole was chosen to be $p_i = 1850$ D while the effective charge of the dipole is $Z_i = 750$. With these choices the amplitude of dipole oscillation is $x_{\omega_0} \simeq 0.51\mathring{A}$.

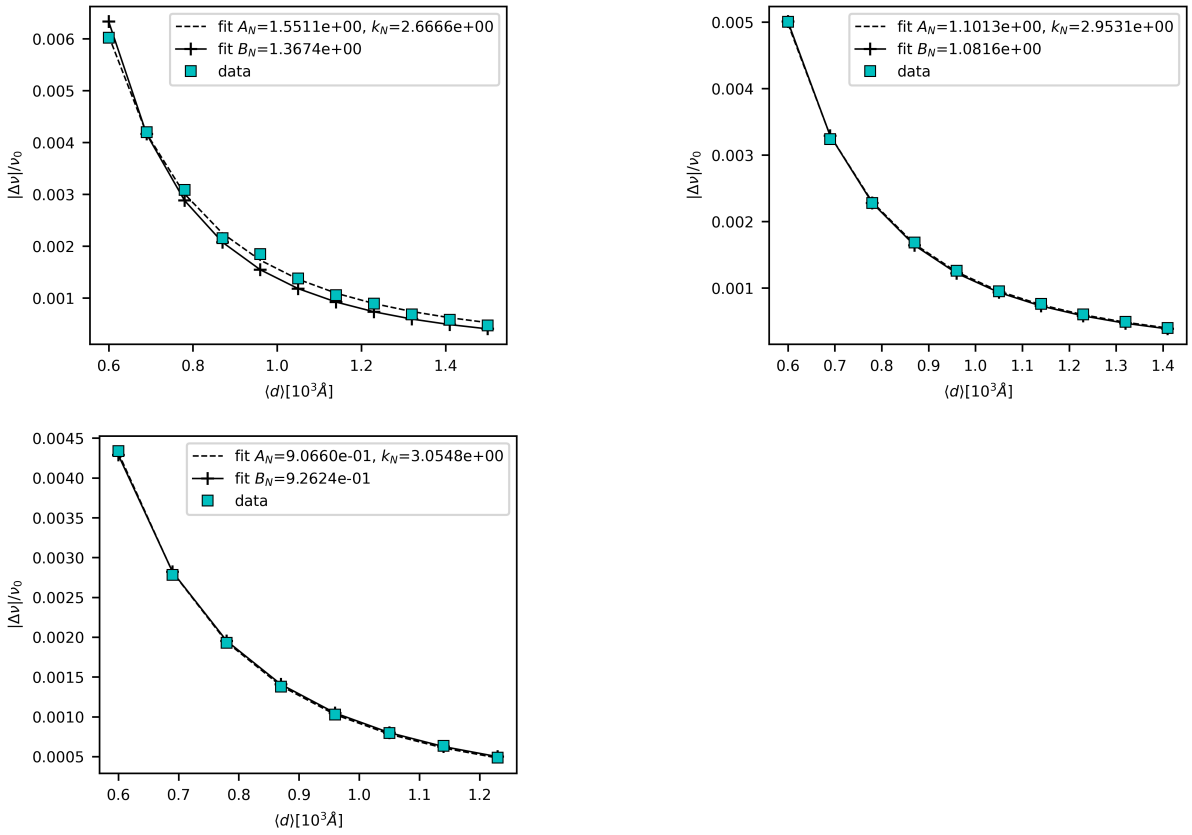

Figure S19: Relative frequency shift as a function of the intermolecular average distance. N = 50 (top), N = 100 (center) N = 200 (bottom). The coefficients A_N and B_N are expressed in $10^6\mathring{A}^3$. The modulus of each dipole was chosen to be $p_i = 1850\ D$ while the effective charge of the dipole is $Z_i = 750$. With these choices the amplitude of dipole oscillation is $x_{\omega_0} \simeq 0.51\mathring{A}$.

4 Remark on possible activation mechanisms of electrodynamic forces in vivo

The light-induced activation of electrodynamic forces has been adopted to warrant a well controllable and reproducible *in vitro* methodology. Even though the R-PE protein is a natural light harvesting protein, we might wonder which kind of activation mechanisms might be hypothesised to take place in living cells. A comment on this point is not out of place because there is no physical reason to consider the activation of electrodynamic forces limited to the creation of "hot points" through the excitation of fluorophores. In fact, the theoretical modelling in Ref.[19] of Main text requires an external source of energy injection into a macromolecule of absolutely generic form.

Actually, in vivo the external energy supply for the activation of electrodynamic attractive forces between cognate partners could be well provided by the cellular machinery itself as energy released by adenosine triphosphate (ATP) or guanosine triphosphate (GTP) hydrolysis by specific enzymes. In fact, the typical intracellular concentration of ATP molecules is given around $1\ \mathrm{mM}$ implying that a protein molecule in the cell undergoes around 10^6 collisions with ATP molecules per second (Ref.[58] of Main text). Given the standard free-energy obtained from ATP hydrolysis estimated around $50\ \mathrm{kJmol}^{-1} = 8.306 \cdot 10^{-13}\mathrm{erg}$, we can assume that 1% of the collisions with ATP will provide energy, which corresponds to a power supply of $8.306 \cdot 10^{-9}\mathrm{erg\ s}^{-1}$ potentially available. Besides ATP hydrolysis, other possible forms of energy supply should be considered in a cellular environment, for example, the energy released from mitochondria in the course of citric acid cycle with a power supply given around $10^{-7}\mathrm{erg\ s}^{-1}$. This source of energy might well be enough to excite long-range electrodynamic forces as the corresponding power is around or larger than the power considered for ATP hydrolysis. Let us also mention the recent experimental evidence for thermally induced "protein

quakes" to initiate enzymatic catalysis (Ref.[59] of Main text) through a solvent-dependent non isotropic momentum transfer due to the collisions of water molecules or ions.

It is worth mentioning that, for a broad class of physical systems, long-living Quasi Stationary States (QSS) can be dynamically generated which keep a system out of thermodynamic equilibrium. Among many other systems where QSS are produced (Ref.[60] of Main text) let us mention a beam of fast particles interacting with the set of waves describing a physical system a situation which is reminiscent, for example, of fast phosphate groups - produced by ATP hydrolysis - colliding against suitable sites of a biomolecule to create "hot points" yielding "protein quakes". In Ref.[19] of Main text, "protein quakes" have been invoked to explain the activation of collective oscillations through light irradiation of the BSA protein.

A Dielectric properties of solution of salt in water

From data reported in Ref.[61] of Main text we have interpolated the complex dielectric constant of water solution of NaCl at 200m M at $T = 30^o C$:

$$\varepsilon_W(\omega) = \varepsilon_{W,\infty} + \frac{\varepsilon_W(0) - \varepsilon_{W,\infty}}{1 - i\omega\tau_W} \tag{102}$$

with $\varepsilon_{W,\infty} = 5$, $\varepsilon_W(0) = 72.7$ and $\tau = 7.0110^{-12} sec.$ It follows that for R-PE ($\omega_{CVM} = 2\pi \times 71\mathrm{GHz} = 0.446\mathrm{THz}$) the value of the electric dielectric constant is

$$\varepsilon_W(\omega_{CVM}) = 11.3 + \mathrm{i}\ 19.6 \tag{103}$$

so that the characteristic wavelength corresponding to the angular frequency of the CVM observed for the R-PE is

$$k(\omega_{CVM,\mathrm{RPE}}) = (6.13 + \mathrm{i}\ 3.55) \times 10^{-7} \mathrm{\AA}^{-1} \tag{104}$$

assuming $\mu_W(\omega_{CVM}) = 1$.

	$\varepsilon_W(\omega_{CVM})$	$\lvert\varepsilon_W(\omega_{CVM})\rvert$	$\phi_{\varepsilon_W}(\omega_{CVM})$	$\varepsilon'_W(\omega_{CVM})$	$\lvert\varepsilon_W(\omega_{CVM})\rvert'$	$\phi'_\varepsilon(\omega_{CVM})$
R-PE	$11.3 + \mathrm{i}\,19.6$	22.6	1.050	$-(25.5 + \mathrm{i}35.8)$	-43.8	0.189
BSA	$5.35 + \mathrm{i}\,4.87$	7.24	0.738	$-(0.35 + \mathrm{i}\,2.44)$	-1.91	-0.217

Table S4: Dielectric properties of water in the regime of frequency of the Collective Vibration Mode (CVM) for R-PE and BSA. The Fourier transform of the relative dielectric constant of water is expressed using polar form $\varepsilon_W(\omega) = |\varepsilon_W(\omega)| \exp[\mathrm{i}\phi_{\varepsilon_W}(\omega)]$. The derivatives (primed quantities) of the different adimensional quantities with respect to ω are expressed in $10^{-12} sec = \mathrm{THz}^{-1}$.

An analogous calculation for BSA ($\omega_{CVM} \simeq 2\pi \times 0.314\mathrm{THz} = 1.97\mathrm{THz}$) gives $\varepsilon_W(\omega_{CVM,\mathrm{BSA}}) = 5.35132 - \mathrm{i}\ 4.8655$, and the wavenumber is

$$k(\omega_{CVM,BSA}) = (1.65 - \mathrm{i}\ 0.638) \times 10^{-6} \mathring{\mathrm{A}}^{-1} \tag{105}$$

For the biomolecules considered, BSA and R-PE, the characteristic attenuation length scale $\kappa_{att} = [\Im(k)]^{-1}$ is smaller than both the wavelength $\lambda_{rad} = [\Re(k)]^{-1}$ of the radiation and the characteristic length scale of the observed systems in THz spectroscopy experiments, i.e. $l_{sys} \approx (1\mu L)^{1/3} = 10^7 \mathring{A} = 10^{-1}\mathrm{cm}$.

	$\omega_{CVM}[THz]$	$\lambda_{rad}(\omega_{CVM})[\times 10^7 \mathring{A}]$	$\kappa_{att}(\omega_{CVM})[\times 10^7 \mathring{A}]$	$\kappa^3_{att}[\mu L]$
R-PE	0.447	1.02	0.281	0.022
BSA	1.97	0.606	0.156	0.0038

Table S5: Characteristic parameters of the BSA and R-PE proteins.

This means that the dissipative properties of the medium define the range of the interaction: inside a volume of κ^3_{att} the dynamical electric field can be considered quasi-static, i.e. retardation and radiation terms in Eq.(68) can be neglected.

B Effective mass for two oscillating charge centers

Here we consider the inner dynamics of a unidimensional oscillator along a fixed axis.

Charge barycenters coordinates $x_\pm$ can be expressed in terms of the relative/center-of-mass coordinates $r = x_+ - x_-$ and $R = (m_+x_+ - m_-x_-)/m_{Tot}$, where $m_{Tot} = m_+ + m_-$ is the total mass of the oscillating dipole. The following assumptions have been made:

- the two charge barycenters are associated to the same effective mass m_+ and m_-;
- the equilibrium position of the system is supposed to correspond to the situation where the two charge barycenters overlap, i.e. $x_+ = x_- = 0$

$$H_{intOsc} = \frac{1}{2}m\dot{x}_+^2 + \frac{1}{2}m\dot{x}_-^2 + \frac{1}{2}m\omega^2(x_+ - R)^2 + \frac{1}{2}m\omega^2(x_- - R)^2 \tag{106}$$

After the change of coordinates $x_\pm = R \pm r/2$, Eq.(106) reads:

$$\begin{aligned} H_{intOsc} &= \frac{1}{2}(2m)\dot{R}^2 + \frac{1}{2}\frac{m}{2}\dot{r}^2 + \frac{1}{2}\frac{m}{2}\omega^2 r^2 = \\ &= \frac{1}{2}m_{Tot}\dot{R}^2 + \frac{1}{2}m_{red}\dot{r}^2 + \frac{1}{2}m_{red}\omega^2 r^2 \end{aligned} \tag{107}$$

Ignoring the contribution to the total energy of the center of mass ($R = 0$, $\dot{R} = 0$), we find that the energy contribution deriving from oscillations reduces to

$$H_{intOsc} = \frac{1}{2}m_{red}\dot{r}^2 + \frac{1}{2}m_{red}\omega^2 r^2\,. \tag{108}$$

For the protein R-PE, we have that $m_{Tot} \sim 2.4 \times 10^2$ kDa, so that $m_{red} = m/2 = m_{Tot}/4 \approx 0.6 \times 10^2$ kDa.

REFERENCES AND NOTES

1. K. Venkatesan, J.-F. Rual, A. Vazquez, U. Stelzl, I. Lemmens, T. Hirozane-Kishikawa, T. Hao, M. Zenkner, X. Xin, K.-I. Goh, M. A. Yildirim, N. Simonis, K. Heinzmann, F. Gebreab, J. M. Sahalie, S. Cevik, C. Simon, A.-S. de Smet, E. Dann, A. Smolyar, A. Vinayagam, H. Yu, D. Szeto, H. Borick, A. Dricot, N. Klitgord, R. R. Murray, C. Lin, M. Lalowski, J. Timm, K. Rau, C. Boone, P. Braun, M. E. Cusick, F. P. Roth, D. E. Hill, J. Tavernier, E. E. Wanker, A.-L. Barabási, M. Vidal, An empirical framework for binary interactome mapping. *Nat. Methods* **6**, 83–90 (2009).

2. L. Bonetta, Interactome under construction. *Nature* **468**, 851–852 (2010).

3. M. Gori, I. Donato, E. Floriani, I. Nardecchia, M. Pettini, Random walk of passive tracers among randomly moving obstacles. *Theor. Biol. Med. Model.* **13**, 13–33 (2016).

4. D. S. Banks, C. Fradin, Anomalous diffusion of proteins due to molecular crowding. *Biophys. J.* **89**, 2960–2971 (2005).

5. Y. Golan, E. Sherman, Resolving mixed mechanisms of protein subdiffusion at the T cell plasma membrane. *Nat. Commun.* **8**, 15851–15866 (2017).

6. S. F. Banani, H. O. Lee, A. A. Hyman, M. K. Rosen, Biomolecular condensates: Organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017).

7. J. Berry, C. P. Brangwynne, M. Haataja, Physical principles of intracellular organization via active and passive phase transitions. *Rep. Prog. Phys.* **81**, 046601–046643 (2018).

8. L. J. Sweetlove, A. R. Fernie, The role of dynamic enzyme assemblies and substrate channelling in metabolic regulation. *Nat. Commun.* **9**, 2136–2148 (2018).

9. O. I. Kulaeva, E. V. Nizovtseva, Y. S. Polikanov, S. V. Ulianov, V. M. Studitsky, Distant activation of transcription: Mechanisms of enhancer action. *Mol. Cell. Biol.* **32**, 4892–4897 (2012).

10. J. Wang, X. Meng, H. Chen, C. Yuan, X. Li, Y. Zhou, M. Chen, Exploring the mechanisms of genome-wide long-range interactions: Interpreting chromosome organization. *Brief. Funct. Genomics* **15**, 385–395 (2016).

11. I. Wheeldon, S. D. Minteer, S. Banta, S. C. Barton, P. Atanassov, M. Sigman, Substrate channelling as an approach to cascade reactions. *Nat. Chem.* **8**, 299–309 (2016).

12. J. Preto, M. Pettini, J. A. Tuszynski, Possible role of electrodynamic interactions in long-distance biomolecular recognition. *Phys. Rev. E* **91**, 052710–052728 (2015).

13. J. R. de Xammar Oro, G. Ruderman, J. R. Grigera, Electrodynamics of interactions in electrolyte media. Possible consequences in biological functions. *Biophysics* **53**, 195–198 (2008).

14. J. C. Maxwell, *A Treatise on Electricity & Magnetism* (Dover Publications Inc., New-York, 1954).

15. P. Debye, Reaction rates in ionic solutions. *Trans. Electrochem. Soc.* **82**, 265–272 (1942).

16. R. M. Noyes, Effects of diffusion rates on chemical kinetics. *Prog. React. Kinet.* **1**, 129 (1961).

17. H. Fröhlich, Long-range coherence in biological systems. *Rivista Nuovo Cimento* **7**, 399–418 (1977).

18. A. Salam, *Molecular Quantum Electrodynamics* (John Wiley and Sons, Inc., New Jersey, 2010).

19. I. Nardecchia, J. Torres, M. Lechelon, V. Giliberti, M. Ortolani, P. Nouvel, M. Gori, Y. Meriguet, I. Donato, J. Preto, L. Varani, J. Sturgis, M. Pettini, Out-of-equilibrium collective oscillation as phonon condensation in a model protein. *Phys. Rev. X* **8**, 031061–031084 (2018).

20. P. Kurian, G. Dunston, J. Lindesayk, How quantum entanglement in DNA synchronizes double-strand breakage by type II restriction endonucleases. *J. Theor. Biol.* **391**, 102–112 (2016).

21. P. Kurian, A. Capolupo, T. J. A. Craddock, G. Vitiello, Water-mediated correlations in DNA-enzyme interactions. *Phys. Lett. A* **382**, 33–43 (2018).

22. Z. Zhang, G. S. Agarwal, M. O. Scully, Quantum fluctuations in the Fröhlich condensate of molecular vibrations driven far from equilibrium. *Phys. Rev. Lett.* **122**, 158101–158107 (2019).

23. J. Preto, E. Floriani, I. Nardecchia, P. Ferrier, M. Pettini, Experimental assessment of the contribution of electrodynamic interactions to long-distance recruitment of biomolecular partners: Theoretical basis. *Phys. Rev. E* **85**, 041904–041916 (2012).

24. I. Nardecchia, L. Spinelli, J. Preto, M. Gori, E. Floriani, S. Jaeger, P. Ferrier, M. Pettini, Experimental detection of long-distance interactions between biomolecules through their diffusion behavior: Numerical study. *Phys. Rev. E* **90**, 022703–022717 (2014).

25. I. Nardecchia, M. Lechelon, M. Gori, I. Donato, J. Preto, E. Floriani, S. Jaeger, S. Mailfert, D. Marguet, P. Ferrier, M. Pettini, Detection of long-range electrostatic interactions between charged molecules by means of fluorescence correlation spectroscopy. *Phys. Rev. E* **96**, 022403–022416 (2017).

26. S. Olmi, M. Gori, I. Donato, M. Pettini, Collective behavior of oscillating electric dipoles. *Sci. Rep.* **8**, 15748–15760 (2018).

27. W. R. Chang, T. Jiang, Z. L. Wan, J. P. Zhang, Z. X. Yang, D. C. Liang, Crystal structure of R-phycoerythrin from *Polysiphonia urceolata* at 2.8 Å resolution. *J. Mol. Biol.* **262**, 721–731 (1996).

28. R. D. Blevins, *Formulas for Natural Frequency and Mode Shape* (Van Nostrand Reinhold Company, New-York, 1979).

29. S. Perticaroli, J. D. Nickels, G. Ehlers, H. O'Neill, Q. Zhang, A. P. Sokolov, Secondary structure and rigidity in model proteins. *Soft Matter* **9**, 9548–9556 (2013).

30. F. Schuster, D. Coquillat, H. Videlier, M. Sakowicz, F. Teppe, L. Dussopt, B. Giffard, T. Skotnicki, W. Knap, Broadband terahertz imaging with highly sensitive silicon CMOS detectors. *Opt. Express* **19**, 7827–7832 (2011).

31. R. Van Wijk, E. P. A. Van Wijk, J. Pang, M. Yang, Y. Yan, J. Han, Integrating ultra-weak photon emission analysis in mitochondrial research. *Front. Physiol.* **11**, 717 (2020).

32. G. L. Celardo, M. Angeli, T. J. A. Craddock, P. Kurian, On the existence of superradiant excitonic states in microtubules. *New J. Phys.* **21**, 023005 (2019).

33. M. Stöhr, A. Tkatchenko, Quantum mechanics of proteins in explicit water: The role of plasmon-like solute-solvent interactions. *Sci. Adv.* **5**, eaax0024 (2019).

34. D. Tischer, O. D. Weiner, Illuminating cell signalling with optogenetic tools. *Nat. Rev. Mol. Cell Biol.* **15**, 551–558 (2014).

35. N. Veljkovic, S. Glisic, V. Perovic, V. Veljkovic, The role of long-range intermolecular interactions in discovery of new drugs. *Expert Opin. Drug Discov.* **6**, 1263–1270 (2011).

36. P. C. Painter, L. E. Mosher, C. Rhoads, Low-frequency modes in the Raman spectra of proteins. *Biopolymers* **21**, 1469–1472 (1982).

37. P. C. Painter, L. Mosher, C. Rhoads, Low-frequency modes in the Raman spectrum of DNA. *Biopolymers* **20**, 243–247 (1981).

38. H. Fröhlich, Long-range coherence and energy storage in biological systems. *Int. J. Quantum Chem.* **2**, 641–649 (1968).

39. H. Fröhlich, Long range coherence and the action of enzymes. *Nature* **228**, 1093 (1970).

40. H. Fröhlich, Selective long range dispersion forces between large systems. *Phys. Lett. A* **39**, 153–154 (1972).

41. E. Haustein, P. Schwille, Fluorescence correlation spectroscopy: Novel variations of an established technique. *Annu. Rev. Biophys. Biomol. Struct.* **36**, 151–169 (2007)..

42. Z. Petrášek, P. Schwille, Precise measurement of diffusion coefficients using scanning fluorescence correlation spectroscopy. *Biophys. J.* **94**, 1437–1448 (2008).

43. T. Laurence, S. Ly, F. Bourguet, N. O. Fischer, M. A. Coleman, Fluorescence correlation spectroscopy at micro-molar concentrations without optical nanoconfinement. *J. Phys. Chem.* **118**, 9662–9667 (2014).

44. S. Khatua, H. Yuan, M. Orrit, Enhanced-fluorescence correlation spectroscopy at micro-molar dye concentration around a single gold nanorod. *Phys. Chem. Chem. Phys.* **17**, 21127–21132 (2015).

45. M. Zhao, L. Jin, B. Chen, Y. Ding, H. Ma, D. Chen, After pulsing and its correction in fluorescence correlation spectroscopy experiments. *Appl. Optics* **42**, 4031–4036 (2003).

46. D. Magde, E. Elson, A. Hyman, W. Webb, Fluorescence correlation spectroscopy. II. An experimental realization. *Biopolymers* **13**, 29–61 (1974).

47. F. Keilmann, FIR microscopy. *Infrared Phys. Techno.* **36**, 217–224 (1995).

48. F. M. Richards, Areas, volumes, packing, and protein structure. *Annu. Rev. Biophys. Bioeng.* **6**, 151–176 (1977).

49. H. Durchschlag, Determination of the partial specific volume of conjugated proteins. *Colloid Polym. Sci.* **267**, 1139–1150 (1989).

50. P.-H. Chavanis, Phase transitions in self-gravitating systems: Self-gravitating fermions and hard-sphere models. *Phys. Rev. E* **65**, 056123 (2002).

51. E. R. Likhachev, Dependence of water viscosity on temperature and pressure. *Tech. Phys.* **48**, 514–515 (2003).

52. J. N. Israelachvili, *Intermolecular and Surface Forces* (Academic press, 2011).

53. M. Farnum, C. Zukoski, Effect of glycerol on the interactions and solubility of bovine pancreatic trypsin inhibitor. *Biophys. J.* **76**, 2716–2726 (1999).

54. M. Lund, B. Jönsson, A mesoscopic model for protein-protein interactions in solution. *Biophys. J.* **85**, 2940–2947 (2003).

55. E. Faraji, R. Franzosi, S. Mancini, M. Pettini, Energy transfer to the phonons of a macromolecule through light pumping. *Sci. Rep.* **11**, 6591–6605 (2021).

56. I. Y. Torshin, R. W. Harrison, Charge centers and formation of the protein folding core. *Proteins* **43**, 353–364 (2001).

57. S. Tsonchev, G. C. Schatz, M. A. Ratner, Screened multipole electrostatic interactions at the Debye–Hückel level. *Chem. Phys. Lett.* **400**, 221–225 (2004).

58. B. Alberts, A. Johnson, J. Lewis, P. Walter, M. Raff, K. Roberts, *Molecular Biology of the Cell 4th Edition: International Student Edition* (Routledge, 2002).

59. J. P. T. Zaragoza, A. Nguy, N. Minnetian, Z. Deng, A. T. Iavarone, A. R. Offenbacher, J. P. Klinman, Detecting and characterizing the kinetic activation of thermal networks in proteins: Thermal transfer from a distal, solvent-exposed loop to the active site in soybean lipoxygenase. *J. Phys. Chem. B* **123**, 8662–8674 (2019).

60. A. Campa, T. Dauxois, S. Ruffo, Statistical mechanics and dynamics of solvable models with long-range interactions. *Phys. Rep.* **480**, 57–159 (2009).

61. V. V. Shcherbakov, Yu. M. Artemkina, E. N. Korotkova, Dielectric properties and high-frequency conductivity of the sodium chloride-water system. *Russ. J. Inorg. Chem.* **59**, 922–926 (2014).