Dynamics of a globular protein and its hydration water studied by neutron scattering and MD simulations

Sow-Hsin Chen \(^{a,\ast}\), Marco Lagi \(^{a,b}\), Xiang-qiang Chu \(^{a}\), Yang Zhang \(^{a}\), Chansoo Kim \(^{c}\), Antonio Faraone \(^{d,e}\), Emiliano Fratini \(^{b}\) and Piero Baglioni \(^{b}\)

\(^{a}\) Department of Nuclear Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA

\(^{b}\) Department of Chemistry and CSGI, University of Florence, Florence, Italy

\(^{c}\) Computational Science Center, KIST, Seongbuk-gu, Seoul, Korea

\(^{d}\) NIST Center for Neutron Research, Gaithersburg, MD, USA

\(^{e}\) Department of Material Science and Engineering, University of Maryland, College Park, MD, USA

Abstract. This review article describes our neutron scattering experiments made in the past four years for the understanding of the single-particle (hydrogen atom) dynamics of a protein and its hydration water and the strong coupling between them. We found that the key to this strong coupling is the existence of a fragile-to-strong dynamic crossover (FSC) phenomenon occurring at around \(T^* = 225 \pm 5\) K in the hydration water. On lowering of the temperature toward FSC, the structure of hydration water makes a transition from predominantly the high density form (HDL), a more fluid state, to predominantly the low density form (LDL), a less fluid state, derived from the existence of a liquid–liquid critical point at an elevated pressure. We show experimentally that this sudden switch in the mobility of hydration water on Lysozyme, B-DNA and RNA triggers the dynamic transition, at a temperature \(T_D = 220\) K, for these biopolymers. In the glassy state, below \(T_D\), the biopolymers lose their vital conformational flexibility resulting in a substantial diminishing of their biological functions. We also performed molecular dynamics (MD) simulations on a realistic model of hydrated lysozyme powder, which confirms the existence of the FSC and the hydration level dependence of the FSC temperature. Furthermore, we show a striking feature in the short time relaxation (\(\beta\)-relaxation) of protein dynamics, which is the logarithmic decay spanning 3 decades (from ps to ns). The long time \(\alpha\)-relaxation shows instead a diffusive behavior, which supports the liquid-like motions of protein constituents. We then discuss our recent high-resolution X-ray inelastic scattering studies of globular proteins, Lysozyme and Bovine Serum Albumin. We were able to measure the dispersion relations of collective, intra-protein phonon-like excitations in these proteins for the first time. We found that the phonon energies show a marked softening and at the same time their population increases substantially in a certain wave vector range when temperature crosses over the \(T_D\). Thus the increase of biological activities above \(T_D\) has positive correlation with activation of slower and large amplitude collective motions of a protein.

1. Introduction

While water has been considered as “life’s solvent” (in a passive sense) for a long time, only in the past 20 years has it become an active constituent of cell biochemistry and not just a uniform background...
One of the most striking examples of the importance of water in biosystems is that proteins cannot perform their function if they are not covered by a minimum amount of hydration water. Hydration can be considered as a process, that of adding water incrementally to dry protein, until a level of hydration is reached beyond which further addition of water produces no change of the essential properties of the protein and only dilutes the protein [49]. The hydration shell can be defined as the water associated with the protein at the hydration end point. This shell represents a monolayer coverage of the protein surface. Water outside the monolayer is perturbed to a significantly smaller extent, typically not detected by measurements of properties such as heat capacity, volume or heat content.

Rupley et al. [50] measured the reaction of lysozyme with the hexasaccharide of N-acetylglucosamine over the full hydration range. The threshold hydration level was \( h = 0.2 \), where \( h \) is the ratio between grams of water and grams of dry protein. They showed that enzymatic activity closely parallels the development of surface motion, which is thus responsible for the functionality of the protein.

Around \( T_D = 220 \) K, the protein has a transition that could be described as a dynamic transition or a so-called glass transition [19,55]. An analogous transition at \( T_L \) can also be detected in the protein hydration water at the same temperature. Some IR data collected by Doster et al. [18] suggested that the transition in the hydration water could be described as the melting of amorphous ice and that this solvent network is composed of water clusters with relatively strong internal bonding. They used this information to address the problem of dynamic coupling of solvent motions with internal protein motions, suggesting that the cooperativity of the solvent network provides the coupling mechanism. Our recent Quasi-Elastic Neutron Scattering experiments [12] suggest however that this dynamic crossover in hydration water is the result of a transition from predominantly low density form of water at lower temperature (LDL), a less fluid state, to predominantly high density form at higher temperature (HDL), a more fluid state, derived from the existence of a second (liquid–liquid) critical point at an elevated pressure [34].

In this article, we shall give a comprehensive description of the Elastic and Quasi-Elastic Neutron Scattering experiments as a function of temperature at different pressures which substantiate the above statement [14]. We shall also supplement these experimental data with a MD Simulation [36] to elucidate the role of coupling between dynamics of hydration water and the protein.

2. Incoherent neutron scattering for studying dynamics of hydration water

Incoherent Elastic Neutron Scattering (IENS) and Quasi-Elastic Neutron Scattering (QENS) methods offer many advantages for the study of hydrogen atom dynamics in a protein and its hydration water. The main reason is that the scattering cross section of hydrogen is about 80 barns, and it is much larger (at least 20 times) than that of other atoms in the protein–hydration–water system, composed of oxygen, carbon, nitrogen and sulfur atoms. Furthermore, neutron scattering cross section of a hydrogen atom is mostly incoherent so that QENS spectra reflect, essentially, the self-dynamics of the hydrogen atoms in the protein or water. Combining this dominant cross section of hydrogen atoms with the use of spectrometers having different energy resolutions, we can study the molecular dynamics of water in a wide range of time-scale, encompassing picoseconds to tens of nanoseconds.

It can be shown generally [13] that the double differential scattering cross section is proportional to the self-dynamic structure factor of hydrogen atoms \( S_H(Q, E) \) through the following relation:

\[
\frac{d^2\sigma_H}{d\Omega dE} = N \frac{\sigma_H}{4\pi \hbar} k_f S_H(Q, E),
\]  

(1)
where $E = E_i - E_f = \hbar \omega$ is the energy transferred by a neutron to the sample in the collision process; and $\hbar \vec{Q} = \hbar \vec{k}_i - \hbar \vec{k}_f$, the momentum transferred in the scattering process; and $N$, the number of hydrogen atoms in the scattering volume. The self-dynamic structure factor, $S_H(Q, E)$ embodies the elastic, quasi-elastic and inelastic scattering contributions. It can be expressed as a Fourier transform of the self-ISF of a typical hydrogen atom according to:

$$S_H(Q, E) = \frac{1}{2\pi \hbar} \int_{-\infty}^{\infty} dt \, e^{-iEt/\hbar} F_H(Q, t).$$  \hspace{1cm} (2)

$F_H(Q, t)$ is the atom density–density time correlation function of the tagged hydrogen atom being measured by the neutron scattering. It is, thus, the primary quantity of theoretical interest related to the experiment. It can be calculated by a model, such as the relaxing cage model (RCM), or by a molecular dynamics (MD) simulation based on a phenomenological potential model of water.

### 2.1. Incoherent elastic scattering ($E = 0$)

The intermediate scattering function (ISF) for a hydrogen atom harmonically bound to a molecule can be written as

$$F_H(Q, t) = \langle \exp(-iQX_H(0)) \exp(iQX_H(t)) \rangle,$$  \hspace{1cm} (3)

where $Q$ is the magnitude of the $\vec{Q}$ vector, pointing in the $x$-direction in the isotropic powder sample. Then it can be shown that in the Gaussian approximation, which is exact for the harmonically bound particle, one can write [9]

$$F_H(Q, t) = \exp(-Q^2\langle X_H^2 \rangle) \exp(Q^2\langle X_H(0)X_H(t) \rangle),$$  \hspace{1cm} (4)

where the first factor, $\exp(-Q^2\langle X_H^2 \rangle)$, is called the Debye–Waller factor, which gives rise to the elastic scattering, and the second factor, which involves the displacement–displacement time correlation function, gives rise to the inelastic scattering such as phonons. In the classical regime, the Eq. (4) can further be written into the form

$$F_{cl}^H(Q, t) = \exp(-\frac{1}{2}Q^2W(t)),$$  \hspace{1cm} (5)

where the width function can be written as [9]

$$W(t) = 2V_0^2 \int_0^{\infty} d\omega \frac{f_H(\omega)}{\omega^2} (1 - \cos \omega t),$$  \hspace{1cm} (6)

where $f_H(\omega)$ is the Fourier transform of the normalized velocity correlation function of a hydrogen atom, which is sometime called the spectral density function of the hydrogen atom.

$$f_H(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dt \, e^{i\omega t} \frac{\langle V_x^H(0)V_x^H(t) \rangle}{\langle (V_x^H)^2 \rangle},$$  \hspace{1cm} (7)
where \((V_{HH}^2) = V_0^2 = k_B T / M_H\). In Eq. (4), the elastic scattering corresponds to \(t = \infty\), where the second factor becomes unity, and \(F_{HH}(Q, \infty) = \exp(-Q^2 \langle X_{HH}^2 \rangle)\), which is just the Debye–Waller factor. Combining this equation with Eqs (5) and (6), we finally obtain a very useful result,

\[
\langle X_{HH}^2 \rangle = \frac{1}{2} W(\infty) = V_0^2 \int_0^\infty d\omega \frac{f_H(\omega)}{\omega^2}.
\] (8)

The mean-square deviation (MSD) of the hydrogen atoms can be obtained from the frequency integral of the reduced spectral density function of the hydrogen atoms.

### 2.2. Incoherent quasi-elastic scattering \((E \approx 0)\) of protein hydration water

In principle, the single-particle dynamics of bulk or confined water should include both the translational and the rotational motions of a rigid water molecule. Given the fact that in the process of QENS data analysis, we only focus our attention to ISF with \(Q \leq 1.1 \text{Å}^{-1}\), we can safely neglect the contribution of the rotational motion to the total dynamics [7], which means \(F_{HH}(Q, t) \approx F_T(Q, t)\), where \(F_T(Q, t)\) is the translational part of the ISF.

During the past several years, we have developed the relaxing cage model (RCM) for the description of the translational and the rotational dynamics of water at supercooled temperatures. This model has been tested with MD simulations of SPC/E water, and has been found to be accurate. It has been used to analyze many QENS data from supercooled bulk water as well as interfacial water [21,22,24,25,42].

On lowering the temperature below the freezing point, around a given water molecule, there is a tendency to form a hydrogen-bonded, tetrahedrally coordinated first and second neighbor shells (cage). At short times, less than 0.05 ps, the center of mass of a water molecule performs vibrations inside the cage. At long times, longer than 1.0 ps, the cage eventually relaxes and the trapped particle can migrate through the rearrangement of a large number of particles surrounding it. Therefore, there is a strong coupling between the single particle motion and the density fluctuations of the liquid. The mathematical expression of this physical picture is the so-called RCM.

The RCM assumes that the short-time translational dynamics of the tagged (or the trapped) water molecule can be treated approximately as the motion of the center of mass in an isotropic harmonic potential well provided by its neighbors. We can, then, write the short-time part of the translational ISF in the Gaussian approximation, connecting it to the velocity auto-correlation function, \(\langle \vec{v}_{CM}(t) \cdot \vec{v}_{CM}(0) \rangle\), in the following way:

\[
F_T^s(Q, t) = \exp \left( -\frac{Q^2}{2} \langle r_{CM}^2(t) \rangle \right)
= \exp \left( -Q^2 \left[ \int_0^t (t - \tau) \langle \vec{v}_{CM}(0) \cdot \vec{v}_{CM}(\tau) \rangle d\tau \right] \right),
\] (9)

which can be further derived as [42]

\[
F_T^s(Q, t) = \exp \left\{ -Q^2 v_0^2 \left[ \frac{(1 - C)}{\omega_1^2} \left( 1 - \exp \left( -\frac{\omega_1^2 t^2}{2} \right) \right) + \frac{C}{\omega_2^2} \left( 1 - \exp \left( -\frac{\omega_2^2 t^2}{2} \right) \right) \right] \right\}.
\] (10)
Equation (10) is the short-time behavior of the translational ISF. It starts from unity at \( t = 0 \) and decays rapidly to a flat plateau determined by an incoherent Debye–Waller factor \( A(Q) \), given by

\[
A(Q) = \exp\left\{ -Q^2 v_0^2 \left( \frac{1 - C}{\omega_i^2} + \frac{C}{\omega_s^2} \right) \right\} = \exp\left\{ -Q^2 a^2/3 \right\},
\]  

(11)

where \( a \) is the root mean square vibrational amplitude of the water molecules in the cage, in which the particle is constrained during its short-time movements. According to MD simulations, \( a \approx 0.5 \) Å is fairly temperature independent [27].

On the other hand, the cage relaxation at long-time can be described by the standard \( \alpha \)-relaxation model, according to the Mode-Coupling Theory (MCT), with a stretched exponential having a structural relaxation time \( \tau_T \) and a stretch exponent \( \beta \). Therefore, the translational ISF, valid for the entire time range, can be written as a product of the short-time part and a long-time part:

\[
F_T(Q, t) = F_s(T)(Q, t) \exp\left\{ -\left( \frac{t}{\tau_T} \right)^\beta \right\}.
\]  

(12)

The fit of the MD generated \( F_T(Q, t) \) using Eq. (12) shows that \( \tau_T \) is \( Q \)-dependent, obeying the power-law:

\[
\tau_T = \tau_0(aQ)^{-\gamma},
\]  

(13)

where \( \gamma \) is \( \leq 2 \), with a slight dependency on \( Q \) and \( T \), and \( \beta < 1 \) is slightly \( Q \) and \( T \) dependent as well. In the \( Q \to 0 \) limit, one should approach the diffusion limit, where \( \gamma \to 2 \) and \( \beta \to 1 \). Thus the translational ISF can be written as: \( F_T(Q, t) = \exp\left\{ -DQ^2 t \right\} \), \( D \) being the self-diffusion coefficient. In QENS experiments, this low \( Q \) limit is not usually reached, and both \( \beta \) and \( \gamma \) can be considered \( Q \)-independent in the limited \( Q \) range of \( 0 \leq Q \leq 1 \) [22].

We define a \( Q \)-independent average translational relaxation time

\[
\langle \tau_T \rangle = \frac{\tau_0}{\beta} \Gamma(1/\beta),
\]  

(14)

which is a convenient quantity to be extracted from the experimental data by the fitting process of RCM. This quantity can be identified to be proportional to the \( \alpha \)-relaxation time which dominates the long-time decay of the ISF in low temperature water [42].

In actual QENS experiment on hydrated biomolecules, we have to take into account the signal coming from the hydrogen atoms in the hydration water only, by taking the difference of the spectra of H2O and D2O hydrated samples [17,51]. Denoting the fraction of the elastic scattering coming from the bound hydrogen atom in protein by \( p \) we can analyze the experimental data according to the following model:

\[
S(Q, \omega) = pR(Q_0, \omega) + (1 - p) FT\{F_H(Q, t)R(Q_0, t)\},
\]  

(15)

where \( F_H(Q, t) \approx F_T(Q, t) \) is the ISF of hydrogen atoms which defines the quasi-elastic scattering, \( R(Q_0, t) \) is the experimental resolution function, and the symbol \( FT \) denotes the Fourier transform from time \( t \) to frequency \( \omega \).
3. The elastic scan and the MSD of hydrogen atoms in protein and its hydration water

To obtain the MSD \( \langle x^2 \rangle \) of hydrogen atoms, we perform fixed window scan (an elastic scattering measurement with a fixed resolution window of FWHM of \( \pm 0.8 \mu \text{eV} \)) [2] in the temperature range from 40 to 290 K, covering completely the supposed crossover temperature \( T_L \). Since the system is in a stationary metastable state at temperature below and above \( T_L \), we make measurements by heating and cooling respectively at a heating/cooling rate of 0.75 K/min and observe exactly the same results.

We can calculate \( \langle x^2 \rangle \) from the Debye–Waller factor, \( S_H(Q, \omega = 0) = \exp[-Q^2 \langle x^2 \rangle] \), by a linear fitting of the logarithm of \( S_H(Q, \omega = 0) \) vs. \( Q^2 \) plot. \( S_H(Q, \omega = 0) \) can be easily calculated by taking the ratio of the temperature dependent elastic scattering intensity \( I_{el}(Q, T, \omega = 0) \) and its low temperature limit,

\[
S_H(Q, \omega = 0) = \frac{I_{el}(Q, T, \omega = 0)}{I_{el}(Q, T = 0, \omega = 0)}.
\]  

In Fig. 1 we show an example of the data taken from the D\(_2\)O and H\(_2\)O hydrated lysozyme samples, from which we can extract both MSDs from lysozyme and its hydration water. In order to show the synchronization of the temperature dependence of the two MSDs thus extracted, we multiply the \( \langle x^2 \rangle_{\text{Lysozyme}} \) by a factor 4.2, so both curves superpose onto each other. This figure nicely illustrates that the crossover temperatures for both protein and its hydration water defined by a sudden change of slope of MSD from a low temperature behavior to a high temperature behavior is coincident within the experimental errors [8].

It is well known that some bacteria can survive under extremely high pressure and low temperature in the deep ocean. The microorganisms living in the deepest ocean yet isolated and characterized were sampled at 11,000 m depth or 1100 bar in the deep-sea sediments of the Marianas trench, where the

![Fig. 1. Comparison of MSDs measured for the protein and its hydration water. Note that the MSD for hydration water is plotted using the scale on the left-hand side and MSD for the protein is using the scale on the right-hand side (the multiplication factor of the left and right scales is 4.2). MSD for the protein is taken from the elastic scan of D\(_2\)O hydrated sample. Note the crossover temperature of the hydration water (\( T_L \)) and the crossover temperature of the protein (\( T_C \)) agree with each other.](image-url)
Pacific oceanic lithosphere subducts into the Earth’s mantle [16]. How can proteins in the microorganisms still function under these extreme conditions? Besides the fact that high pressure denatures most of the dissolved proteins above 3000 bar, the behaviors of proteins under pressure below the denaturation limit (<2000 bar) both for structure and dynamics are relevant to the biological functions of proteins and are of great interest [3,30,46]. We show by measured MSD that the temperature dependence of the protein dynamics closely follows that of the hydration water under different pressures. In Fig. 2 we show the temperature dependence of MSD in lysozyme and its hydration water at different pressures up to 1600 bar.

Fig. 2. Reduced plot of pressure dependence of MSD of protein and its hydration water. It is to be noted in this figure that the crossover temperature of the protein and its hydration water is closely synchronized at a range of pressures below 2000 bar. The crossover temperature is seen to decrease as the pressure becomes higher.
4. Incoherent quasi-elastic scattering of hydration water in biopolymers

In this section, we discuss the qualitative and quantitative analyses of quasi-elastic spectra of different hydrated biopolymers, protein (lysozyme) [12], DNA [10] and RNA [15]. Our objective is first to show that the peak height and the peak width of the incoherent quasi-elastic spectrum is necessarily related to each other because the area under the \( S_H(Q, E) \) is in principle normalized to unity at each \( Q \) value. From this property we can already show by plotting the peak height of the \( S_H(Q, E) \) which is \( S_H(Q, E = 0) \) as a function of temperature that there is a dynamic crossover phenomenon without detailed analysis of the spectrum. In Fig. 3 we use hydrated DNA to illustrate the above mentioned fact that by plotting the peak height as a function of temperature we can already detect the presence of the dynamic crossover temperature at \( T = 225 \) K, in a qualitative way.

We now illustrate the detailed analyses of the quasi-elastic peak using RCM. Figure 4 shows a series of spectra of lysozyme hydration water taken at different temperatures at a pressure of 400 bar and \( Q = 0.56 \) and 1.11 Å\(^{-1}\). One can notice immediately that the peak height increases as temperature decreases, indicating the narrowing of the peak width, which is a qualitative measure of the \( \alpha \)-relaxation time.

In Fig. 5 we plot \( \log(\langle \tau_T \rangle) \) vs. \( 1/T \) in the same scale for 6 different pressures 1, 400, 800, 1200, 1500 and 1600 bars. Since the protein dynamics is strongly coupled to that of its hydration water, a short structural relaxation time of the hydration water enables the protein to maintain its flexibility and thus it is able to sample more conformational substates. For lower temperatures, \( \langle \tau_T \rangle \) obeys an Arrhenius...
Fig. 4. An example of the RCM analysis for lysozyme hydration water at 400 bar at two different temperatures \( T = 210 \, \text{K} \) (below \( T_L \)) and 240 K (above \( T_L \)), respectively. Note for the higher temperature case (\( T = 240 \, \text{K} \)), the quasi-elastic components are much broader. And we can see clearly that for higher \( T \), the peak height is much lower than the lower temperature case (\( T = 210 \, \text{K} \)).

behavior, which can be fitted by a straight line in the \( \log(\langle \tau_T \rangle) \) vs. \( 1/T \) plot; while for high temperatures, the behavior of \( \langle \tau_T \rangle \) switches over to obey a Vogel–Fulcher–Tammann (VFT) law,

\[
\log(\langle \tau_T \rangle) = \log \tau_0 + \frac{D T_0}{T - T_0},
\]

which are shown with dashed curves in Fig. 5.

A very distinct phenomenon in our experimental results is that \( \langle \tau_T \rangle \) shows a completely different behavior at pressure above 1600 bar. The crossover phenomenon disappears above this pressure and \( \langle \tau_T \rangle \) appears to be a smoothly bending over curve (concave downwards). Moreover, for the same temperature, \( \langle \tau_T \rangle \) is no longer decreasing as pressure increases to 1600 bar. This suggests that the hydration water must have crossed a “liquid–liquid critical point” around 200 K and 1600 bar [14,47].

Previous experiments on confined water in MCM-41-S porous silica material [23,40] have shown that an increased applied pressure will shift the FSC temperature to a lower value. We showed in Fig. 5 that this is also true for the interfacial water on the surfaces of protein, and that while a well-defined FSC phenomenon is observed for the applied pressure up to 1500 bar, when exceeding this pressure, the FSC phenomenon disappears. We thus identify a Widom line [56] in the \( T–P \) plane with an end point for
Fig. 5. Extracted relaxation time plotted in a log scale against $T_0/T$ (where $T_0$ is the VFT temperature) under six different pressures $P$ – ambient pressure, 400, 800, 1200, 1500 and 1600 bar. Note for the pressures up to 1500 bar, there is a well-defined crossover temperature; but at the pressure of 1600 bar, $\langle \tau_T \rangle$ appears to be a smooth curve, neither having the super-Arrhenius behavior at high temperature nor Arrhenius behavior at low temperature.

the case of protein hydration water which is nearly identical to that of the confined water in MCM-41-S. This implies the existence of liquid–liquid critical point in both the 1-D and 2-D confined water.

We have previously shown by a molecular dynamics (MD) simulation [56] that this super-Arrhenius to Arrhenius crossover is due to crossing of the Widom line in the one phase region. Upon the crossing of the Widom line, the local structure of water evolves from a predominately high density form (HDL, fragile liquid) to a predominately low density form (LDL, strong liquid) as the temperature crosses this characteristic temperature $T_L$ [43]. At the pressure of 1600 bar, $\langle \tau_T \rangle$ vs. $1/T$ plot appears to be a smooth curve, neither having the super-Arrhenius behavior at high temperature nor Arrhenius behavior at low temperature. We may attribute it [40] to the phase separated mixture of the HDL and LDL due to the crossing of the hypothetical liquid–liquid first-order transition line [47]. If these arguments are valid, then the disappearance of the FSC phenomena signals the crossing of the state point from the Widom line to the first-order liquid–liquid transition line. These two lines are separated by the liquid–liquid critical point if it exists.
In Fig. 6, we thus plot the trajectory of the crossover temperature $T_L$ as a function of $P$ (red circles). It is remarkable to see that this Widom line of the protein hydration water seems to coincide with the Widom line of the confined water in MCM-41-S found in our previous experiment [40].

5. Extension of the study of the crossover phenomenon to hydration water on DNA and RNA

At ambient pressure, we measured the average translational $\alpha$-relaxation time $\langle \tau_T \rangle$ of the hydration water molecules by QENS, and found a dynamic crossover in hydration water at a temperature $T_L = 225 \pm 5$ K in three biomolecules – lysozyme [12], B-DNA [10] and RNA [15]. Thus we have shown that $T_D \approx T_L$ at ambient pressure. As we discussed in Section 3, traditionally the dynamic transition temperature of a protein is discussed in terms of the turning point of MSD vs. temperature plot [48]. One can also use a similar plot to discuss the dynamic crossover temperature in protein hydration water, or more generally for hydration water in biopolymers [4,5]. The question naturally arises whether the dynamic crossover temperatures measured by these two different methods are identical or not. In Fig. 7, we present a plot of MSD and the average relaxation time of hydration water in DNA and RNA together in the same figure. It can be seen clearly from the graph that the crossover temperatures as determined from the elastic scan and the quasi-elastic scattering methods are identical within the experimental error bars.

Figure 8 shows the MSD of the three biopolymers and their hydration water in the form of scaled plots. From these plots we can see nicely the synchronization of $T_D$ (the glass transition temperature of the three biopolymers) and $T_L$ (the dynamic crossover temperatures of their hydration water).
6. Molecular dynamics simulations of hydrated protein powder

To better understand our experimental results on hydrated protein powder, we decided to perform MD simulations on the random powder model developed by Tarek and Tobias [54]. This realistic model can reproduce experimental data within the statistical error bars, including the measured mean-square displacements of the protein and its hydration water and the translational $\alpha$-relaxation time of the hydration water, $\langle \tau_T \rangle$, calculated from the self-intermediate scattering functions (SISF). The dynamic crossover we observed in experiments can thus be attributed solely to the long-time decay (in the range of 100 ps–50 ns) of the SISF of the hydrogen atoms attached rigidly to a typical water molecule [11,52,53], not to the long-range proton diffusion coupled to the motion of the so-called Bjerrum-type defects [52,53].
Fig. 8. Each panel shows the temperature dependence of the MSD of hydrogen atoms in both the biopolymer and its hydration water, respectively. It shows evidence that the crossover temperatures of the two systems, the biopolymer and its hydration water, are closely synchronized. (A) MSD of hydrated lysozyme; (B) MSD of the hydrated B-DNA; (C) MSD of the hydrated RNA. The arrow signs indicate the approximate positions of the crossover temperature in both the biopolymer ($T_D$) and its hydration water ($T_L$). Note that the scale on the left-hand side is for MSD of the hydration water and that on the right-hand side is for the biopolymer.
We put in a box two OPLS-AA [31] lysozyme molecules randomly oriented and 484 TIP4P-Ew water molecules \((h = 0.3\) for each protein, Fig. 11): after an energy minimization of 5000 steps with the Steepest Descent algorithm, we equilibrated the system in a NPT ensemble (isobaric–isothermal). We then performed 11 simulations at different temperatures (from 180 to 280 K, with 10 K of interval) with a parallel-compiled version of GROMACS [38], starting each simulation from the final configuration of the closest temperature. Each MD simulation length was 50 ns after the equilibration time. While there are only a few water molecules sandwiched between the two proteins, there are more water molecules around other parts of protein surface. But on the average, \(h = 0.3\) is supposed to be only one monolayer of water covering each protein.

### 6.1. Demonstration of the dynamic crossover phenomenon in protein hydration water

Figure 9a shows the calculated water hydrogen self-intermediate scattering functions (SISF) as a function of time at fixed \(Q\)-value (0.6 Å\(^{-1}\)), while the inset shows the ISF at \(T = 220\) K for different \(Q\)-values. The solid lines are the best fits to the ISF according to the RCM described above. The RCM fits of the ISF allow us to extract \(\langle \tau_T \rangle\) as a function of temperature as shown in Fig. 9b. The crossover feature is clearly visible looking at the decay of the ISF below and above \(T_L\). The crossover temperature is determined to be \(T_L = 221\) K, very close to the experimental value of 220 K [12].

### 6.2. Hydration dependence of the crossover phenomenon

But how does the relative amount of water that hydrates the protein powder affect its dynamics? To answer this question we increased the hydration level to \(h = 0.45\) (726 water molecules) and \(h = 0.6\) (968 water molecules). Details of the calculation are the same as in the \(h = 0.3\) case. When more and more water is added, the solvent molecules become more bulk-like. The hydration \(h = 0.3\) corresponds in fact to the average monolayer coverage of the protein surface. When this parameter is increased, water is forced to keep its distance from the macromolecule. Merzel and Smith [45] showed that the first hydration layer (\(\sim 2\) Å from the protein surface) is about 15\% more dense respect to bulk water, but that the normal density is recovered in the second hydration layer (\(\sim 4.5\) Å). Therefore, we expect that going from the \(h = 0.3\) to the \(h = 0.6\) case would shift water properties toward the bulk case.

In Fig. 10 we show the temperature dependence of the \(\alpha\) relaxation time for the various hydration levels, which is fitted with a VFT equation (at high temperatures) and an Arrhenius equation (at low temperatures). One can see that, as \(h\) increases, the dynamics becomes faster. This is in agreement with the view that water–water interactions are less strong than protein–water interactions, so that the bulk water limit corresponds to minimum relaxation times. Three results are evident from this picture: as \(h\) increases:

1. Average \(\alpha\)-relaxation time \(\langle \tau_T \rangle\) decreases.
2. The crossover temperature \(T_L\) decreases.
3. The activation energy \(E_A\) of the Arrhenius part decreases.

Both the first and the second point confirm the hypothesis that the bulk water case is a limit case. In fact, our published results show that a box of 512 TIP4P-Ew water molecules has \(T_L = 215\) K [57] \((T_L = 222, 218\) and \(216\) for \(h = 0.3, 0.45\) and \(0.6\), respectively).

We can then conclude that the protein–water interactions shift the temperature dependence of water dynamics to higher \(T\), but the essential characteristics and phenomena are qualitatively preserved.
Fig. 9. (a) Water proton incoherent self-intermediate scattering functions calculated at six different temperatures. (Inset) ISF at five different $Q$-values (from top to bottom, 0.4, 0.5, 0.6, 0.7 and 0.8 Å$^{-1}$). The choice of the $Q$ range was dictated by the low-limit value of $Q = 0.2$ Å$^{-1}$ imposed by the simulation box dimensions and the high-limit value of $Q = 1$ Å$^{-1}$, below which rotational motions can be neglected. The solid curves are fits to the relaxing cage model in a wide time range of 7 orders of magnitude, between 2 fs and 20 ns. (b) Temperature dependence of the average translational relaxation time, $\langle \tau_T \rangle$, calculated from MD simulation; $T_0$ is the ideal glass transition temperature. Numerical data are fitted with a Vogel–Fulcher–Tammann (VFT) law at high temperatures and with an Arrhenius law at low temperatures (solid lines) but with the same prefactor.

6.3. Single-particle dynamics of protein interior: $\alpha$ and $\beta$ relaxations

Protein dynamics can be divided into two main types according to its timescale [29]: (1) the slow timescale dynamics (µs–ms or the $\alpha$-relaxation) defines fluctuations between many conformational substates (CS) [26]. The CS are a group of large-amplitude collective motional states separated by energy barriers of $E_A \gg k_B T$; (2) the fast timescale dynamics (ps–ns or $\beta$-relaxation) defines local fluctuations between structurally similar states that are separated by energy barriers of $E_A < k_B T$. 
Fig. 10. The hydration level dependence of the crossover temperature $T_L$ for lysozyme hydration water. Note as the hydration level increases, the crossover temperature decreases and also the relaxation time $\langle \tau_T \rangle$ at $T_L(h)$ decreases [33].

By means of MD, we show how the time dependence of the $\beta$-relaxation in hydrated lysozyme powder follows a logarithmic decay [35], while the time dependence of the $\alpha$-relaxation decays exponentially in Fig. 11.

We are then able to fit the self-intermediate scattering functions of the center-of-mass of protein amino acids between 1 ps and 100 ns with the two-term equation

$$
\phi_q(t) \sim [f_q - H'_q \ln(t/\tau^{\beta}) + H''_q \ln^2(t/\tau^{\beta})] \exp(-t/\tau^{\alpha}_q),
$$

(18)

where the first term originates from the asymptotic solutions of mode coupling theory (MCT) for systems close to a high-order singularity [28], while the second term accounts for the slow relaxation in protein [6]. The factors $H'_q$ and $H''_q$ depend both on the wave vector $q$ and on the distance of the state point from the singularity (also known as separation parameter, $|x - x_c|$). The characteristic time $\tau^{\beta}$, instead, depends only on the separation parameter. $f_q$, on the other hand, depends on the state point only if the system is not close to the transition.

We show in Fig. 12 that the behavior of protein $\beta$-relaxation is in agreement with the predictions of MCT (Eq. (18)). The MCT scaling plots are also valid: $H'_q$ can be factorized in $H'_q \cong h(q)B'(x)$ and
Fig. 11. Complete time dependence of the self-intermediate scattering function of protein amino acid center-of-mass, at $T = 310$ K and $q = 4.0 \, \text{Å}^{-1}$. The inset shows the simulation box, highlighting the two lysozyme molecules of the protein powder model.

Fig. 12. $\beta$-relaxation in protein powder. Upper panels: protein SISF at $T = 280$ and 340 K. Continuous lines are best fits with Eq. (18), in the interval delimited by the dashed vertical lines (2 ps–6 ns). Lower panels: MCT scaling plots, $H'_q$ (left) and $(\phi_q - f_q)/H'_q$ (right).
Fig. 13. α-relaxation in protein powder. (A) Protein SISF at $T = 310$ K. Continuous lines are best fits with Eq. (18) in the interval 2 ps–100 ns. (B) Inverse of the α-relaxation time as a function of $q^2$. Dashed line is the best fit with $1/\tau^\alpha = D q^2$.

the ISF rescaled as $(\phi_q - f_q)/H_q'$ collapses to a form $-\ln(t/\tau^\beta)$, in the region where the first order approximation holds.

In Fig. 13, we show that at $T = 310$ K, the protein α-relaxation follows a diffusive behavior for $q > \pi/R$, with $R = \text{protein diameter} \sim 15$ Å: $\tau^\alpha$ follows the $\tau^\alpha = D q^2$ law, with a diffusion constant $D = 3.1 \times 10^{-10}$ cm$^2$/s, indicating a glassy liquid-like diffusive behavior for the constituents of the protein at physiological temperatures. Compared to a common glass-forming liquid like o-terphenyl, that also shows a logarithmic decay [6], this magnitude of the diffusion constant would correspond to $T \sim 290$ K, around the crossover temperature $T_c$ [44].

7. Detection of intra-protein phonon-like collective excitations in two globular proteins by high-resolution inelastic X-ray scattering (IXS) spectroscopy

In this section, we use inelastic X-ray scattering (IXS) to investigate the collective atomic motions in two globular proteins, lysozyme (LYZ) and bovine serum albumin (BSA). We show that the $Q$ dependent intra-protein collective vibrational frequencies exhibit a substantial softening above the transition temperature $T_D$ [41].

A measured IXS spectrum $I(Q, E)$ can be expressed as $I(Q, E) = S(Q, E) \otimes R(Q, E)$, where $S(Q, E)$ is the dynamic structure factor, $R(Q, E)$ the resolution function, and $\otimes$ means numerical convolution. In the present paper we shall use a simple damped harmonic-oscillator (DHO) model to analyze the normalized dynamic structure factor given by [37]

$$S(Q, E) = A \frac{E}{k_B T} (n(E) + 1)$$

$$= \left[ f_g(Q) \delta(E) + (1 - f_g(Q)) \frac{1}{\pi} \frac{\Gamma(Q) \Omega(Q)^2}{(E^2 - \Omega(Q)^2)^2 + (\Gamma(Q) E)^2} \right],$$

(19)

where $\Omega(Q)$ is the phonon energy at wave vector transfer $Q$, and $\Gamma(Q)$ is the phonon damping, $n(E) + 1$ is the detailed balance factor, $n(E) = (e^{E/(k_B T)} - 1)^{-1}$ the Bose factor, $A$ a normalization constant, and
Fig. 14. Lysozyme structure factor $S(q)$ calculated from the crystallographic structure (red) and from the MD trajectories (blue). Arrows point at the main peaks in the $S(q)$ spectrum: (A) 7 nm$^{-1}$, (B) 14 nm$^{-1}$, (C) 30 nm$^{-1}$, (D) 55 nm$^{-1}$. Only the MD trajectories can reproduce the experimentally observed shoulder peak A. Peak B corresponds to typical distances of the protein secondary structure, peak C to some medium range order between residues and peak D to the protein covalent bonds. The meaning of the shoulder peak A is still uncertain. (The colors are visible in the online version of the article.)

$f_g(Q)$ represents the fraction of the elastic peak. At the low $Q$ region $f_g(Q) > 0.95$ and at the high $Q$ region $f_g(Q) \approx 0.6–0.8$.

To verify the experimental $S(Q)$ for lysozyme, we calculated it from (1) the crystallographic structure (1AKL.pdb), and (2) an MD simulation at 250 K using a hydrated lysozyme powder model [36,54]. Only $S(Q)$ calculated from MD shows the 4 expected peaks, centered at $Q = 7, 14, 30$ and 55 nm$^{-1}$ [20,32], while the shoulder peak A cannot be seen reproduced from the crystallographic structure. Peak B, which should be the measured peak of the structure factor, arises from the order of the secondary structure ($\beta$-sheets average distance and $\alpha$-helix repeat and width), as displayed in Fig. 14. Peak C is ascribed to the intermediate range order between amino acidic residues, while peak D is attributed to the covalent bonds between heavy atoms (like C–C, C–N, etc.). Finally, the sharp increase at low $Q$ indicates that, on average, a protein can be regarded as a fluctuating continuum.

We show in Fig. 15 the analysis of measured IXS spectra $S(Q, E)$ of BSA and LYZ by the DHO model at three different temperatures $T = 170, 220$ and 250 K. From the results of the model fit at $Q = 4.5, 5.6$ and 6.7 nm$^{-1}$ (low $Q$ range) at 220 K, we can see that the two Brillouin peaks are clearly identified and the positions of the peaks are located at around 11 meV for all the $Q$ values. However, for $Q = 24.6, 27.9$ and 31.2 nm$^{-1}$ (high $Q$ range), clear $Q$-dependence of the phonon excitation energy is observed. At a given $Q$ value, when the temperature exceeds $T_D$, we see a substantial decrease in the phonon energy. This softening of the phonons in this $Q$ range suggests that the slowing-down of motions involving the intermediate-range order in a protein may be the origin of the restoration of its biological activities.

We plotted the phonon dispersion relations in the full $Q$ range for LYZ and BSA in Fig. 16. Two different phonon dispersion behaviors below and above $Q \approx 15$ nm$^{-1}$ (the position of the major peak of the measured structure factor $S(Q)$ are clearly shown in both figures. In the low $Q$ range of the BSA dispersion curves, there is no clear temperature dependence of the excitation energies. In the high $Q$ range, excitation energies have similar dispersion curves as that had been observed in other bio-systems [39]. Since the lower $Q$ phonons in the $Q$ range of 4–10 nm$^{-1}$ correspond to the length scale between 7 and
Fig. 15. Model fitting of the measured IXS spectra of BSA. The blue circles, cyan solid line, green dashed line and pink solid line represent respectively the measured data, DHO model fitted curve, resolution function and Brillouin component of the DHO model. The arrow signs (red) show the Stokes and the anti-Stokes components of the phonon-like excitation energy at this $Q$. The upper panel shows the fitted results at $Q = 4.5, 5.6, 6.7 \text{ nm}^{-1}$ (from left to right) at 220 K. No dispersion was observed in this $Q$ range. The lower panel shows the fitted results at $Q = 24.6, 27.9, 31.2 \text{ nm}^{-1}$ (from left to right) at $T = 170, 220, 250 \text{ K}$ (from up to down). At each temperature, clear $Q$-dependence of the phonon excitation energy is shown in the figure. At a given $Q$ value, the phonon energy decreases substantially when temperature exceeds $T_D$. (The colors are visible in the online version of the article.)

15 Å, which reflects the inter-protein motions, we can conclude that the inter-protein motions are temperature independent and do not induce the temperature dependent biological functions of the proteins. On the other hand, the high $Q$ phonons in the $Q$ range of 20–30 nm$^{-1}$, which correspond to the length scale of about 2–3 Å, are close to the length scale of the typical distance of the secondary structure of the proteins (4–5 Å) and reflect the intra-protein collective atomic motions. We can clearly observe the
Fig. 16. (a) Structure factor $S(Q)$ of LYZ and BSA in a relative scale. We can see two major peaks at $Q \approx 5$ and 15 nm$^{-1}$. (b) Dispersion of the intra-protein phonon-like energy excitations of LYZ and BSA as a function of $Q$, below and above the second major peak in the structure factor. The square (blue), triangle (orange) and circle (yellow) symbols indicate the phonon energies at 170, 220 and 250 K. (c) The phonon damping as a function of temperature. (d) Fractional area of the Brillouin peak vs. $Q$. The factor $(1 - f_g)$ in Eq. (2) is found to increase proportionally to $Q^2$, $1 - f_g(Q) = \alpha(T)Q^2$, where $\alpha(T)$ is an increasing function of $T$. (The colors are visible in the online version of the article.)

quantitative degree of the softening of phonons for both proteins as temperature exceeds $T_D$ in Fig. 16. This fact strongly suggests that these temperature dependent intra-protein motions are intimately related to the biological activities of proteins, which are also temperature dependent and strongly decrease below $T_D$.

The presence of the structure factor peak at $Q \approx 15$ nm$^{-1}$ leads to a large damping of the phonons, as we can see in panel (c) of Fig. 16, so we cannot detect the phonons in this region. The dotted lines in panel (b) of Fig. 16, which connect the dispersion curves, are guide lines drawn based on the previous studies of phonons in other bio-systems [39]. The phonon damping $\Gamma(Q)$ does not show a temperature dependence and is nearly constant at high $Q$ region. This fact shows that the temperature dependence of $\Omega$ is reliable.

We calculated the factor $1 - f_g(Q)$ in Eq. (19) and fit it by a power law. Since the factor $f_g(Q)$ is the fraction of the elastic component, it should be proportional to the Debye–Waller factor $\exp(-Q^2\langle x^2 \rangle)$. So the low $Q$ expansion of $f_g(Q)$ should be proportional to $Q^2$. Our data show this $Q^2$ dependence in Fig. 16. The factor $1 - f_g(Q)$ represents the fractional spectral intensity of the phonon-like excitation in the total intensity. At temperature higher than $T_D$, $1 - f_g(Q)$ is larger, which indicates that there are more phonon population due to the onset of the conformational flexibility at $T_D$.

In conclusion, we believe that these phonon-like modes are the result of the collective vibrational motions of the atoms in the $\alpha$-helices and $\beta$-sheets. Below $T_D$ the vibrational frequency is too high and the population of the modes is too low to be able to facilitate the biological function. That is the reason why proteins are not good enzymes below the dynamic transition temperature.
8. Summary

In this last section, we shall recapitulate the important new findings of our neutron and X-ray scattering experiments that we described in the above sections. They can be summarized in the following seven points:

(1) We can detect the dynamic crossover temperature $T_L$ by either the MSD vs. $T$ plot or by the Arrhenius plot of the $\alpha$-relaxation time vs. $1/T$. The two methods give the same crossover temperature.

(2) The pressure dependence of an Arrhenius plot of the relaxation time in lysozyme hydration water gives a plausible evidence of the existence of the second critical point in supercooled confined water.

(3) The dynamic transition temperature is the same in all biopolymers because it is synchronized with the dynamic crossover temperature $T_L$ in their hydration water suggesting a “slaving” phenomenon.

(4) MD simulation of a realistic hydrated powder model of lysozyme reproduces the dynamic crossover phenomenon in the protein hydration water.

(5) As the protein hydration level increases from $h = 0.3$ to 0.45 to 0.6, the crossover temperature decreases from 222 to 218 to 216 K. The crossover temperature gradually approaches the bulk water (TIP4P-Ew) limit of 215 K.

(6) The $\beta$-relaxation of protein dynamics shows an exotic logarithmic decay over 3 decades of time, from 2 ps to 6 ns, predicted by mode coupling theory for systems close to a higher-order singularity.

(7) The protein phonon energy shows a marked softening at $T > T_D$. The increase of biological activities above $T_D$ has positive correlation with activation of slower and large amplitude collective motions of a protein.

Acknowledgements

The research at MIT is supported by DOE Grant No. DE-FG02-90ER45429. This work utilized facilities supported in part by the National Science Foundation under Agreement No. DMR-0086210. We appreciate technical support during experiments from V. Garcia-Sakai, T. Jenkins, M. Tyagi, S. Poulton and J. Leao of the NIST Center for Neutron Research. Emiliano Fratini and Piero Baglioni acknowledge financial support from CSGI and MIUR. We benefited from affiliation with EU funded Marie-Curie Research and Training Network on Arrested Matter.

References

S.-H. Chen et al. / Dynamics of a globular protein and its hydration water 23


