Supporting Information

Mazza et al. 10.1073/pnas.1104299108

SI Text

I. Quality of Experimental Setup and Data Analysis. A "third check" of the validity of our data analysis procedure is provided by comparing our results with recent data from other experiments that do not use blocking electrodes. In Fig. S1, we compare our data with recent dielectric data on the same hydrated lysozyme powder and on similarly hydrated myoglobin powders. These data were collected by researchers (1, 2) using a different experimental setup—one without blocking electrodes—and were analyzed using a different deconvolution method. Note that our data for the "side-chain relaxation" and the "main relaxation" overlap with the data from these other experiments, thus validating both our experimental setup and our data analysis.

In particular, a recent dielectric study on the same lysozyme protein (1) revealed the presence of two relaxation processes when the protein water content was $h \leq 0.3$ g H_2O/g dry protein. In ref. 3, the slower process has been assigned to the motion of the protein "side chains," or to a cross-term due to the relaxation of water and protein dipoles, and the faster process has been designated "main." Here, we adopt the same nomenclature for the relaxation process. The associated relaxation times (Fig. S1) are clearly consistent with the relaxation times we labeled as "side chain" and "main," both in terms of their temperature dependence and their absolute values.

In addition, the relaxation time observed for myoglobin embedded in poly(vinyl) alcohol at $h = 0.4$ g H_2O/g dry protein (2) has absolute values and T dependence that are identical to our side-chain relaxation (identified as the "β-relaxation" observed in many glass-forming liquids) assigned to the dynamics of the protein hydration shell (2).

Note once again that the experimental setup adopted in refs. 1 and 2 makes no use of blocking electrodes. Moreover, the authors of refs. 1 and 2 used a single relaxation mode to fit the low-frequency process that coincides with our side chain, despite the fact that it has a remarkably broad frequency spectrum. We instead decompose it into two contributions and include the "proton" relaxation. Yet our results agree well with those reported in refs. 1 and 2 for temperatures that have also been studied by other researchers. When we also examine the behavior below 180 K, we find a crossover not observed by other researchers, possibly due to lack of data.

Because these relaxation times are consistent with our findings, in terms of both temperature dependence and absolute values, we conclude that our data analysis of the results obtained with blocking electrodes and with our deconvolution method is sufficiently robust to reproduce data collected using a different experimental setup with no blocking electrodes and with a different method of analysis.

To better quantify the accuracy of our fitting procedure, which includes side-chain relaxation and proton relaxation to account for the remarkably broad frequency spectrum of the low-frequency process, we provide in Figs. S2 and S3 the details of our deconvolution method.

Fig. S2 shows the data for the real (left axis) and imaginary (right axis) components of the complex permittivity (open symbols) at $T = 245.2$ K. We choose this T because it is one of the worst fit cases (see Fig. 3 in the main text), with the deconvolution that results in two relaxation processes with a considerable overlap. Nevertheless, even in this worst-case scenario the error of our analysis is minimized when compared to an analysis that includes only one relaxation mode for the low-frequency process.

The experimental accuracy is within the size of the symbols used in Fig. S2. The continuous lines through the symbols are the result of the simultaneous fit of both components of complex permittivity. The vertical axis on a log scale reveals small deviations in the fits. The thick black line and the dashed line are the fitting results for the imaginary component, due to the proton relaxation and the side-chain relaxation, respectively. For the sake of clarity, other low-frequency contributions to the measured data (electrode polarization, Maxwell–Wagner relaxation, local conductivity) are not shown, as well as the high-frequency main relaxation that is entering the frequency window investigated at this temperature.

Fig. S3 shows the fit residue of each component of the complex permittivity (both components have the same residue within the numerical precision) as a function of the data point, 0 being the lowest angular frequency and 81 the highest, for the same data-set shown in Fig. S2. The fitting procedure minimizes iteratively the difference between the measured $\epsilon^*_m(\omega)$ and the calculated $\epsilon^*_c(\omega)$ values simultaneously for the real, $\Delta\epsilon' = \epsilon'_m - \epsilon'_c$, and imaginary, $\Delta\epsilon'' = \epsilon''_m - \epsilon''_c$, components of the complex permittivity.

The fit residue is scattered around zero, within $\pm 2 \times 10^{-3}$ for almost the entire frequency range investigated, indicating that the fit describes the data well with a relative error of $\lesssim$0.2% for the fitted points and the fitting parameters. For $1/T > 6 \times 10^{-3}$ K^{-1}, the proton relaxation process moves outside the experimentally accessible frequency window, but its shape changes continuously with T, allowing a consistent estimate of the fitting parameters with a relative error that is at least one order of magnitude larger than that for higher T.

Two further considerations are in order here. First, the quality of decomposition is at a maximum if two relaxation modes—not one or three—are used (see section 2 of ref. 4 for details). In particular, although the fit with two relaxation terms is statistically better than the fit with a single relaxation term, the addition of a third relaxation term in the same frequency range as the first two yields a nonphysical negative relaxation amplitude ($\Delta\epsilon_j$ in Eq. **6** of the main text) for one of the processes. Thus, the possible number of relaxation terms is restricted to one or two. As shown in figure 2 of ref. 4 for similar experimental conditions, the resulting distribution function of relaxation times derived for a single relaxation process using a standard method (5) does not totally account for the distribution function obtained in the time domain (6) using an inverse Laplace transform with no a priori assumption of the type and number of relaxation processes from raw $\epsilon^*_m(\omega)$ data (7). Instead, adding a second relaxation is sufficient to completely overlap the experimental distribution function of relaxation times.

Second, we perform an analysis that takes into account the data at all temperatures, with a deconvolution of the data that preserves a smooth variation of the values of the fitting parameters. Once the analysis of the distribution function of relaxation times allows us to conclude that there is a proton relaxation that contributes with a characteristic shape to the total relaxation, we follow the temperature dependence of the proton relaxation, along with that of the other contributions, for the entire dataset available. Only this global fitting procedure can correctly describe the data.

II. Simulations. ***Monte Carlo for the thermodynamics.*** To avoid the slow equilibration times at low T and the critical slowing of standard Monte Carlo (MC) dynamics (e.g., Metropolis) in the vicinity of a critical point, we extend the cluster algorithm proposed by Wolff (8) to the model considered here (9). The MC dynamics

consists of randomly selecting a bonding index σ_{ij}, and growing a cluster of thermodynamically correlated bond variables following the Wolff algorithm (8). If the remaining bonding indices of the same initial molecule are in the same Potts state, then they are added to the stack with probability $p_{\text{same}} \equiv \min(1, 1 - \exp(-\beta J_\sigma))$, where $\beta \equiv (k_B T)^{-1}$ and k_B is the Boltzmann constant. This choice for the probability p_{same} depends on the interaction J_σ and guarantees that the connected bonding indices are thermodynamically correlated (10–12). We next consider the bonding index of a new molecule that faces the initially chosen bonding index. To guarantee that connected facing bonding indices correspond to thermodynamically correlated variables, we link them with the probability $p_{\text{facing}} \equiv \min(1, 1 - \exp(-\beta J'))$ where $J' \equiv J - P v_{\text{HB}}$ is the P-dependent effective coupling between two facing bonding indices that result from the enthalpy of the system. Depending on P, J' can be positive or negative. If $J' > 0$ and the two facing bonding indices are in the same state, then the new bonding index is added to the cluster with probability p_{facing}. If $J' < 0$ and the two facing bonding indices are in different states, then the new bonding index is added with probability p_{facing}. Only after every possible direction of growth for the cluster has been considered are the values of the bonding indices changed in a stochastic way. Again we consider two cases:

i. If $J' > 0$, all bonding indices are set to the same new value

$$\sigma^{\text{new}} = (\sigma^{\text{old}} + \phi) \qquad \text{mod}\, q, \qquad \textbf{[S1]}$$

where ϕ is a random integer between 1 and q.

ii. If $J' < 0$, the state of every bonding index is changed (rotated) by the same random constant $\phi \in [1, \ldots q]$

$$\sigma_i^{\text{new}} = (\sigma_i^{\text{old}} + \phi) \qquad \text{mod}\, q. \qquad \textbf{[S2]}$$

To implement an isobaric ensemble we consider increments ΔV (positive or negative) to the volume. This global change in volume is accepted with probability

$$\min(1, \exp[-\beta(\Delta E + P\Delta V - N k_B T \ln(V_f / V_i))]), \qquad \textbf{[S3]}$$

where $\beta \equiv (k_B T)^{-1}$; V_i and V_f are, respectively, the initial and final values of the volume; $\Delta V \equiv V_f - V_i$; $\Delta E \equiv E_f - E_i$; and E_i and E_f are, respectively, the initial and final values of the internal energy.

We define an MC step as $4N$ updates of the bond indices followed by a volume update (i.e., as $4N + 1$ steps of the algorithm). We control the equilibration by making sure that variables such as energy, volume, and number of hydrogen bonds (HBs) show a constant average value with respect to the MC steps. We also make sure that response function values, e.g., specific heat, calculated from the temperature derivative of enthalpy are the same (i.e., within the error bars) as the values calculated from the energy fluctuation using the fluctuation-dissipation theorem. To compare the MC results with the experiments, we rescale them as described below.

We simulate the system for 10^7 MC steps and observe that, because the cluster MC can make nonlocal moves in the configuration space, the Wolff algorithm is 10^4–10^6 times faster than the standard Metropolis MC (9). The Wolff MC thus allows jumps between configurations that are far apart in the energy landscape, but with an energy difference of the order of $k_B T$.

MC for the dynamics. When we study the correlation function

$$C_M(t) \equiv \frac{1}{N} \sum_{i=1}^{N} \frac{\langle M_i(t_0 + t) M_i(t_0) \rangle - \langle M_i \rangle^2}{\langle M_i(t_0)^2 \rangle - \langle M_i \rangle^2}, \qquad \textbf{[S4]}$$

we cannot use the Wolff algorithm because of its nonlocal characteristics. Instead we use the local Metropolis MC in which the state of each randomly chosen variable σ_{ij} is randomly set to a new value σ_{ij}^{new} with a probability given by Eq. **S3**, where the old and the new states (and their volume and energy) are determined by σ_{ij} and σ_{ij}^{new}, respectively.

For the Metropolis algorithm we also define an MC step as $4N$ updates of the bond indices followed by a volume update (i.e., as $4N + 1$ steps of the algorithm) and we calculate the correlation time τ_{MC} as described in the text. To compare with experimental data, we rescale τ_{MC} with respect to the experimental data, adopting a logarithmic rescale between the MC steps and seconds: $\ln(\tau_{\text{MC}}[\text{s}]) = -31.3 + 1.74 \ln \tau[\text{MC Steps}]$. This rescaling implies that the lower the T, the longer will be the time in seconds for a Metropolis MC step, and that the increase in the time scale for decreasing T is logarithmic, consistent with the logarithmic collapse of time scales for supercooled water.

1. Khodadadi S, Pawlus S, Sokolov AP (2008) Influence of hydration on protein dynamics: Combining dielectric and neutron scattering spectroscopy data. *J Phys Chem B* 112:14273–14280.
2. Frauenfelder H, et al. (2009) A unified model of protein dynamics. *Proc Natl Acad Sci USA* 106:5129–5134.
3. Khodadadi S, et al. (2008) The origin of the dynamic transition in proteins. *J Chem Phys* 128:195106.
4. Bruni F, Pagnotta SE (2004) Dielectric investigation of the temperature dependence of the dynamics of a hydrated protein. *Phys Chem Chem Phys* 6:1912–1919.
5. Raicu V (1999) Dielectric dispersion of biological matter: Model combining Debye-type and "universal" responses. *Phys Rev E Stat Nonlin Soft Matter Phys* 60:4677–4680.
6. Alvarez F, Alegria A, Colmenero J (1991) Relationship between the time-domain Kohlrausch-Williams-Watts and frequency-domain Havriliak-Negami relaxation functions. *Phys Rev B Condens Matter* 44:7306–7312.
7. Provencher SW (1982) A constrained regularization method for inverting data represented by linear algebraic or integral equations. *Comput Phys Commun* 27:213–227.
8. Wolff U (1989) Collective Monte Carlo updating for spin systems. *Phys Rev Lett* 62:361–364.
9. Mazza MG, Stokely K, Strekalova EG, Stanley HE, Franzese G (2009) Cluster Monte Carlo and numerical mean field analysis for the water liquid-liquid phase transition. *Comput Phys Commun* 180:497–502.
10. Coniglio A, Peruggi F (1982) Clusters and droplets in the Q-state Potts-model. *J Phys A Math Gen* 15:1873–1883.
11. Franzese G (1996) Cluster analysis for percolation on a two-dimensional fully frustrated system. *J Phys A Math Gen* 29:7367–7375.
12. Cataudella V, Franzese G, Nicodemi M, Scala A, Coniglio A (1996) Percolation and cluster Monte Carlo dynamics for spin models. *Phys Rev E Stat Nonlin Soft Matter Phys* 54:175–189.

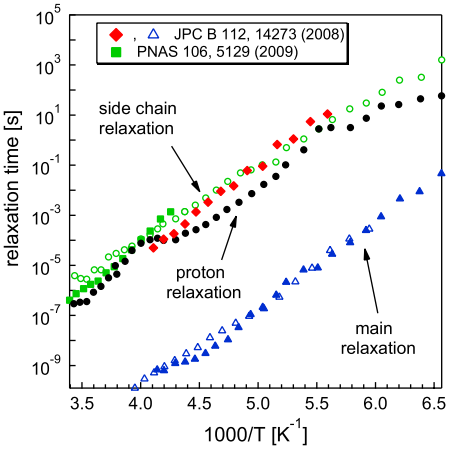

Fig. S1. The comparison of our data with those from previous experiments shows that they follow the same temperature dependence and overlap within the experimental precision for those temperatures at which data from others are available. In particular, although our data for side-chain relaxation (open circles), proton relaxation (solid circles), and "main" relaxation (solid triangles) extend below the crossover at about 181 K, side-chain relaxation from ref. 1 (solid diamonds) and from ref. 2 (solid squares), and "main" relaxation from ref. 1 (open triangles) do not. Our data are for lysozyme powder with hydration level $h = 0.30$ g H_2O/g dry protein, as described in the manuscript. Data from ref. 1 are for the same protein powder at the same hydration. Data from ref. 2 are for embedded myoglobin at similar hydration.

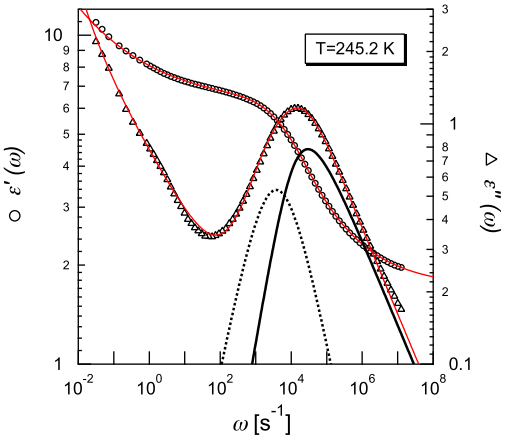

Fig. S2. Example of fitting procedure for our dielectric relaxation data of hydrated lysozyme. Open symbols are the data for the real component (ϵ', circles and left axis) and imaginary component (ϵ'', triangles and right axis) of the measured complex permittivity at $T = 245.2$ K. Lines through the symbols are the result of the simultaneous fitting of both components. The black thick line indicates the imaginary component of the proton relaxation, and the dashed line indicates the imaginary component of the side-chain relaxation. Other contributions to the spectrum (electrode polarization, Maxwell–Wagner relaxation, and local conductivity at low frequency; main relaxation at high frequency) are not shown for the sake of clarity.

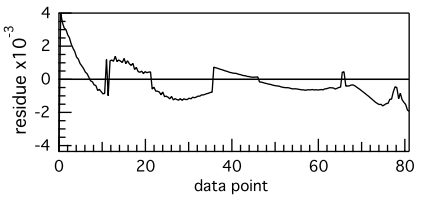

Fig. S3. Fitting residue of the dataset in Fig. S2. The x axis indicates the data point (namely, 0 corresponds to the lowest angular frequency, and 81 to the highest reported in Fig. S2). Deviation from zero is of the order of $\pm 2 \times 10^{-3}$ over almost the entire frequency window investigated, indicating relatively small errors associated to the fitting procedure.