Microscopic Protein Diffusion at High Concentration

Sebastian Busch[1], Wolfgang Doster[1], Stéphane Longeville[2], Victoria García Sakai[3,4] and Tobias Unruh[5]
[1] Physics-Department E13, Technische Universität München, James Franck Straße 1, D-85747 Garching, Germany
[2] Laboratoire Léon Brillouin, CEA-CNRS, DSM-DRECAM, CEA Saclay, 91191 Gif-sur-Yvette, France
[3] NIST Center for Neutron Research, National Institute of Standards & Technology, Gaithersburg, MD 20899, U.S.A.
[4] Department of Materials Science and Engineering, University of Maryland, College Park, MD 20742, U.S.A.
[5] Forschungs-Neutronenquelle Heinz Maier-Leibnitz (FRM II), TU München, Lichtenbergstr.1, D-85747 Garching, Germany

ABSTRACT

The self-diffusion of myoglobin in concentrated solutions was investigated up to volume fractions of 0.4 by neutron back-scattering spectroscopy. The quasi-elastic spectrum can be decomposed into two Lorentz curves: (1) a narrow line, where the width increases with Q, which is assigned to translational diffusion, and (2) a broad Q-independent line, reflecting protein-internal motions. The apparent diffusion coefficient decreases with increasing concentration and wave-vector, suggesting that protein diffusion deviates at high Q (1.75 $Å^{-1}$) from its long-time Brownian limit. Jump diffusion, sample heterogeneity and time-dependent diffusion are discussed as possible explanations.

INTRODUCTION

Proteins are often studied in dilute, nearly ideal solutions. In contrast, biological fluids contain macromolecules at high concentration: for example, the volume fraction of hemoglobin in red blood cells is close to $\varphi = 0.25$. In most other cells, a large variety of macromolecules is present. These environments are termed “crowded” rather than “concentrated”. In both cases, excluded volume effects play an essential role in vivo, affecting reaction rates, association-dissociation equilibria and molecular transport [1]. In general, studies of protein dynamics in solution require information on protein diffusion [2]. In this contribution we focus on the transport properties of myoglobin, a small globular protein with a molecular weight of 18 kDa and a radius of gyration of 1.48 nm [3]. Myoglobin stores oxygen in muscle tissue and facilitates oxygen transport by macro-molecular diffusion.

Concentration gradients (more generally: gradients in the chemical potential) are compensated by collective diffusion D_c whereas the motion of the individual particle is described by self-diffusion D_s. We address the question whether protein self-diffusion in crowded systems can be understood based on concepts of colloidal solutions, with the size of the particles being in the range of several nanometers. In these systems dynamic forces such as hydrodynamic

interactions, in addition to direct interactions (van der Waals and electrostatic), play an important role [4].

Direct and hydrodynamic interactions can be separated by experiments which are sensitive to various time and length scales: hydrodynamic interactions are established almost with the speed of sound. In this hydrodynamic regime, the configuration of particles is essentially constant and the particles are performing short-time diffusion. Direct interactions enter after structural relaxation on a time scale corresponding to diffusion across the inter-particle distance: $\tau_I = d^2/6D^S_s$. d denotes the average center to center distance and D^S_s is the short-time self-diffusion coefficient. For times $\tau \approx \tau_I$, long-time diffusion is established with $D^L_s < D^S_s$ due to obstruction by direct interactions [4].

We have previously determined the inter-molecular structure factor of concentrated myoglobin solutions from small angle neutron scattering (SANS) data and a mean spherical analysis (MSA) assuming a modified hard-sphere potential [5]. Concentrated solutions up to volume fractions above 0.4 could be characterized as colloidal suspensions of hard-sphere monomers and are thus a good system to study concentration-dependent effects.

We also investigated protein diffusion on the scale of the inter-molecular distance by neutron spin-echo spectroscopy [5]. The respective diffusion coefficients decrease with increasing wave-vectors. As the method is based on coherent scattering, it records relative particle motions (collective diffusion). However, for $Q >> Q_{max}$, the wave vector of the structure factor maximum, S(Q) approaches 1 so that the self-diffusion coefficient is determined. The observed wave-vector dependence thus reflects the transition between collective- and self-diffusion. Moreover it was found that the effective friction corrected by the structure factor varies with the wave-vector as predicted for hydrodynamic interactions [5].

Due to a limited Q-range, the question of short-and long-time diffusion could not be addressed properly in the spin-echo experiments. To discriminate between different time-regimes it is necessary to extend the Q-range, while preserving the resolution of the spin-echo method. This is partially possible with back-scattering spectroscopy, where Q-values up to 1.75 $Å^{-1}$ at a resolution below 1 μeV are achieved. Moreover with this method incoherent scattering dominates and one determines the self-diffusion coefficient at all Q-values. For this purpose we employed the high flux back-scattering spectrometer at NIST [6].

EXPERIMENT

Met-myoglobin (equine skeletal muscle, M0630) was obtained from Sigma Chemicals. The protein was dissolved in H_2O and dialyzed against de-salted water to remove residual buffer. Subsequently, the solution was freeze-dried, then dissolved in D_2O for H/D exchange. The D-exchanged sample was freeze-dried again. For the experiment the protein powder was dissolved at the desired volume fraction ($\varphi = 0.13$, 0.26, and 0.39) in 100 mM NaCl/D_2O and one sample ($\varphi = 0.26$) was prepared with pure D_2O. The volume fractions were determined by weighing the sample.

Approximately 1ml of the solution, containing about 200 mg protein, was filled into a hollow cylinder (aluminum) with a diameter of 2.9 cm and a gap of 0.2 mm. Vacuum tightness was checked each time before the experiment. All experiments were performed at 293 K.

The High Flux Backscattering Spectrometer at NIST [6] covers a Q-range from 0.25 to 1.75 $Å^{-1}$. The 4 low-angle analyzers need special consideration, since they are not oriented

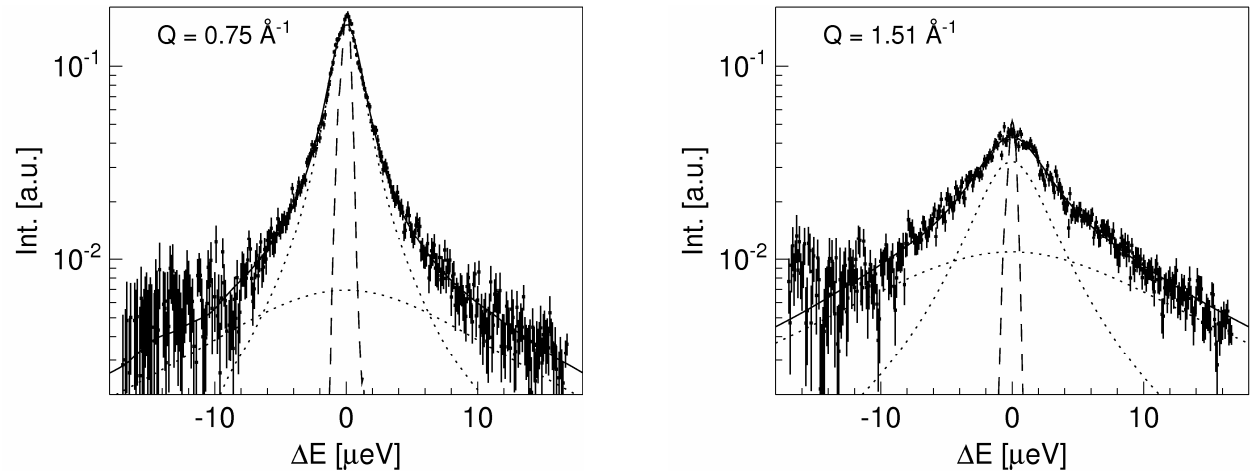

Figure 1: Backscattering spectra (dots with error bars) of a myoglobin solution at φ = 0.26 (in 100mM NaCl/D_2O) and Q = 0.75 Å^{-1} (left) resp. 1.5 Å^{-1} (right). Also shown are the two Lorentzian spectral components, L_1 and L_2 (dotted lines) and the fit (full line) as well as the resolution function (dashed line).

exactly in back-scattering geometry and are of different design. At the higher scattering angles, a resolution of 0.8 μeV (FWHM) was achieved using an energy range of ±17 μeV. Measurements with pure D_2O were performed to correct for the background, cell-, and solvent-scattering. The broad spectrum due to fast diffusion of water appears as an almost constant background in the frequency window. Subsequently, the scattering functions were normalized to vanadium.

RESULTS

Figure 1 shows a typical quasi-elastic spectrum obtained with a myoglobin solution at φ = 0.26 for two Q-values. For display, a logarithmic scale was chosen to expose the properties of the line-shape more clearly at high energy transfer. Also, the data were weighted logarithmically (disregarding their experimental errors) for the fit. The effective width of the spectrum, being much wider than the resolution function, increases with Q. A convolution of at least two Lorentzians with the following properties is required to fit the data: ($L_{1'}$) a narrow component with a line-width increasing with Q, it is assigned to long-range diffusion and ($\delta+L_{2'}$) a broad component where the linewidth is kept fixed. The second component thus reflects a localized molecular process and is assigned to protein-internal motions. The two-component dynamical structure factor can be written as [8]:

$$\begin{aligned} S(Q,\omega) &= a \cdot L_{1'}(\omega,\Gamma_{diff}) \otimes \left[A_0(Q)\cdot\delta(\omega) + \{1-A_0(Q)\}\cdot L_{2'}(\omega,\Gamma_{\text{int}})\right] \\ &= a \cdot \left[A_0(Q)\cdot L_1(\omega,\Gamma_{diff}) + \{1-A_0(Q)\}\cdot L_2(\omega,\Gamma_{diff}+\Gamma_{\text{int}})\right] \end{aligned} \qquad \text{equ. (1)}$$

$$\text{with} \quad L(\omega,\Gamma) = \frac{1}{\pi}\cdot\frac{\Gamma}{\omega^2+\Gamma^2}$$

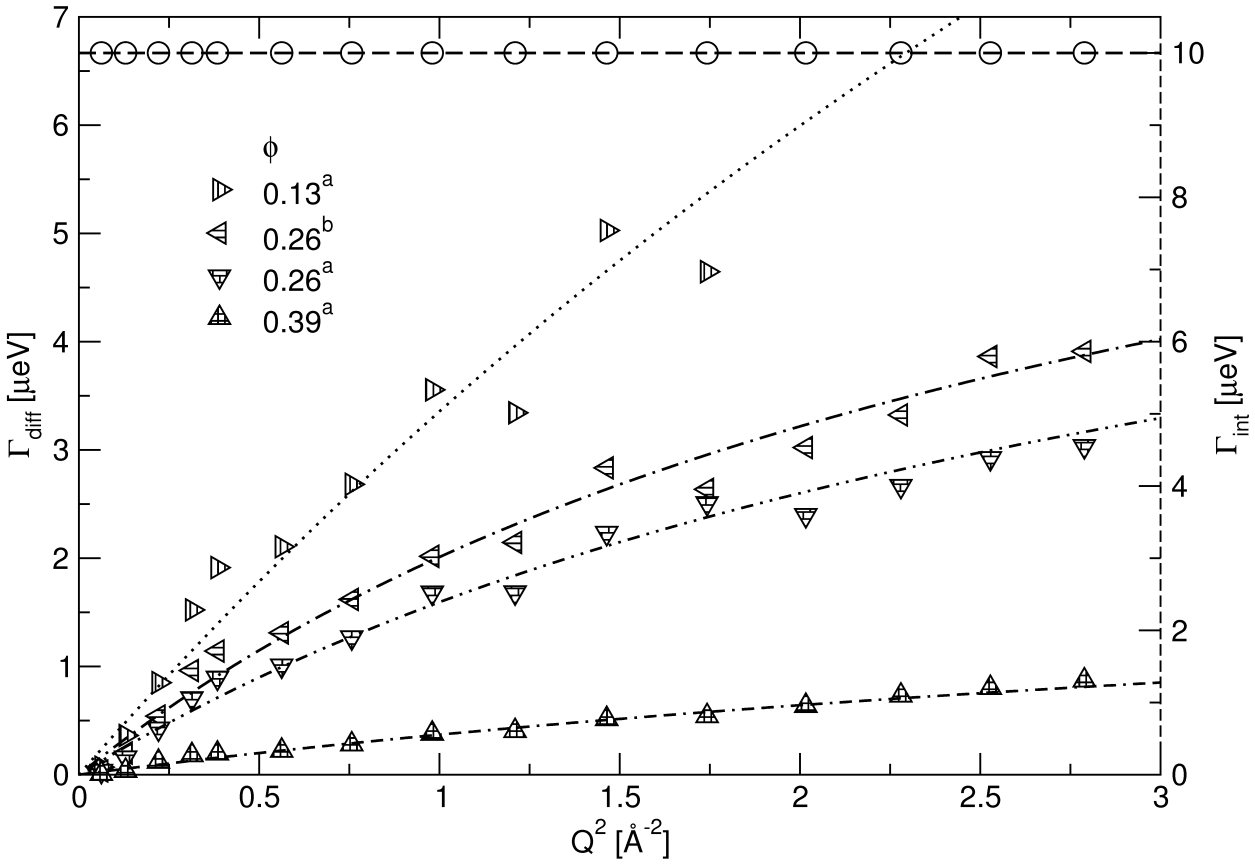

Figure 2: Linewidth (half width at half maximum, HWHM) versus Q^2 of the 2-component fit. The broad (internal) component Γ_{int} is represented by the data points and the dashed line at 10 μeV (scale on the right). The fit of the narrow (translational) component Γ_{diff} is represented by the data points at the indicated volume fractions (scale on the left hand side). [a]: with 100 mM NaCl, [b]: without NaCl. The lines are fits assuming the jump diffusion model of Singwi and Sjölander [7]. The respective parameters are given in Table I.

$A_0(Q)$ is the elastic incoherent structure factor (EISF) appearing here as the fraction of the amplitude of the translational component relative to the total scattering intensity. The broad line L_2 is the convolution of internal and diffusive motions. Γ_{diff} is the linewidth of translational diffusion, Γ_{int} denotes the (Q-independent) width of the internal process.

Long-range diffusion

The resulting linewidths Γ at various volume fractions of myoglobin are displayed versus the squared momentum transfer in Figure 2. For the internal processes a constant width of Γ_{int} = 10 (±2) μeV was obtained, which was kept fixed, while $\Gamma_{diff}(Q)$ increases with Q. The fact that $\Gamma_{diff}(Q \rightarrow 0) = 0$ shows that the broadening is indeed due to long-range diffusion.

For regular diffusion, a linear relationship of the type $\Gamma_{diff} = \hbar D_s Q^2$ is expected, where D_s denotes the self-diffusion coefficient. Such a Q^2-dependence is not observed, the slope rather decreases with increasing Q. This effect cannot be attributed to rotational diffusion: as discussed in ref. [2], at high Q-values or short length scales the displacements due to translational or rotational diffusion cannot be distinguished. Thus the linear relation between linewidth and Q^2 is not affected. One obtains an average diffusion coefficient including rotational effects.

The peculiar behavior of D(Q) saturating at high Q resembles results which have been obtained for water at various temperatures and on a much shorter time scale [7]. This effect was explained by assuming that diffusion on a short length scale is discontinuous: for small molecules the increase in linewidth is limited due to a finite jump rate, which is the asymptote at high Q. The associated length-scale is on the order of the size of the water molecule.

In our case of macromolecular diffusion the corresponding parameter is the volume fraction and not the temperature. From the class of jump diffusion models we select the Singwi-Sjölander model [8]. A particle executes an oscillatory motion for a time τ. Then it diffuses by continuous motion for a time τ_D, which is repeated. If the time of localization is much longer than the diffusion time, $\tau >> \tau_D$, the Singwi-Sjölander model yields the following Lorentzian line-width:

$$\Gamma_{diff} = \hbar \cdot \frac{D_s \cdot Q^2}{1 + D_s Q^2 \tau} \qquad \text{equ. (2)}$$

The line-width saturates at high Q yielding $\Gamma_{diff}(Q \to \infty) = \hbar/\tau$. This model fits the data (see Figure 2). The respective parameters are given in Table I. It is striking that the resulting jump length is for all volume fractions on the order of 1 Å.

Table I: Parameters obtained from the fit of the Singwi-Sjölander model to Γ_{diff}. [a]: with 100 mM NaCl, [b]: without NaCl

ϕ	$D\left[10^{-11}\,\frac{m^2}{s}\right]$	$\tau[ns]$	$\sqrt{\langle l^2 \rangle} = \sqrt{6D\tau}\left[\mathring{A}\right]$
0.13[a]	5.79	0.02	0.83
0.26[b]	4.07	0.08	1.40
0.26[a]	3.13	0.09	1.30
0.39[a]	0.66	0.27	1.03

Internal motions

Because of their finite spatial range, internal motions give rise to a finite elastic peak. The respective elastic fraction (EISF) was denoted above by $A_0(Q)$. Moreover, the line-width of a localized process is constant, $\Gamma_{int} = 10\ (\pm 2)\ \mu eV$, and independent of protein concentration. The simplest case of a localized process is described by diffusion of a particle in a harmonic potential. This model predicts a Gauss distribution of displacements, and the following elastic fraction [10]:

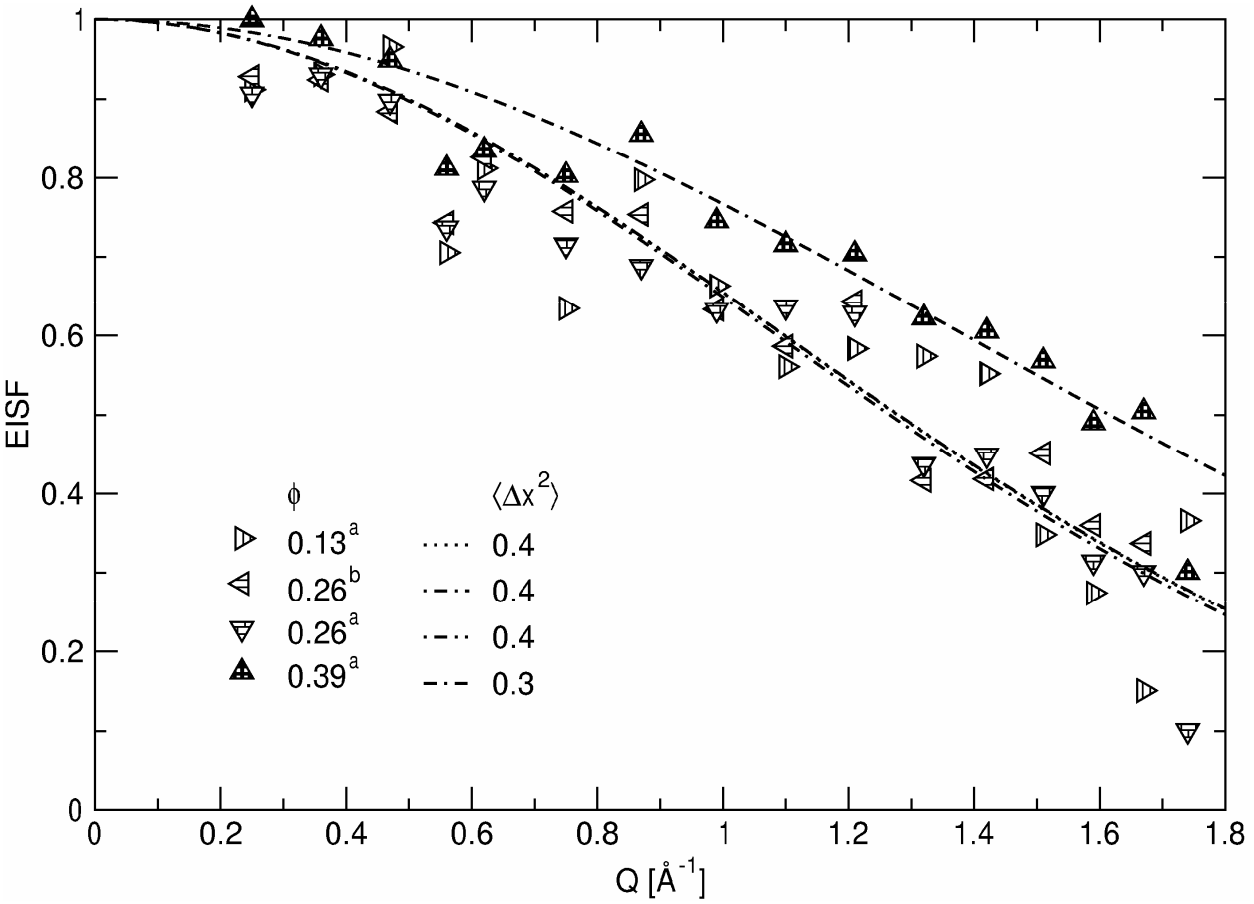

Figure 3: The EISF $A_0(Q)$ of the 2-component fit. [a]: with 100 mM NaCl, [b]: without NaCl. The lines are fits to a Gauss distribution, see equation 3. The resulting squared displacements $\langle \Delta x^2 \rangle$ [Å^2] are shown in the legend. Their error is below 0.05 Å^2.

$$A_0(Q) = \exp\left(-Q^2 \langle \Delta x^2 \rangle\right) \qquad \text{equ. (3)}$$

This equation fits the data quite well within experimental error as shown in Figure 3.

As expected for a harmonic potential, the correlations decay to zero at high Q instead of assuming a constant value. This result, together with the correlation time of 65 ps resolved by the spectrometer, suggests collective motions and not a reorientation of a molecular group. The motional amplitudes appear to be independent of concentration, except for the highest volume fraction, where the displacements are lower (Figure 3). The resulting displacements of $\sqrt{\langle \Delta x^2 \rangle} \approx 0.6$ Å are slightly larger than those of internal motions in hydrated proteins ($\sqrt{\langle \Delta x^2 \rangle} \approx 0.4$ Å) [2,10].

DISCUSSION

The most interesting result of this study is that the line-width associated with protein diffusion is not increasing with Q^2 but saturates at high Q, as shown in Figure 2. A jump diffusion model fits the data quite well, which implies discrete steps on a scale of 1 Å. The physical basis of discrete steps is not clear in the case of a macromolecule. One would have to consider a break-down of the continuum approximation due to the finite size of the water molecule. However even a small macromolecular displacement will displace many water molecules, the primary solvation shell of myoglobin contains at least 400 water molecules. We thus briefly discuss two alternative models, time-dependent- and heterogeneous diffusion.

Figure 2 shows that the effective diffusion coefficient decreases with increasing Q, i.e. decreasing length- and herewith time-scale. The diffusion coefficient thus seems to increase with time. As briefly discussed in the introduction, there are different typical time scales for the system [4]: the shortest time scale is set by the structural relaxation of the solvent τ_s, which is on the order of 10 ps for protein hydration water [9]. This is outside the time-window of our measurements ($\tau > 35$ ps). On the Brownian time scale of $\tau_B = M/f_0$ velocity correlations play a role. Here M is the mass of the protein and f_0 denotes the protein-solvent friction coefficient. The longest time scale is the interaction time τ_I, which is on the order of a few nanoseconds, depending on protein concentration.

In the Brownian time scale, the time-dependent diffusion coefficient D(t) is proportional to the integral of the velocity correlation function. Hence, D(t) increases with time, which could explain a decrease in D(Q) at high Q-values or short times [4]:

$$D(t) = \int_0^t \langle u(0)u(t')\rangle dt' = D_s \cdot \left(1 - \exp\left(-t/\tau_B\right)\right) \qquad \text{equ. (4)}$$

However, for myoglobin τ_B is presumably not much larger than τ_s, which is outside of our time window $\tau >> \tau_B$. For times $\tau > \tau_B$, hydrodynamic interactions are fully established and the particles are performing short-time self-diffusion D_s^S.

The spectral component L_1, is thus possibly heterogeneous, combining short- and long-time diffusion. Due to the finite energy resolution of the spectrometer, only the fast component is resolved at low Q, giving rise to a large initial slope in $\Gamma(Q)$. The slower process of long-time diffusion will contribute to the spectrum fully at high Q, which leads to an apparent saturation effect.

The component L_1, could also be heterogeneous for a different reason: protein solutions at high concentration may contain a variety of oligomers differing in molecular weight. The resulting distribution of diffusion coefficients can give rise to deviations from a Lorentzian line-shape. However, as mentioned above, in our previous structural and dynamic studies we could not detect oligomers or larger aggregates with myoglobin solutions even at high concentrations [5].

The differences observed between the samples with and without NaCl could result from electrostatic interactions or slight differences in concentration because of the different sample preparation. The effective diffusion coefficient (Table I) is slightly larger than those derived with the spin echo method at lower Q-values [5]. In this preliminary report we discuss some

characteristic results; a more detailed account, including further data, will be presented elsewhere.

ACKNOWLEDGMENTS

The project is supported by a grant of the Deutsche Forschungsgemeinschaft SFB 533 TP 10, which is gratefully acknowledged. This work utilized facilities supported in part by the National Science Foundation under Agreement No. DMR-0454672. For data analysis the program FRIDA by Joachim Wuttke was used. See http://sourceforge.net/projects/frida/.

The use of the commercial products identified in this paper does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

REFERENCES

1. R.J. Ellis, *Trends in Biochemical Sciences* **26,** 597 (2001)
2. J. Pérez, J-M Zanotti, D. Durand, *Biophysical Journal* **77**, 454 (1999)
3. W. Doster and M. Diehl (1994) unpublished
4. J.K.G. Dhont, *An Intoduction to the Dynamics of Colloids,* Elsevier, 1996, p.227–314
5. S. Longeville, W. Doster and G. Kali, *Chemical Physics* **292**, 413–424 (2003)
6. A. Meyer, R. M. Dimeo, P. M. Gehring, and D. A. Neumann, *Review of Scientific Instruments*, **74**, 2759 (2003)
7. J. Teixeira, M.-C. Bellissent-Funel, S. H. Chen, A. J. Dianoux, *Physical Review A* **31**/3, 1913–1917 (1985)
8. M. Bée, *Quasielastic Neutron Scattering,* Adam Hilger, 1988, p. 167 et seqq.
9. M. Settles and W. Doster, *Faraday Discussions*, **103**, 269–279 (1996)
10. W. Doster and M. Settles, *Biochimica et Biophysica Acta – Proteins & Proteomics*, **1749**, 173-186 (2005), Special Edition: Solvent Effects