Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions

Wolfgang Doster[1]

[1]*Technical University Munich, Physics-Department, D-85748 Garching, Germany and Jülich Centre for Neutron Science, FRM II, Lichtenberg-str. 1, D-85747 Garching, Germany.*

Abstract

Dynamic neutron scattering is a technique to record time resolved the molecular displacements of hydrogen atoms in biomolecules. With spectrometers, measuring the energy exchange, $\hbar\omega$, and momentum exchange, $\hbar Q$, of neutrons with the hydrogen atoms of the sample, one determines the scattering function S ($\mathbf{Q}$, ω). The central physical quantity of interest, which can be derived from the scattering function by a double Fourier transform, is the displacement distribution $G_s(r, t)$. In this article we explain by three theoretical methods, how to determine the displacement distribution of hydrogen atoms in proteins from experimental data: (1) a moment expansion, (2) a numerical transform of the scattering function and (3) an analytical model of two principal dynamic components. The expansion yields a temperature dependent second moment of the displacement distribution and a Gauss-deviation, suggesting asymmetric motions. The numerical transform reveals a temperature dependent bimodal distribution function of small scale diffusive displacements and of rotational jumps on a larger scale. The principal components of the analytical model are thus rotational transitions of side chains and local diffusion of residues in a quasi-harmonic potential. The local diffusion is modelled by an over-damped Brownian oscillator derived from the Smoluchowski equation. The rotational transitions are treated by a three equivalent site model of the methyl group. Comparison of the model predictions with experimental data in the time domain show a remarkable agreement within a wide range of time scales, momentum exchange and temperature. The microscopic dynamics are compared with functional kinetic studies of ligand binding to myoglobin on a longer time scale. Dynamical heterogeneity is reduced to a bimodal distribution of molecular processes. Multiple scattering calculations provide an alternative explanation to heterogeneous quasi-elastic neutron scattering spectra and energy landscapes.

Keywords: Protein dynamical heterogeneity; Neutron scattering; Myoglobin; Methyl groups; Hydration water; Energy landscape, Ornstein-Uhlenbeck process.

Introduction

Proteins are hetero-polymers exhibiting well-defined, compact structures in contrast to artificial homo-polymers. Molecular dynamics in proteins thus differs from disordered open polymer structures. Depending on the sequence of amino acids one expects a heterogeneous system of local sites, which gives rise to considerable dynamical heterogeneity. Site sensitive techniques like neutron scattering and molecular dynamic simulations should be able to characterize the properties of fast motions in proteins. Is heterogeneity a dominant feature of protein dynamics or is it possible to focus on a small set of essential motions? The biological activity of enzymes is generally characterized by few well defined rate coefficients, while heterogeneity plays a minor role. This raises the question, whether the catalytic power of proteins requires the assistance of molecular dynamics. Arieh Warshel, co-Nobel prize winner of 2013 in Chemistry, argues that "flexibility is interfering with rate acceleration, catalysis requires rigid stereo-chemical structures"[1]. However conformational changes play an important role in regulation and specific catalytic steps. This open question can be addressed by combining dynamic studies with functional experiments. The idea of functional heterogeneity came from low temperature kinetic studies of ligand binding to the oxygen storage protein myoglobin. Flash photolysis experiments in various cryo-solvents revealed non-exponential kinetic steps with distributed rate coefficients [2]. Wide temperature range neutron scattering experiments and simulations of myoglobin were performed since 1988 to identify molecular motions on a pico-second time scale [4-7]. Instead of protein solutions, thermally stable hydrated powders were investigated [3]. This method avoids crystallization and reduces solvent scattering, moreover, the lack of solvent inhibits blurring global translational and rotational motions, while preserving the biological function. Two kinds of molecular motions, rotational transitions of side chains and small scale diffusive displacements were identified by low temperature transitions and the analysis of the non-Gaussian elastic scattering function [4]. An alternative interpretation of "dynamical heterogeneity" in terms of continuous displacement distributions was proposed simultaneously by Smith [6]. The two interpretations persist independently up to date, which is one of the main topics in this work. The new approach with myoglobin was soon applied to numerous other proteins. But instead of

Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions

Copyright:
© 2018 Doster W.

expanding the original spectral analysis of protein dynamics [4], the main effort until today is devoted to elastic scattering, which reflects mainly the 'rigid' part of the molecule [8-10]. The elastic spectrum is the most intense and thus most easily measurable component, allowing a simple analysis in analogy to static low angle scattering. Even more recent publications analyze elastic scattering curves complemented by simulations, which provide the missing spectral information [11]. Elastic scans were performed with numerous protein hydrated powders. Absolute Mean Square Displacement (MSD) can be determined from the slope of the scattering function versus momentum exchange (Figure 2) at low values as discussed below. Figure 1 shows

Figure 1: Second moment of hydrogen atom displacement distribution in D2O-hydrated myoglobin (green diamonds) and lysozyme (blue squares), instrument IN13, τ_{res} = 140 ps [4, 8, 9], heme iron of myoglobin, crystal: red circles, in sucrose solution: red triangles by Mössbauer spectroscopy [12, 13], short dashed line: vibrational displacements calculated from density of states [5]. Full line: model calculation, type II. The onset temperatures of type I and II motions are indicated. Dashed arrow: onset of vibrational displacements (Bose factor), arrow: unresolved transition, magenta circles: δ_B^2 of the TR model.

Figure 2: Normalized elastic scattering profiles of D2O-hydrated myoglobin versus Q^2 and temperature, backscattering spectrometer IN13 [4, 8, 7]. The full lines were adjusted to the TR model of equation (19-21)

typical temperature scans of the hydrogen atom mean square displacements comparing two proteins with different structures, hydrated lysozyme (β-sheets) and myoglobin (α-helices) [4, 8, 9]. There is little difference in MSD between the two proteins at low temperature, slight deviations emerge however above 240 K. The low temperature plateau is terminated at $T_L \cong 50$ K by the condition $\hbar\omega_L \cong K_B T_L$, the Bose factor of low frequency vibrations, crossing over to a linear regime of harmonic excitations [5]. A first deviation from linearity occurs at about 180 K, which is denoted as type I transition identical for both proteins. A second transition emerges at 240 K, hardly visible as a deviation from the dashed line, which we call type II transition. The type II displacements in the α-helical protein seem to be slightly larger than in the β-sheet protein. But the discrepancy could equally well reflect a different degree of hydration. The figure is meant to illustrate that the popular method based on the interpretation of displacement scans alone is difficult. The two transition temperatures were already identified in 1989 [4]. The assignment was only possible with the assistance of an analytical model of the complete elastic scattering function as shown in Figure 2. A combination of two molecular processes had to be assumed, rotational transitions (type I) and local translational diffusion (type II). The distinction became convincing, when it was realized that type II requires the presence of water, while type I is nearly independent of the protein environment [4, 8]. This model and its extensions will be discussed below.
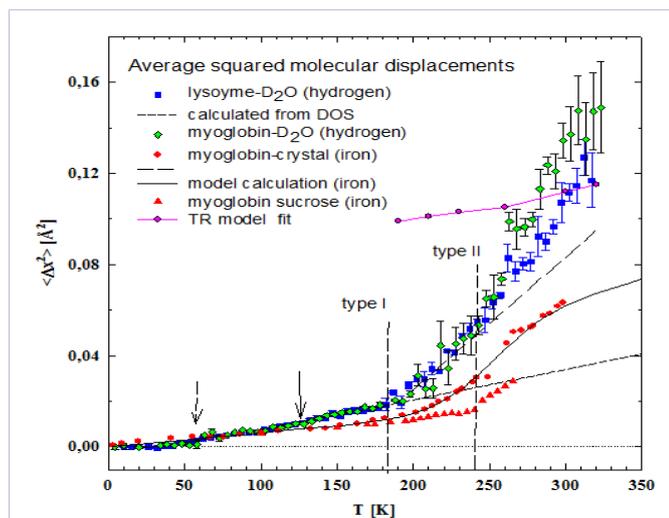
Elastic scans are based on a fixed energy window method, which implies that the resulting displacements carry the time tag of the elastic instrumental resolution. For the back-scattering data of IN13 above, the relevant resolution time is $\tau_{res} \cong$ 140 ps. Fixed energy window scans have been used with other methods. For myoglobin, the γ-resonance spectroscopy of the heme iron (Mössbauer) is often compared with neutron scattering, which is therefore displayed in Figure 1 [12, 13]. Although the γ- MSD scans resemble those obtained with neutrons, there are three main differences: (1) Mössbauer spectroscopy records the local displacements of the heme iron, while neutrons record the hydrogen atoms, which are distributed evenly across the protein, (2) the spatial scale, about 0.1 Å, is much smaller than for neutron scattering, which is in the range of a few Å. The γ-resonance thus allows one to measure protein global diffusion on a very small scale at very high resolution. From this method we know, that global diffusion is arrested in protein crystals and hydrated powders below 0.45 g /g [14]. (3) the instrumental resolution τ_{res}^{γ} = 142 ns is by a factor of 1000 higher than with neutron backscattering (IN13), $\tau_{res}^{n} \cong$ 140 ps. Protein crystals have about the same water content as fully hydrated powders, suggesting similar dynamical properties. In contrast to neutron scattering, there is only a single transition with Mössbauer at $T^{\gamma} \cong$ 205 K. To investigate the effect of the solvent on the transition, we have performed γ-resonance experiments with myoglobin in a highly viscous sucrose solution [13]. As a result, the onset temperature is shifted to 240 K. This experiment shows that the heme interacts strongly with the surrounding solvent and is thus associated with type II motions.
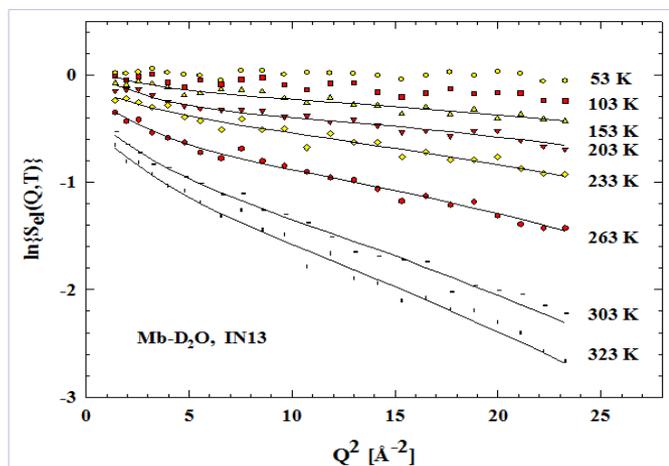
Citation: Doster W (2018) Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions. Int J Mol Theor Phy. 2(1):1-14.

Page 2 of 14

Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions

Copyright:
© 2018 Doster W.

The difference by 35 degrees in the onset temperatures of γ/n in the hydrated/crystal case is a consequence of the different resolution times of the two methods. With this concept the temperature dependent iron displacement could be reproduced, with the 142 ns resolution time and the water relaxation time as input [10]. Frauenfelder et al. by contrast consider neutron and γ-resonance displacements as being essentially identical, which ignores the vastly different resolution times of the two methods [15]. A more detailed discussion of these discrepancies is given in reference [16]. Blinded by the narrow view of elastic scattering, one easily is led to incorrect conclusions. Similar experiments with different solvents have been performed with neutron scattering: Myoglobin in D-exchanged sucrose did not show a single transition and was classified as rigid [17]. By contrast, with myoglobin embedded in per-deuterated glucose glass, the type I transition could be fully observed as with dehydrated myoglobin [8, 9]. Sucrose contains non-exchangeable protons, which dominate the elastic intensity of the glassy myoglobin-solvent system, overwhelming the protein-internal motions.

After 30 years of effort, combining experiments and molecular dynamic simulations, all major questions should have been resolved. Instead, up to now, there is no accepted molecular model available in the literature, combining elastic and inelastic neutron scattering spectra of proteins. At present, this field is controlled by phenomenological models, which play a major role in the standard list of citations of neutron scattering work:

The Bicout-Zaccai model derives protein dynamics exclusively from elastic displacements and their temperature dependence as in Figure 1 [17, 18]. Reducing protein dynamics to harmonic oscillators, the central goal is to determine a global protein force constant from the displacement temperature slope. A change in the slope at a particular temperature is interpreted as indication of a softening transition of the general force constant [9]. Energy landscape models picture protein dynamics as single particle random walks across a complex surface. Frauenfelder has launched recently a massive attack on well-established neutron scattering theory, which is supposed to fail for complex systems [15, 19]. In this view, quasi-elastic neutron scattering spectra have to be analyzed by a heterogeneous superposition of resolution-limited narrow lines, dictated by the energy landscape. We show below that some of these conclusions are based on incorrect assumptions, ignoring for instance the effects of multiple scattering.

More detailed information can be derived from molecular dynamic simulations, by calculating the correlation function for each site. Protein structures are lacking translational symmetry, which suggests a broad distribution of molecular motions. In the limit of small momentum exchange, the scattering function may be expanded by a sum of Gaussian local site distributions. This is the 'dynamical heterogeneity' concept of protein sites [6, 20, 26]. Several distributions, Gamma, exponential, bimodal, derived by simulations, were found to be compatible with the data, the experimental curves however are site-averaged. Experimentally, we have deduced a bimodal distribution (8), which suggests a discrete set of processes.

In this article we follow the concept of a pluralistic methodology: It is shown, that the dynamic displacement distribution can determined theoretically by applying three methods: (1) a moment expansion of the distribution function, (2) a numerical transform of the elastic scattering functions and (3) an analytical model assuming two principal components of protein dynamics. The methods are explained by using a novel analysis of previously published data.

Methods

Neutron scattering requires big machines, reactors or accelerators, which are located in a dozen research facilities scattered over the world. The access is free to anybody, who can come up with good proposals for experiments. One of the goals of this article is to explain the method, trying to generate interest and understanding for its potential application in a biological context. With the advent of reactors and more recently of accelerators, neutrons have become a standard tool of investigating density fluctuations in condensed matter. The de Broglie wavelength of thermal neutrons, a few Å, is close to interatomic distances in liquids and solids. Thus interference effects occur, which yield information on the molecular structure of the scattering system. Second, since the neutron is uncharged, there is no Coulomb barrier to overcome. The scattering of nuclear forces can be treated by a delta-function Fermi pseudopotential [21-23]. For certain nuclides the scattering cross section is large, the hydrogen atom has a nearly ten times larger cross section than the other atoms of organic matter and the deuteron. In proteins typically 84% of the scattering fraction is due to the hydrogen atoms [24, 25]. Because of the low cross section of the deuteron, one can diminish the contribution of water by using D_2O to about 4% at full hydration, 0.38g water per g protein [25]. Thirdly, the energy of thermal neutrons compares favorably with thermal excitations in biomolecules, which can be studied by energy exchange experiments on pico- to nanosecond time scale.

Below we focus on experiments performed mainly with the back-scattering spectrometer IN13 at the ILL in Grenoble [23]. Back-scattering by Bragg reflection maximizes the instrumental resolution (FWHM =7μeV), corresponding to a resolution time window, τ_{res}= ħ/FWHM ≅140 ps. Neutrons are selected from the primary beam by Bragg reflection at λ = 2.23 Å. The scattered beam hits the same type of analyzer crystal, whose lattice constants are modified by heating. A temperature scan of the mismatched crystal produces the spectral information from the scattered beam. The most important property of IN13 is the exceptionally large Q^2 range up to 25 Å^{-2} combined with a high energy resolution due to the short primary wavelength. This combination allows in particular conclusive tests of spatial displacement distributions of molecular motions. A comparable back-scattering spectrometer, HFBS at NIST, has a Q^2 range of only 4 Å^{-2}.

Citation: Doster W (2018) Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions. Int J Mol Theor Phy. 2(1):1-14.

Page 3 of 14

Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions

Copyright:
© 2018 Doster W.

Theoretical background [10, 21-23, 26]

The basic theoretical quantity of incoherent neutron scattering is the single particle density-density correlation function, $G_s(\vec{r},t)$ of displacements $\vec{r}(t)$. Ignoring quantum effects, it may be interpreted as the conditional probability, averaged over all initial positions $\vec{r}_0$, that a particle has moved across a distance $\vec{r}$. during the time interval 't'.

$$G_s(\vec{r},t) = \int d^3 r_0 \, p_o(\vec{r}_0) \, p(\vec{r}_0, \vec{r}_0 + \vec{r}, t) \qquad (1)$$

where $p\left(\vec{r}_0, t_0 / \vec{r}_0 + \vec{r}, t_0 + t\right)$ denotes the conditional probability for this transition.

The Fourier transform of equation-(1) defines the self-intermediate scattering function $\Phi_s(\vec{Q}, t)$ versus momentum exchange $\vec{Q}$, which is measured in a scattering experiment:

$$\Phi_s(\vec{Q}, t) = \int d^3 \vec{r} e^{-i\vec{Q}\vec{r}} G_s(\vec{r}, t) \qquad (2)$$

The scattering function of equation-(2) can be expanded in powers of $\vec{Q}.\vec{r}$ and the spatial moments of $G(\vec{r},t)$ resulting in [8, 21-23, 26]:

$$\Phi_s(\vec{Q},t) = 1 - \frac{1}{2}\int d^3\vec{r}\left(\hat{Q}\vec{r}\right)^2 G(\vec{r},t) + \frac{1}{24}\int d^3\vec{r}\left(\hat{Q}\vec{r}\right)^4 G(\vec{r},t) - \ldots \quad (3)$$

$$= 1 - \frac{1}{2}Q^2 \cdot \left\langle\left(\hat{Q}\vec{r}\right)^2(t)\right\rangle + \frac{1}{24}Q^4 \cdot \left\langle\left(\hat{Q}\vec{r}\right)^4(t)\right\rangle - \ldots$$

For isotropic systems the average covers all orientations leading to:

$$\Phi_s(Q,t) \cong 1 - \frac{1}{2\cdot 3}Q^2 \left\langle r^2(t)\right\rangle + \frac{1}{24\cdot 5}Q^4 \left\langle r^4(t)\right\rangle - \ldots \qquad (4)$$

For a Gaussian G_s(r, t) one has $\left\langle r^4(t)\right\rangle = \frac{5}{3}\left\langle r^2(t)\right\rangle$ which of course gives:

$$\Phi_s(Q,t) = \exp\left\{-\frac{Q^2}{6}\left\langle r^2(t)\right\rangle\right\} \qquad (5)$$

Often it is useful to split the intermediate scattering function into a time-dependent component F (Q, t) with F (Q, t → ∞) = 0 and a time-independent part, the EISF (Q) according to:

$$\Phi_s(Q, t) = \left[1 - EISF(Q)\right] F(Q, t) + EISF(Q) \qquad (6)$$

The EISF (Q) is also known as the "elastic incoherent structure factor" of the spectrum, since, in the frequency domain, it gives rise to elastic scattering. It is defined as the long time limit of the averaged scattering function according to equation- 2:

$$EISF(Q) = \Phi(Q, t \to \infty) = <\int d^3\vec{r} e^{-i\vec{Q}\vec{r}} G_s(\vec{r}, t \to \infty) >_{powder} \quad (7)$$

From the EISF (Q) one can reconstruct the powder averaged displacement distribution of proteins, based on experiments, which is the focus of this work. The Gaussian distribution of equation (5) yields at long times:

$$EISF_G(Q) = \exp(-Q^2\left\langle\Delta r_G^2\right\rangle / 3) \qquad (7a)$$

With $\left\langle(\Delta r)_G^2\right\rangle = \frac{1}{2}\left\langle r^2(t\to\infty)\right\rangle$. On a semi-log scale versus Q^2 one thus expects for the EISFG (Q) straight lines with slopes yielding the Gaussian MSD = <Δx²> = <Δr$_G$²>/3.

Experimentally, one is limited by the finite resolution time τ_{res} of the spectrometer. Thus instead of equation-7, one needs to consider a resolution corrected function, the REISF (Q), which accounts for the displacements at a particular time τ_{res} [10]:

$$REISF(Q, \tau_{res}) = <\int d^3\vec{r} e^{-i\vec{Q}\vec{r}} G_s(\vec{r}, t \approx \tau_{res}) >_{powder} \qquad (8)$$

Elastic scattering experiments are always performed in the frequency domain, which involves the Fourier transform of equation (2), given by the incoherent dynamic structure factor SR (Q, ω), convoluted with the instrumental resolution function R (ω, Δω = 1/τ_{res}):

$$S_R(Q,\omega) = REISF(Q,\tau_{res})\cdot R(\omega,\Delta\omega) + QREISF(Q)\cdot S_{qel}(Q,\omega) * R(\omega,\Delta\omega) \quad (9)$$

S$_{qel}$ (Q, ω) denotes the quasi-elastic spectrum, the Fourier transform of F (Q, t) in equation (9), the star is a short notion of the convolution integral. In the following we analyze the normalized elastic fraction REISF (Q, τ_{res}) with respect to a particular observation time, τ_{res}. A low temperature spectrum (10 K), whose shape is defined by the resolution function R (ω,Δω), serves to normalize the data obtained at the higher temperatures:

$$S_R^0(Q,0,\Delta\omega) = \frac{1}{\pi}\int_0^\infty dt \, R(t,\tau_{res} = 1/\Delta\omega) \qquad (10)$$

The normalized elastic intensity is then determined from the experiment:

$$S_{el}^N(Q,\tau_{res}) = S_R(Q,0,\Delta\omega) / S_R^0(Q,0,\Delta\omega) \qquad (11)$$

The normalized elastic intensity in frequency space can be interpreted as an approximate intermediate scattering function in the time domain at the resolution time t ≅ τ_{res}, which is one basic result of "elastic resolution spectroscopy" [10, 27]:

$$S_{el}^N(Q,\tau_{res}) \approx \Phi_s(Q, t = \tau_{res}, \tau_c(T)) = REISF(Q,\tau_{res},\tau_c) \qquad (12)$$

τ_c(T) denotes the correlation time of a molecular process. The REISF (Q, τ_{res}, τ_c) differs from S$^N_{el}$ (Q, τ_{res}, τ_c), since the latter includes a quasi-elastic intensity at ω = 0, while the former only records the true elastic intensity. This difference, usually ignored in elastic scattering studies, will be small only for small ratios τ_{res}/τ_c.

Results

Elastic backscattering experiments with D$_2$O-hydrated myoglobin [4, 8]

The elastic intensity was determined with a fixed energy window method at a resolution of 7 μeV, full width at half maximum, applying the back-scattering spectrometer IN13 at

Citation: Doster W (2018) Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions. Int J Mol Theor Phy. 2(1):1-14.

Page 4 of 14

Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions

Copyright:
© 2018 Doster W.

the incident neutron wavelength of 2,23 Å: The primary beam crystal and the analyzer crystal are kept at the same temperature. The sample, 350 mg of D_2O-hydrated horse myoglobin (SIGMA Chemical, h = 0.38 g/g), was continuously heated from 4 K to 330 K over a period of 24 h. The samples were measured in a thin walled, vacuum tight aluminum cell, diameter 50 mm and interspacing of 1 mm with the sample at 135° to the incident beam. With 350 mg of hydrated sample the neutron transmission was close to a tolerable 90%. Raw data were corrected for detector response and cell scattering. The elastic scattering curves, normalized at the lowest temperature according to equation-11, are displayed in figure 2. With normalization at 10 K, one removes the zero point vibrations, which are not relevant in the present context [4, 8]. However the second normalization step at Q = 0 Å^{-1} at each temperature as in references 4 and 19 was omitted. The data of figure 2 exhibit three main features:

(1) The elastic intensity decreases with increasing Q and temperature. The loss in intensity is balanced by an increasing quasi-elastic intensity and line broadening due to resolved molecular motions (equation 9).

(2) The elastic scattering curves on a logarithmic scale deviate from straight lines except at the lowest temperatures, revealing that the protein hydrogen displacements are not Gaussian distributed.

(3) Extrapolated to zero Q, the elastic intensity tends to decrease with increasing temperature, which is a clear sign of multiple scattering. These features are identical for nearly all hydrated and lyophilized proteins, irrespective of their structure.

The data in figure 2 vary smoothly with Q, the deviations from a straight line are subtle. Most workers are interested only in MSD values versus the temperature, which requires evaluating the slope at low Q [17, 18]. This can be done in many ways. Thus there is little consensus in the literature, how to properly analyze elastic scattering profiles. Since there is more information available than just MSD values, we approach the data of figure 2 from three different points of view.

The moment (Plazcek) expansion of elastic scattering data

The most general method is to perform a moment expansion without introducing specific assumptions about the nature of the molecular processes according equations 3&4. The moment expansion was previously applied to neutron scattering experiments in ref. 8, 28 and in the context of protein simulations and dynamical heterogeneity in reference 26.

The data of figure 2 are reasonably well approximated by a second order polynomial in Q^2 ecording to equation (4):

$$S_{el}^N(Q,T) \approx A(T) - B(T) \cdot Q^2 + C(T) \cdot Q^4 \quad (13)$$

While B (T) and C (T) involve the second and forth order spatial moment, the zero order term A (T) emerges in all experiments with a magnitude depending on the sample transmission. It thus reflects second and higher order scattering contributions, which are approximately angle independent [5, 29]. Figure 3 displays an example of how equation-(13) is adjusted to the data on a logarithmic scale applying different fitting intervals. Fits on a linear scale are not stable and increase at higher Q.

Figure 4 shows the resulting second moment (MSD) and the so-called Gauss deviation G (t), which is constructed from the second and the forth moment and can depend on time[8]:
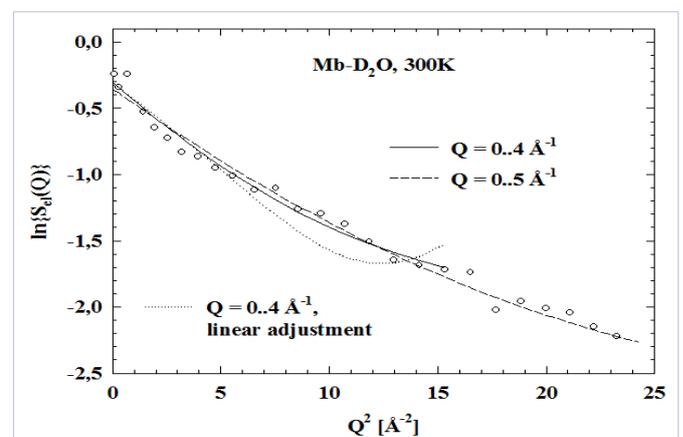
Figure 3: Polynomial fit on a logarithmic scale according to equation 13 with different fitting intervals. Adjustments on a linear scale become unstable at higher Q. The very large Q range of IN13 is crucial to the success of this method.
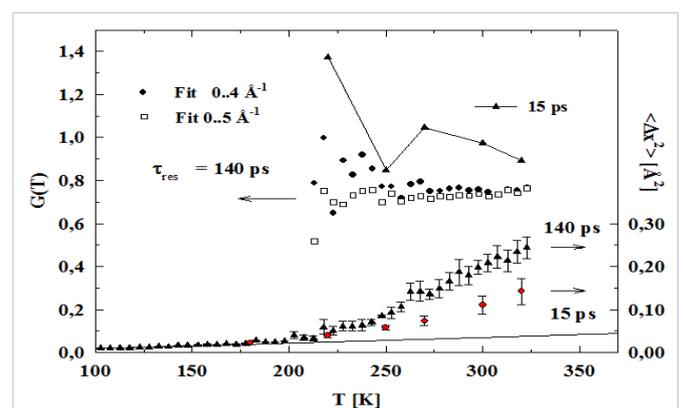
Figure 4: Mean square displacements and Gauss deviation G (T) versus the temperature forIN13, triangles (τ_{res} = 140 ps) and IN6 (red circles, τ_{res} = 15 ps) from fits of the data in figure 2 to equation13.

$$G(t) = \frac{3\langle r^4(t)\rangle}{10\langle r^2(t)\rangle} \quad (14)$$

Citation: Doster W (2018) Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions. Int J Mol Theor Phy. 2(1):1-14.

Page 5 of 14

Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions

Copyright:
© 2018 Doster W.

For myoglobin, the Gauss deviation amounts to 0.75, which is significantly above 0.5, the values for a standard distribution. It is independent of the temperature and the fitting range. Protein displacements are thus either locally anisotropic or heterogeneous. The deviation becomes even larger at shorter observation times, suggesting anisotropy relaxation. The second moments $<\Delta x^2>= <\Delta r^2>/3$ are temperature dependent and display the well-known "dynamical transition" around 180 K at $\tau_{res}= 140$ ps. Also shown are the short time displacements at $\tau_{res} \cong 15$ ps, measured with the time of flight spectrometer IN6 (ILL Grenoble). The displacements start to deviate from linearity above 200 K, but on a lower level. It is more difficult to identify a transition temperature. The full line denotes the vibrational MSD calculated from the density of states [5, 29].

Figure 5 shows that the zero order term of the expansion A(T) in equation-13 deviates from unity increasingly above 50K. Thus, the normalized elastic intensities do not extrapolate to unity at Q = 0 except at the lowest temperatures. This important result, which has been observed frequently, is not compatible with single scattering theory. Below we show for the first time, that A (T) can be reproduced temperature dependent by second scattering calculations. Multiple scattering effects are largely ignored in the bio-neutron scattering field. In 1990 we have published a protein vibrational density of states, which was corrected for multiple scattering [5]. Also the data in our analysis of displacement distributions, the MSD values were corrected for multiple scattering [8].

Direct numerical Fourier transform of elastic scattering data

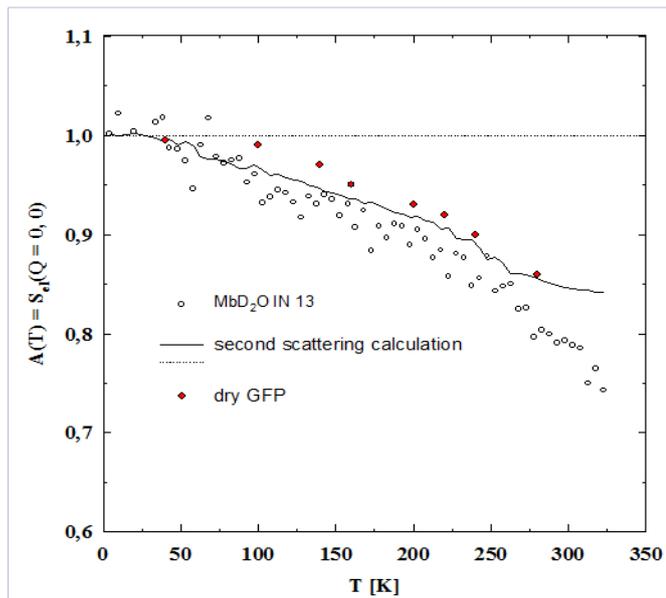By inverting the Fourier transform of equation (2) or its isotropic average of equation 7&8 numerically, one could in principle derive G(r, t ≅ τ_{res}) from experimental data. Figure 2 displays powder-averaged elastic scattering profiles from amorphous samples, thus only isotropic distributions can be obtained. After switching the powder average with the Fourier integral in equation (8), an isotropic displacement distribution is obtained according to:

$$\left\langle G_s\left(r,t=\tau_{res}\right)\right\rangle_{powder} = \frac{1}{2\pi}\int_{-\infty}^{\infty} dQ \cdot \exp(iQr) \cdot REISF(Q,\tau_{res}) \quad (15)$$

with REISF (Q, τ_{res}) being symmetric in ± Q. Unfortunately, even the very large experimental Q-range of IN 13 is not sufficient to calculate meaningful Fourier transforms. The alternative is to approximate the experimental data by analytical functions, which can be easily transformed. The most appropriate choice, according to equation (7a), is a sum of Gaussian distributions. The elastic scattering profiles in figure 2 are well approximated by a sum of two Gaussians with temperature dependent parameters $f_i(T)$ and $\gamma_i(T)$ [8]:

$$REISF_{GG}(Q,T) = f_1\exp\left(-Q^2\gamma_1\right) + f_2\exp\left(-Q^2\gamma_2\right) \quad (16)$$

In Figure 6 the result of this procedure is displayed as $4\pi r^2< G(r, t_{res})>_{power}$ at various temperatures.

This quantity denotes the conditional probability that a particle, which started at the origin at t = 0, has moved to a shell between r and r+dr during an interval. t = τ_{res}.

Figure 5: Zero-Q value A (T) derived from the polynomial fit to equation 13 (open circles) together with calculations of the multiple scattering intensities (full line). Also shown are the reconstructed elastic intensities at zero Q of dry GFP by Frauenfelder et al. (red circles) [19].
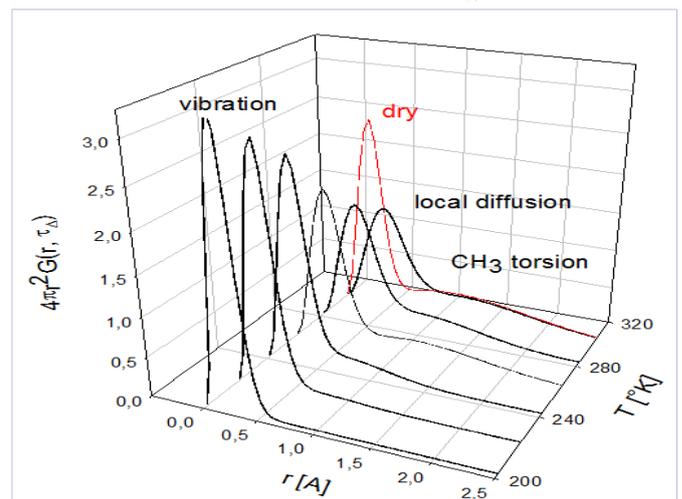
Figure 6: Displacement distribution derived by analytical inversion of the data of figure 2 adjusted to equation 16, versus the temperature: 200, 220, 240, 260, 280 and 300 K. The classification as local diffusion and methyl torsion is only suggestive at this stage. The black curves reflect the hydrated sample at 0.38 g/g, while the red curve is derived from a lyophilized sample (<0.05 g/g). Local diffusion is arrested, but rotational transitions are preserved as in the hydrated case.

Citation: Doster W (2018) Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions. Int J Mol Theor Phy. 2(1):1-14.

Page 6 of 14

Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions

Copyright:
© 2018 Doster W.

The resulting displacement distribution exhibits two maxima, which evolve with the temperature: Below 180 K only a single maximum is recorded at $r \cong 0.2$ Å, which broadens with increasing temperature. It reflects mostly vibrational motions. At about 240K, this maximum starts shifting to $r \cong 0.5$ Å and its width increases. This effect is compatible with a Gaussian distribution of small scale displacements, which become resolved above this temperature. It is only observed with D_2O-hydrated samples, dry samples do not show this feature [8]. Above 180 K a second maximum emerges around 1.5 Å, which can be found in all data sets independent of how the protein environment was prepared, hydrated, dehydrated or vitrified. Such large excursions point to rotational transitions.

This suggests that the bimodal distribution reflects two specific molecular processes. It should be stressed, that the method of inverting the Fourier transform using a sum of two Gaussians differs formally from the "dynamical heterogeneity" concept [20, 24, 26]. There a sum of Gaussians is used to explain the non-Gaussian nature of the scattering function by site heterogeneity. Since scattering is a local process, each atom "j" has its own dynamical structure factor. In this model all the local structure factors are assumed to be Gaussian with distributed second moments:

$$EISF_{DH}(Q) = \frac{1}{N} \sum_{j} \exp\left\{-Q^2 \left\langle \Delta x_j^2 \right\rangle\right\} \quad (17)$$

Intrinsic non-Gaussian scattering is thus excluded. This model is well adjusted to simulations, since the atomic scattering functions are available. To interpret experiments, one introduces groups of similar sites associated with MSD distribution functions $g(\alpha)$ [20, 24, 26]:

$$S_{DH}(Q,T) = \int d\alpha \cdot g(\alpha) \cdot \exp\left(-Q^2 \cdot \alpha\right) \quad (18)$$

with $\alpha = \langle\Delta x^2\rangle$. Exponential, Gamma and also bimodal distributions have been suggested in dynamical heterogeneity work of globular proteins [20]. The $g(\alpha)$ is a temperature independent distribution of weakly temperature dependent second moments $\langle\Delta x^2\rangle$. The well resolved bimodal distribution of figure. 6 suggests two types of processes, I and II, with different dynamic characteristics and dependence on the protein environment. This result is the main motivation to suggest a minimal model of protein dynamics based on just two processes.

A principle component model of protein dynamics

The third possibility to evaluate the dynamic displacement distribution is to adjust an analytical model to the data. Several attempts to derive dynamical models from a structural classification of hydrogen atoms and MD simulations have been published. Up to five components were proposed [24, 30-32].

It is fair to conclude, that the available data base and the many parameter fits were not sufficient to establish conclusive results. In the "three classes of motion" paper by Hong et al. a wide frequency range combining three spectrometers is presented, but the analysis is based a single averaged Q-value [32]. Dellerue et al. combine inelastic neutron scattering experiments with analytical models and simulations [31]. To represent local diffusion, they use the Dianoux-Volino model of free diffusion inside a sphere [33]. Because of the limited Q-range of the Mibemol spectrometer, they could not test the validity of their model. The authors display simulated intermediate scattering functions of main and side chain atoms, depending on the distance from the surface. Engler et al. derive a three-Gaussian model of displacements from the neutron structure of myoglobin [24]. By modelling the temperature dependence and accounting for the structural disorder, good fits of the elastic profiles in fig. 2 are reported. Unfortunately, we failed to reproduce these fits, in our view the rotational transitions are underestimated, which are not Gaussian distributed. It is thus not the lysine residue, which gives the strongest contribution to the displacements.

Motivated by the results of figure 6, we start with a bimodal distribution of sites associated with two kinds of motions: (I) internal rotational transitions and (II) local translational diffusive displacements. Thus a bimodal correlation function of principle components is assumed:

$$\Phi_s(Q,t) = \sigma_I \cdot \Phi_I(Q,t) + (1-\sigma_I) \cdot \Phi_{II}(Q,t) \quad (19)$$

σ_I denotes the fractional cross section of the type I sites.

Equation-(19) has to be complemented by a term accounting for global protein diffusion, if the degree of hydration exceeds 0.4 g/g.

According to the neutron structure of myoglobin 25 % of the total number of hydrogen's are organized in methyl groups, thus $\sigma_I = \sigma_m \cong 0.25$ [24].

More specifically, type I motions are defined by a three site jump model of methyl groups reorienting by 120° jumps about their three-fold symmetry axis [8, 23]:

$$\Phi_{rot}(Q,t) = \frac{1}{3}\left\{1 + 2j_0(Q) + 2 \cdot (1-j_0(Q)) \cdot \exp(-t/\tau_{rot})\right\} \quad (20)$$

$$j_0(Q) = \sin\left(Q\sqrt{3} \cdot r\right) / \left(Q\sqrt{3} \cdot r\right) \quad (20\,a)$$

is the zero order Bessel function, with r = 1.03 Å is the length of the C-H bond. Exponential relaxation or a sharp barrier height is assumed for simplicity. Note that $\tau_{rot} = \tau_{Met}/3$ and that the $EISF_{met}(Q) = \Phi_{rot}(Q, t \gg \tau_{rot})$ is not Gaussian.

Citation: Doster W (2018) Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions. Int J Mol Theor Phy. 2(1):1-14.

Page 7 of 14

Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions

Copyright:
© 2018 Doster W.

By contrast, type II processes involve small local displacements of residues confined by a quasi-harmonic potential. Instead of the Dinanoux-Volino model of free diffusion inside a rigid sphere[33], a continuous confining potential and Brownian motion seems more physical. The intermediate scattering function of the Brownian oscillator associated with an Ornstein-Uhlenbeck process is derived in the Supplement. The result is [36]

$$\Phi_B(Q,t) = \exp\left\{-Q^2\delta_B^2 \cdot \left(1 - \exp(-t/\tau_B)\right)\right\} \quad (21)$$

$\delta_B^2 = D/\gamma = K_BT/(m\,\omega_0^2)$ is the translational mean square displacement. D is the local diffusion coefficient and γ is the damping constant. The rate is given by $1/\tau_B = D/\delta_B^2$. Φ_B (Q, t >> τ_B) turns into a Gaussian elastic scattering function at long times:

$$\Phi_B\left(Q, t >> \tau_B\right) = \exp\left(-Q^2 \cdot \delta_B^2\right) \quad (22)$$

The analytical model comprising two principal components is thus defined by equation-(19), combining Φ_I (Q,t) = Φ_{rot} (Q,t) according to equation-(20), and Φ_{II} (Q,t) = Φ_B (Q,t) = Φ_{trans} (Q,t), of equation-(21), which we call briefly the "TR model" of translation and rotation.

TR model of elastic scattering profiles

Elastic scattering is collected in the frequency domain using a fixed window method, $\Delta E_{res} = \hbar/\tau_{res}$ at $\omega = 0$: The normalized elastic intensity S_{el}^N (Q, τ_{res}, τ_c) is determined by equation-(11). Within the approximation of equation-12, assuming a delta-correlated resolution function, this equals the intermediate scattering function at t = τ_{res}:

$$REISF(Q,T,\tau_{res}) \cong \Phi(Q,\ t = \tau_{res})$$

Figure 7a shows a decomposition of the elastic scattering profile at 270 K (figure 2) into a rotational and a Gaussian component according to the TR model. That the experimental data agree quite well with the predictions, is the first important test of the spatial properties of the TR model at a fixed temperature. The next step is to deduce the temperature dependence of the correlation times by fitting the data to the model equation as shown in figure 2. An Arrhenius plot of the resulting correlation times τ_{rot} (T) and $t_B = \tau_{trans}$ (T) are displayed in figure 7b. The third parameter, the translational MSD, $\delta_B^2 \cong 0.1$ (± 0.02) Å^2, shown in figure 1, depends very little on the temperature. That the strong temperature dependent displacements can be reduced to a nearly temperature independent model displacement, is a striking success of the TR model. The correlation times thus carry the full temperature dependence of the model. Both correlation times, as determined from elastic scans, compare well with the spectral analysis of methyl rotational transitions (I) and hydration water (II) [8, 10, 37-42]. The average methyl barrier amounts to H_I = 11 (± 1) kJ/mol, with a pre-exponential of A_I = 1.6 (± 0.5)10^{13}s^{-1}. Also, the translational correlation times, τ_{trans} (T), superimpose within experimental error with those derived for protein hydration water by neutron spectroscopy [10, 37]. The type II barrier amounts to H_{II} = 17 (± 1) kJ/mol with a pre-exponential of A_{II} = 5 (± 1)10^{13} s^{-1}.
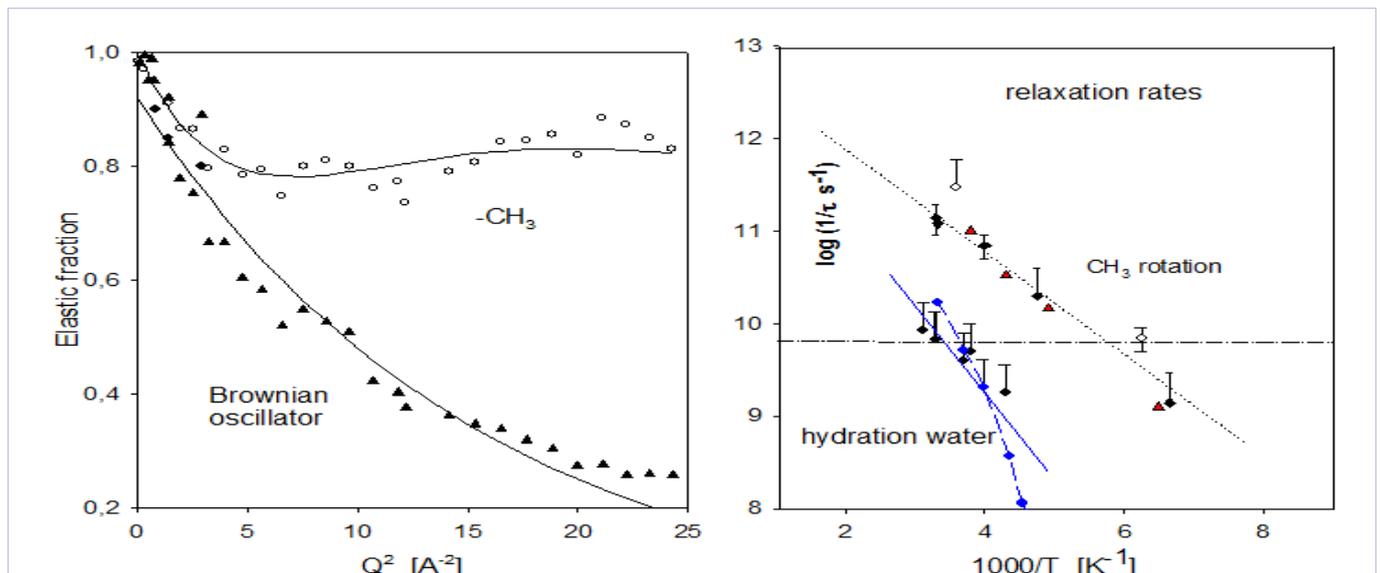
Figure 7(a): Decomposition of the elastic scattering function of figure 2 at 270 K according to the TR model. The lines reflect the theoretical adjusted components of equ. 20 (t >> τ_{rot}) and equation (22).
7(b): Arrhenius plot of the correlation times τ_{rot} (red triangles) and τ_{trans}(black circles) determined from the TR fit in comparison with a spectral analysis of methyl group rotation(black diamonds and open circles) and hydration water (blue circles) [10, 37].

Citation: Doster W (2018) Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions. Int J Mol Theor Phy. 2(1):1-14.

Page 8 of 14

Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions

Copyright:
© 2018 Doster W.

Time domain predictions of the TR model

The analysis of the elastic scattering data of figure 2 is now expanded to include the quasi-elastic spectra of IN13 and their associated time domain correlation functions, applying the previous considerations to the physiological temperature range. The relevant spectra were first published in 1989, together with a preliminary version of the minimal molecular model [4, 27, 43]. For the present purpose the spectral data were numerically Fourier transformed to the time domain.

The recorded dynamical structure factor, S (θ, ℏω), versus scattering angle θ and energy exchange ℏω, is first reorganized at a constant Q format, S (Q, ℏω), which is subsequently symmetrized by the detailed-balance factor. The time domain correlation function Φ_s (Q, t) is determined numerically by turning the Fourier integral of S (Q, ω) into a sum of discrete points using the experimental spectrum S (Q, ω_j):

$$\Phi_s(Q,t) = \hbar \int d\omega \, e^{i\omega \cdot t} S(Q,\omega) \cong \hbar \cdot \Delta\omega \cdot \sum_j e^{i\omega_j \cdot t} S(Q,\omega_j) \quad (23)$$

To avoid aliasing effects the smooth spectra S (Q, ω_j) are interpolated with a maximum number of n data points according to $t_n = n \cdot \dfrac{\pi}{\omega_{max}}$ (FFT), where ω_{max} is the cut-off of frequency of the spectrum. As a consistency check the transformed time domain result is back-transformed to the frequency domain. To de-convolute the data from the resolution function, a time domain low temperature spectrum of Mb-D_2O at 100 K was determined.

To exclude border line effects, the full time window of 140 ps was reduced to 50 ps. Beyond this limit, the noise is rapidly increasing due to low intensity spectral wings. We focus on a single temperature at 270 K, which is not affected by ice formation at h = 0.38 g/g. Apart from the 50 ps time window it is in particular the large Q-range of IN13, which imposes severe restrictions on the molecular model.

The property of time and frequency domain representation of spectra has been discussed in reference (10).

As Figure 8 shows, the agreement between experimental time domain data and the predictions of the TR molecular models is convincing. The derived time constants,

$\tau_{rot} \cong 11$ (±2) ps and $\tau_{trans} \cong 155$ (±20) ps, agree with those determined from the elastic analysis within experimental error. They define two types of molecular motions well separated in time. The correlation times cannot be fixed with higher precision due to limitations of the time window. By contrast, the constraints, imposed on the fits by the model function and the wide Q-range are significant. For comparison, the dashed lines in figure 8 show the result for the Brownian oscillator model, equation (21), which predicts a much stronger Q-dependence.
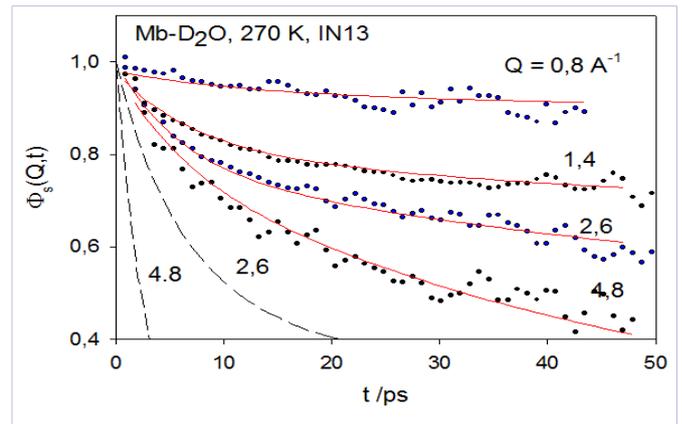
Figure 8: Time domain density correlation function Φ_s (Q, t) of myoglobin-D_2O (0.35 g/g), (IN13) at various Q values. The red lines are the predictions of the TR model with $\delta^2 = 0.11(\pm 0.02)$ Å^2, $\tau_{rot} \cong 11$ (±2)ps and $t_{trans} \cong 155$ (±20) ps at 270 K. The dashed lines are the predictions of the Brownian oscillator model of equation 21

Figure 9 finally shows the intermediate scattering function of hydrated myoglobin determined by combining the transformed spectra of three spectrometers. The figure illustrates the advantage of the time domain analysis. Although the decay extends over at least three decades, it is straightforward to recognize two main processes. The initial decay below 1 ps reflects the vibrational dephasing of low frequency vibrations [4, 5]. The fit at Q = 1.95 Å^{-1} reproduces the data very well. Also shown are results at other Q-values with significant deviations. Again the time constants are compatible with the Arrhenius plot of figure 7b. The long time value of type I, Φ_I (Q, t >> τ_{rot}) $\cong$ 0.8 (dashed) approximates the predicted value assuming a partial cross section of $\sigma_I = 0.28$ and equation 23 at Q = 2Å^{-1}. This indicates that type I reflects mostly methyl group rotational transitions. The time constants are τ_{rot} = 9 (±1) ps and τ_{trans} = 135 (±10) ps at 300 K. The local diffusion coefficient of type II amounts to $D_B = \tau_{trans}^{-1} \delta_B^2 \cong 10^{-7}$ cm^2/s, which
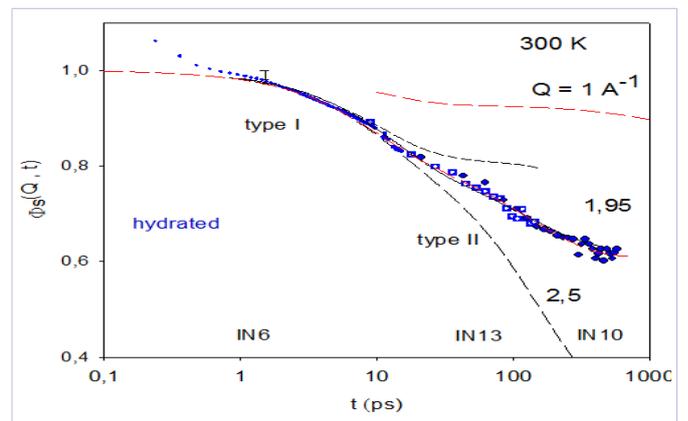
Figure 9: Correlation function of hydrated (0.35 g/g) myoglobin combining the transforms of spectra with three spectrometers at Q = 1.95 Å^{-1}, IN6 (blue crosses), IN13 (open blue squares) and IN10 (blue circles). Lines: fit to the TR model, τ_{rot} = 9(±2) ps, τ_{trans} = 145(±10) ps, δ^2 = 0.11(±0.01) Å^2. Dashed: predictions at alternative Q-values, short dashed line: methyl relaxation only (equation 20)

Citation: Doster W (2018) Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions. Int J Mol Theor Phy. 2(1):1-14.

Page 9 of 14

Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions

Copyright:
© 2018 Doster W.

is by a factor of 10 less than the diffusion coefficient of hydration water, $D_{hyd} = \Gamma_{hyd}/Q^2$, $D_{hyd} \cong 10^{-6}$ cm^2/s. Bulk water is in the range of 2.5×10^{-5} cm^2/s [8]. The protein-water fluctuations occur on a similar time scale, but the squared average displacements of water and protein residues within this time interval differ by a factor of ten.

Multiple scattering corrections to single scattering theory [5, 23, 29]

The density correlation function of equation (1) is normalized due to particle conservation at all times, which applies also to its Fourier transform Φ_s (Q, t) at Q = 0:

$$S_{el}^N (Q = 0, \tau_{res}) = \Phi_s (Q = 0, \tau_{res}) = \int d^3\vec{r} \cdot G(\vec{r}, \tau_{res}) \equiv 1 \quad (24)$$

Equation 24 shows, that elastic scattering functions, normalized to the lowest temperature, must extrapolate to unity at zero Q: $S_{el} (0,T)/ S_{el} (0,T_{low}) \equiv 1$. S_{el} (Q, T_{low}) defines a Q-independent constant in the absence of quasi-elastic scattering. It follows, that the observed temperature dependent zero-Q intensity of figure 5, A (T) $\leq$ 1, is incompatible with single scattering theory. Depending on the transmission, there is always a multiple scattering fraction, mainly of the type elastic-elastic, which creates a Q-independent background and thus additional intensity at Q = 0. Since the elastic single scattering intensity decreases with the temperature, the second scattering effect will become smaller as well (figure 5). We have explicitly calculated the effect of two sequential elastic scattering processes on the elastic scattering intensity at Q = 0 from an infinite plane slab sample of thickness d. As input, one needs the single scattering structure factor. For this purpose the experimental elastic scattering profiles of figure 2 were parametrized by two Gaussian functions at each temperature. The resulting second scattering intensities versus Q and temperature are displayed in figure 10.
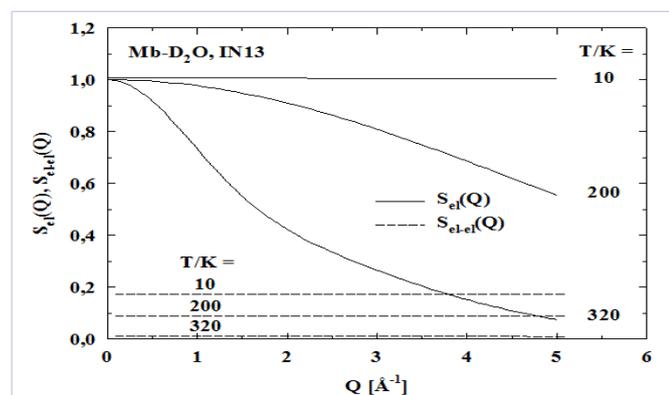
S_{el-el} (Q) is approximately independent of Q, and decreases with the temperature. The associated S_{el-el} (Q = 0, T), displayed in figure 5, agrees rather well with the experimental A(T) values, determined from the moment expansion. The deviations at high temperature result from increasing quasi-elastic scattering. The reconstructed elastic intensity of the green fluorescent protein (GFP) extrapolated to Q = 0 by Frauenfelder et al. fit quite well into this picture [19]. The temperature dependence and the absolute values of A_{GFP} (T) are compatible with multiple scattering, given the uncertainties of the extrapolation procedure and the unknown transmission.

Protein function

Inside biological cells proteins operate in concentrated solution. We have recently investigated, using neutron scattering, protein diffusion in concentrated systems including the diffusion of hemoglobin inside red blood cells, the latter facilitates the oxygen exchange in the lung [44]. Due to the dominant solvent scattering in this milieu, it is rather difficult to observe protein-internal motions on the time scale of bulk water. Hydrated samples at reduced solvent scattering and global diffusion are more suitable. We have compared hydrated samples with those in solution with proteins of different structure [45]. This raises the question, whether protein function is modified by increasing the concentration and by reducing the solvent content. Are hydrated proteins active? The motivation to study the dynamics of myoglobin was powered by the possibility to record its biological function within a wide range of temperatures and solvent conditions. My group combines functional studies with neutron scattering and Mössbauer spectroscopy [13, 46-49]. Here we focus on the effect of hydration on protein function. Myoglobin binds dioxygen reversibly, but also and more strongly other ligands like carbon monoxide. Photo-dissociation of myoglobin-

Figure 10: Full line: the elastic scattering functions of figure 2 were parametrized by a sum of two Gauss functions, from which the second scattering was calculated versus the temperature (dashed). Double scattering is nearly Q-independent, yielding finite value at Q = 0, decreasing with increasing temperature.

Figure 11: Flash photolysis experiments of hydrated myoglobin- CO films at different humidity and 300 K. The bound state (Fe-CO) is restored after photolysis employing a 20 ns laser flash at 532 nm. The ligand CO binds in two kinetic steps denoted by I (internal) and II (surface).

Citation: Doster W (2018) Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions. Int J Mol Theor Phy. 2(1):1-14.

Page 10 of 14

Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions

Copyright:
© 2018 Doster W.

CO is achieved by a 20 ns NdYaG laser pulse at 532 nm (Figure 11). Using a home-made flash system, ligand recombination is recorded between 20 ns and 1 s on a logarithmic time scale [46]. The kinetic experiments can be explained by a simplified kinetic scheme:

$$Mb\text{-}CO \leftrightarrow (Mb+CO)_{int} \leftrightarrow Mb + CO$$

$$I \qquad\qquad II$$

After complete photolysis by a laser flash, the ligand CO initially resides inside the protein matrix. Fast recombination from internal states can occur in competition with ligand escape across the protein surface. Figure 11 shows the recombination kinetics of photo-dissociated myoglobin +CO in hydrated films, which were fabricated by spin coating. At full hydration (100 % humidity) a single process is observed, which resembles ligand recombination from the solvent in liquid samples [46, 47]. Decreasing the degree of hydration, reduces the amplitude of the solvent process, while fast internal recombination gains in importance. This implies a reduced escape rate across the surface at lower hydration. This trend is enhanced at 0% humidity, the escape rate to the solvent and the recombination rate is further diminished. Note the logarithmic scale of the dissociated fraction. But a limited escape occurs even in the lyophilized state. The internal binding process is thus little affected by further drying as expected for a type I processes. The kinetic steps involving the protein surface and the solvent are of type II. Most neutron scattering experiments were performed with samples exposed to 90-100 % of humidity. In addition 0% samples were investigated. In conclusion, the hydrated films show protein function, which approaches the native kinetic properties in solution at full hydration. Dehydration diminishes ligand escape across the protein surface, while internal recombination is only weakly affected. The rate constants of the elementary kinetic steps range from 100 ns to seconds, overall binding at full hydration occurs within milliseconds. This is much slower than the time scale of microscopic molecular dynamics observed with neutron scattering. For type II processes a direct correlation between microscopic dynamics and ligand escape was established. The escape rate depends on the viscosity η_s and the solvent relaxation time τ_B near the protein surface according to Kramers law of activated escape [46-48]:

$$k_{esc} \sim 1/\tau_B exp\left(-H_{esc}/RT\right) \quad with \quad \tau_B \sim \eta_s \qquad (26)$$

The mechanism of internal ligand displacements and recombination with the heme iron is less well understood. It seems likely however, that ligand migration between internal cavities is assisted by rotational transitions of side chains. Molecular dynamic simulations indicate that the methyl groups that exhibit many rotational transitions are located near xenon cavities, which play a role in internal migration of ligands in myoglobin [42]. The authors also consider the rotational transitions of methyl groups and local diffusion as major excitations of protein dynamics.

Discussion

Energy-resolved neutron scattering allow to record space-time density fluctuations of hydrogen atoms in biomolecules. The basic quantity is the hydrogen-weighted incoherent density-density correlation function G_s (r, t), describing single particle displacements of isotropic samples. In this article three methods were discussed, how to reconstruct dynamic displacement distributions in proteins from incoherent neutron scattering data. The most general approach is the moment expansion, which can provide the second and the forth moment of the distribution at fixed time. The second moment increases above a certain transition temperature, where the respective molecular process is resolved by the spectrometer. Consequently the second moment carries the time tag of the instrumental resolution time. Smaller displacements are recorded at 15 ps compared to 150 ps as shown in Figure 4. The non-Gaussian nature of the elastic scattering profile results in a Gauss-deviation of 0.75 well above the standard value of 0.5. This suggests asymmetric molecular displacements, which are not temperature dependent above 220 K. But the asymmetry decreases with time.

The second approach is focused on a numerical inversion of equation 2. Instead of using a direct numerical transform, we transform a sum Gaussian distributions, which was adjusted to the elastic scattering profiles of figure 2. Since already two well separated Gaussians can account for the data, we obtain a bimodal G_s(r, t), with displacements around 0.5 Å (type I) and 1.5 Å (type II). This result suggests that two specific molecular processes may be sufficient to reproduce the neutron scattering spectra of proteins. Thus a principal component analytical model was proposed, which accounts rather well for the data, both in the time domain but also versus momentum exchange. The latter is a unique property of neutron scattering. The two processes comprise the well-known methyl rotation at 10 ps and local diffusion of residues on a 150 ps time scale. The second momen $\delta_B^2 \cong 0.1$ Å^2 is approximately temperature- independent as Figure1 shows. The correlation times thus carry the full temperature dependence of the apparent displacements in figure 1: The onset temperatures disappear with the TR model. As consequence, the Zaccai-Bicout model, which suggests intrinsic discontinuities in the protein force constants at the transition temperatures, does not apply. Both molecular motions, methyl rotation and hydration water relaxation are well known processes determined by other methods [37, 38-42]. If quasi-elastic spectra of proteins

Citation: Doster W (2018) Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions. Int J Mol Theor Phy. 2(1):1-14.

Page 11 of 14

Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions

Copyright:
© 2018 Doster W.

were indeed inhomogenious, the conventional space-time approach of correlation functions would fail [19]. Moreover, in the Frauenfelder model an elastic peak does not exist. This would invalidate the dynamic analysis of the data in Figure 2. Moreover the elastic intensity extrapolated to zero momentum exchange must first be corrected for multiple scattering, before it can be used to disprove scattering theory. The "ELM" approach was not very successful in enhancing our understanding of molecular motions in proteins during the last 30 years. Not a single energy landscape of a protein could be established. We thus prefer the more efficient molecular approach of minimal complexity. The heterogeneous scattering model of Frauenfelder et al. thus provides "no case against scattering theory" [15, 19 50].

The dynamical heterogeneity, that we observe, concerns the existence of two specific molecular processes. The two kinds of motions were identified before in functional studies: Flash photolysis experiments of ligand binding to myoglobin show that internal ligand displacements are entirely decoupled from the solvent (type I), while ligand exchange processes across the protein solvent interface (type II) vary with the solvent viscosity[46-48]. Moreover, multiple flash experiments demonstrate that heterogeneous non-exponential kinetic turns into homogenous and exponential behavior above the glass temperature of the solvent [48]. According to Arieh Warshel, complex energy landscapes are not the reason of the catalytic power of proteins [1]. Our ligand binding experiments demonstrate however, that the exchange of substrate and product between the active site and the solvent requires structural flexibility and the assistance of type II fluctuations. Proteins are designed to provide a catalytic environment perfectly isolated from the solvent, which is supported by solvent independent type I processes.

With an improved data base one my discover further motional components: The arrow in figure 1 at 130 K points to such an unresolved transition. One likely possibility is the presence of a small amount of acetate in the buffer, since the fast acetyl methyl groups give rise to low temperature onsets [10, 39]. Time domain neutron scattering will probably play a more important role in the future analysis of high quality experiments. The unique potential of neutron scattering to derive time resolved mean square displacements was first demonstrated for protein hydration water [8, 28]. Here a direct comparison with simulations is possible, unexplored up to now.

Acknowledgements

Supplement: derivation of equation (21) [35]

Starting with the Fokker-Planck equation of the Ornstein-Uhlenbeck process, Kneller has derived intermediate scattering functions for coherent and incoherent neutron scattering by macromolecules [34]. He demonstrated that the Langevin description on a coarse-grained time scale is identical with the Smoluchowski description, if friction dominates.

Over-damped motion in a simple harmonic potential $V(x) = \frac{1}{2}\, m\omega_0^2 x^2$ is described by the one-dimensional Langevin-Smoluchowski equation [35]:

$$\dot{x} = -\frac{\omega_0^2}{\gamma} x + \Gamma(t)/\gamma \qquad (27)$$

with a force constant $K_{BO} = m\omega_0^2$ and the friction coefficient γ. $\Gamma(t)$ denotes the δ-correlated random force. The structural relaxation time is given by $\tau_B = \gamma/\omega_0^2$. The local diffusion coefficient is defined by $D = K_B T / (m\gamma)$. This model was applied previously to neutron scattering spectra of proteins [36]. To derive the relevant distributions, we start with the conditional probability $p(x, t| x_0, t`)$ of equ. 1 determined by a one-dimensional Ornstein-Uhlenbeck process [35]:

$$\frac{\partial p}{\partial t} = \frac{1}{\tau_B} \frac{\partial}{\partial x}(x \cdot p) + D \frac{\partial^2}{\partial x^2} p \qquad (28)$$

complemented by the initial condition:

$$p(x, t_0 / x_0, t_0) = \delta(x - x_0) \qquad (29)$$

For scattering applications one deals with the Fourier transform $\Phi_B(Q, t)$ of $p(x, t|x_0, t_0)$:

$$p(x, t / x_0, t_0) = \frac{1}{2\pi} \int e^{iQx} \Phi_B(Q, t / x_0, t_0) dQ \qquad (30)$$

which, if inserted into equation (28) yields:

$$\frac{\partial \Phi_B(Q,t)}{\partial t} = -\frac{1}{\tau_B} \cdot Q \frac{\partial}{\partial Q} \Phi_B - D \cdot Q^2 \cdot \Phi_B \qquad (31)$$

The solution to equation (31) is the requested result of equation (21):

$$\Phi_B(Q, t) = \exp\left\{-Q^2 \delta_B^2 \cdot \left(1 - \exp(-t/\tau_B)\right)\right\} \qquad (32)$$

Citation: Doster W (2018) Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions. Int J Mol Theor Phy. 2(1):1-14.

Page 12 of 14

$\delta_B{}^2 = D / \gamma = K_B T / (m\,\omega_0{}^2)$ is the translational mean square displacement. The rate is given by $1/\tau_B = D/\delta_B{}^2$. Inserting equation (32) into a powder averaged equation (30) yields:

$$p\left(r,r_0,t\right) = \left[2\pi\delta_B^2\left(1-\exp(-2\cdot t/\tau_B)\right)\right]^{-3/2} \cdot \exp\left[-\frac{(r - r_0\cdot\exp(-t/\tau_B))^2}{2\delta_B^2\left(1-\exp(-2\cdot t/\tau_B)\right)}\right] \quad (33)$$

Here the width of the equilibrium distribution is given by $\delta_B{}^2 = K_B T/(m\omega_0{}^2)$. The center of the distribution relaxes from r_0 at $t = 0$ to $r = 0$ at long times, while the width of the distribution increases from zero to $\delta_B{}^2$ during this time interval. Inserting equation (33) into equation (1) yields the time dependent displacement distribution function of the powder-averaged three-dimensional Brownian oscillator:

$$G_B\left(r,t\right) = \left(\pi\delta_B^2\cdot\left(1-\exp\left(-t/\tau_B\right)\right)\right)^{-3/2}\cdot\exp\left\{-\frac{r^2}{4\delta_B^2\cdot\left(1-\exp\left(-t/\tau_B\right)\right)}\right\} \quad (34)$$

This is a Gaussian displacement distribution with a time dependent width. The time evolution of $4\pi r^2\,G_B(r,\,t/\tau_B)$ is illustrated in figure 12 assuming $\delta_B{}^2 = 0.1$ Å^2. The maximum shifts with time to larger displacements, reaching the stationary distribution at long times. This result can be compared with the Gaussian type II distribution of figure 6, which, as a function of temperature at fixed time, shows a similar behavior. Higher temperatures are associated with larger time ratios $t/\tau_B(T)$. Note that the distribution of vibrational displacements is convoluted with the type II distribution in figure 6, which is not present in figure 12.
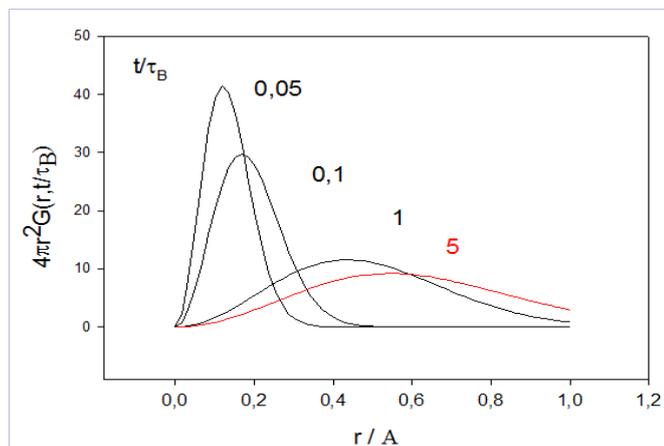

Figure 12: F: Displacement distribution of the powder averaged Brownian oscillator $4\pi r^2 G$ (r, t/τ_B) according to equation (34) with $\delta_B{}^2 = 0.1$ Å^2 at various time ratios, t/τ_B, the red curve represents the stationary distribution at t/$\tau_B \gg 1$.

References

1. Warshel A and Bora RP. Perspective: Defining and quantifying the role of dynamics in enzyme catalysis. J. Chem. Phys. 2016;144(18):180901. Doi: 10.1063/1.4947037.

2. Austin RH, Beeson KW, Eisenstein L, Frauenfelder H and Gunsalus IC. Dynamics of ligand binding to myoglobin. Biochemistry. 1975;14(24):5355-5373.

3. Doster W, Bachleitner A, Dunau R, Hiebl M and Lüscher E. Thermal properties of water in myoglobin crystals and solutions at subzero temperatures. Biophys. J. 1986;50(2):213-219.

4. Doster W, Cusack S, and Petry W. Dynamical transition of myoglobin revealed by inelastic neutron scattering. Nature. 1989; 337:754-756. Doi: 10.1038/337754a0.

5. Cusack S and Doster W. Temperature dependence of the low frequency dynamics of myoglobin. Biophys. J. 1990;58(1):243-251.

6. Smith JC. Protein dynamics: comparison of simulations with inelastic neutron scattering experiments. Q. Rev. Biophys. 1991;24(3):227-291

7. Steinbach PJ, Loncharich RJ and Brooks BR. The effects of environment and hydration on protein dynamics: a simulation study of myoglobin, Chem. Phys. 1991;158(2-3):383-394.

8. Doster W and Settles M. Protein-water displacement distributions. Biochim. Biophys. Act. 2005;1749(2):173-186.

9. Doster W. Concepts and misconceptions of the protein dynamical transition. Eur. Biophys. J. 2008;37(5):591-602. Doi: 10.1007/s00249-008-0274-3.

10. Doster W, Nakagawa H and Appavou MS. Scaling analysis of bio-molecular dynamics derived from elastic neutron scattering experiments. J. Chem. Phys. 2013;139(4):45105. Doi: 10.1063/1.4816513.

11. Hong L, Glass DC, Nickels JD, Perticaroli S, Zheng Y and Madhusudan T, et. al. Elastic and conformational softness of a globular protein. Phys. Rev. Lett. 2013;110(2):028104 .

12. Parak F, Frolov EN, Mössbauer RL and Goldanskii VI. Dynamics of metmyoglobin crystals investigated by nuclear gamma resonance. J. Mol. Biol. 1981;145(4):825– 833.

13. Lichtenegger H, Doster W, Kleinert T, Birk A, Sepiol B and Vogl G. Heme solvent coupling, a Mössbauer study of myoglobin in sucrose. Biophys. J. 1999;76(1 pt 1):414-422.

14. Keller H and Debrunner PG. Evidence for conformational mean square displacements in frozen aqueous solution of oxymyoglobin. Phys. Rev. Lett. 1980;45:68–71.

15. Frauenfelder H, Fenimore PW and Young RD. A wave-mechanical model of incoherent quasi-elastic scattering in complex systems. Proc Natl Acad Sci U S A. 2014;111(35):12764-12768. Doi: 10.1073/pnas.1411781111.

16. Doster W. Critical review of published papers in bio-neutron scattering. Available from: www.bioneutron.de.

17. Zaccai G. How soft is a protein, a protein dynamics force constant measured by neutron scattering. Science 2000;288(5471)1604-1607.

18. Gabel F, Bicout D, Lehnert U, Tehei M, Weik M and Zaccai G. Protein dynamics studied by neutron scattering. Qu. Rev. Biophys. 2002;35(4):327-367.

19. Frauenfelder H, Young RD and Fenimore PW. The role of momentum transfer during incoherent neutron scattering is explained by the energy landscape model. Proc. Natl. Acad. Sci. USA. 2017;114(20):5130-

Citation: Doster W (2018) Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions. Int J Mol Theor Phy. 2(1):1-14.

Page 13 of 14

Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions

Copyright:
© 2018 Doster W.

5135. Doi: 10.1073/pnas.1612267114.

20. Peters J and Kneller GR. Motional heterogeneity in human acetylcholesterinesterase revealed by a non-Gaussian model for elastic incoherent neutron scattering. J. Chem. Phys. 2013;139(16):165102. Doi: 10.1063/1.4825199.

21. Lovesy SW. Dynamics of solids and liquids by neutron scattering. Topics in current physics. Springer Verlag Heidelberg. 1977;1-10.

22. Squires GL. Introduction to the theory of thermal neutron scattering, Dover Books in Physics. 1978;101.

23. Bee M. Quasi-elastic Neutron Scattering. Adam Hilger, Bristol. Philadelphia. 1988;17, 107-147, 195.

24. Engler N. Ostermann O, Niimura N and Parak F. Hydrogen atoms in proteins: positions and dynamics. Proc. Natl. Acad. Sci. USA. 2003;100(18):0243-10248.

25. Gaspar AM, Busch S, Appavou MS, Häussler W, Georgii R and Su Y, et. al. Using polarization analysis to separate coherent and incoherent scattering from protein samples. Biochim. Biophys. Act. 2010;1804(1):76-82.

26. Vural D, Hong L, Smith JC and Glyde HR. Motional displacement in proteins: the origin of wave-vector dependent values. Phys Rev E Stat Nonlin Soft Matter Phys. 2015;91(5):052705.

27. Doster W, Diehl M, Petry W and Ferrand M. Elastic resolution spectroscopy: a method to study molecular motions in small biological samples. Physica B. 2001;301(1-2):65-68.

28. Settles M and Doster W. Anomalous diffusion of adsorbed water, a neutron scattering study of hydrated myoglobin, Faraday Disc. 1996;103:269-280.

29. Settles M and Doster W. Iterative calculations of the vibrational density of states from incoherent neutron scattering data with the account of double scattering. Biological Macromolecular Dynamics. 1996:3-8.

30. Fitter J, Lechner RE, Büldt G and Dencher NA. Internal molecular motions of bacteriorhodopsin: hydration induced flexibility studied by quasi-elastic incoherent neutron scattering using oriented purple membranes. Proc Natl Acad Sci U S A. 1996;93(15):7600-7605.

31. Dellerue S, Petrescu AJ, Smith JC and Bellissent-Funel MC. Radially softening diffusive motions in a globular protein. Biophys J. 2001;81(3):1666-1676.

32. Hong L, Smolin N, Lindner B, Sokolov AP and Smith JC. Three classes of motion in the dynamic neutron scattering susceptibility of a globular protein. Phys Rev Lett. 2011;107(14):148102.

33. Volino F and Dianoux A. Neutron incoherent scattering law for diffusion in a potential of spherical symmetry: general formalism and application to diffusion inside a sphere. Mol. Phys. 1980;41(2):271–279.

34. Kneller GR. Inelastic neutron scattering from damped collective vibrations of macromolecules. Chem. Phys. 2000;261(1-2):1-24.

35. Risken H. The Fokker-Planck equation; methods of solution and applications. Springer. Berlin. 1984.

36. Doster W. Brownian oscillator analysis of molecular motions in biomolecules in Neutron Scattering in Biology, techniques and applications. Editors: Fitter J, Gutberlet T and KatsarasJ, Springer. Biological and medical physics, Biomedical engineering. 2006;chapter 20:461-482.

37. Doster W, Busch S, Gaspar A, Appavou MS, Wuttke J. and Scheer J. Dynamical transition of protein hydration water. Phys. Rev. Lett. 2010;104(9):098101-0981014.

38. Curtis JE, Tarek M and Tobias DJ. Methyl group dynamics as a probe of the protein dynamical transition. J Am Chem Soc. 2004;126(49):15928-15929.

39. Roh JH, Novikov V, Gregory R, Curtis J, Chowdhuri Z and Sokolov A. Onsets of An-harmonicity in protein dynamics. Phys Rev Lett. 2005;95(3):038101.

40. Chiro G, Caronna Ch, Natali F and Cupane A. Direct evidence of the amino acid side chain and backbone contributions to protein anharmonicity. J. Am. Chem. Soc. 2010;132(4):1372-74. DOI: 10.1021/ja908611p.

41. Telling MTF, Clifton L, Combet J, Frick B, Howells S and Garcia Sakai V. Lyophilised protein dynamics: more than just methyls? Soft Matter. 2012;8(37):9529-9532.

42. Krishnan M, Schulz R and Smith JC. Protein dynamical transition, role of methyl dynamics and local diffusion, AP Conference Proceedings. 2009;1102(1):122-131. DOI: 10.1063/1.3108363.

43. Doster W, Cusack S and Petry W. Dynamic instability of liquid-like motions in a globular protein observed by inelastic neutron scattering. Phys Rev Lett. 1990;65(8):1080-1083.

44. Doster W and Longeville S. Microscopic diffusion and hydrodynamic interactions of hemoglobin in red blood cells. Biophys. J. 2007;93(4):1360-1368.

45. Gaspar AM, AppavouM S, BuschS, Unruh T and Doster W. Dynamics of well folded and disordered proteins in solution: a time of flight neutron scattering study. Eur. Biophys. J. 2008;37(5):573-582.

46. Kleinert T, Doster W, Leyser H, Petry W, Schwarz V and Settles M. Solvent composition and viscosity effects on the kinetics of CO-binding to horse myoglobin. Biochem. 1998;37(2):717-733.

47. Doster W. The protein-solvent glass transition. Biochim. Biophys. Act. 2010;1804(1):3-13. Doi: 10.1016/j.bbapap.2009.06.019.

48. Longeville S and Doster W. Protein dynamics and function in 'Dynamics of Soft Matter, Neutron Applications', Chapter: 8. Springer Science and Business Media. 2012.

49. Post F, Doster W, Karvounis G and Settles M. Structural relaxation and non-exponential kinetic of CO-binding to horse myoglobin. Biophys. J. 1993;64(6):1833-1864.

50. Wuttke J. No case against scattering theory. Proc Natl Acad Sci USA.

Citation: Doster W (2018) Are Proteins Dynamically Heterogeneous? Neutron Scattering Analysis of Hydrogen Displacement Distributions. Int J Mol Theor Phy. 2(1):1-14.

Page 14 of 14