Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The case for an empirical 'high-throughput' neutron scattering approach to protein dynamics

Giuseppe Zaccai

Institut Laue–Langevin, 6 Rue Jules Horowitz, BP 156, 38042 Grenoble CEDEX 9, France

Correspondence e-mail: zaccai@ill.fr

Received 14 April 2010
Accepted 2 June 2010

Recalling that there are currently more than 63 000 structures deposited in the Protein Data Bank, it is time for neutron scatterers studying protein dynamics to show a similar interest in the diversity that is the basis of biology. Sound experimental data that can underpin and complement molecular-dynamics simulations for fundamental research and health applications such as drug design can and should be provided. A proposal is presented to fulfil the two conditions that are required to enable such an approach: (i) the identification of measurable dynamics parameters that are correlated to biological function and activity and (ii) the design of experiments to measure these parameters efficiently with reasonable throughput.

1. Introduction

Neutron scattering has contributed significantly to understanding in the field of protein dynamics (Gabel *et al.*, 2002). It has an important advantage over other spectroscopies in that it provides experimental values for atomic displacements as a function of time (trajectories), which are the parameters that are calculated directly in molecular-dynamics (MD) simulations (*e.g.* Tarek & Tobias, 2000; Tournier & Smith, 2003; Glass *et al.*, 2010; Wood, Grudinin *et al.*, 2008; Kneller & Hinsen, 2009; Wood, Frolich *et al.*, 2008; Tobias *et al.*, 2009; Nakagawa *et al.*, 2008; Joti *et al.*, 2008). Neutron scattering elastic window temperature scans have revealed the harmonic behaviour of myoglobin (Doster *et al.*, 1989) and bacteriorhodopsin (Ferrand *et al.*, 1993) at low temperature and the dynamical transitions to more complex motions at around 200 K. Elastic window temperature scans have also revealed the dynamic coupling between a protein and its solvent environment (*e.g.* Cordone *et al.*, 1999; Cornicchi *et al.*, 2006; Gabel & Bellissent-Funel, 2007; Wood, Frolich *et al.*, 2008; Wood, Lehnert *et al.*, 2008). Physical models for protein dynamics have been proposed from neutron data and have been controversially debated (Doster, 2010).

The physical approach has mainly been concerned with mechanisms that are common to all soluble proteins, such as those involved in thermodynamic stabilization and hydration effects. The generic protein, however, does not exist in biology. By their very nature, their selection by evolution and the specificity of their structure–function relation, proteins are necessarily highly heterogeneous in their structures and dynamics, as attested by the more than 63 000 structures currently deposited in the Protein Data Bank (http://www.pdb.org). Dynamics models that fit the experimental neutron data by taking into account the heterogeneity of protein structures have been developed successfully by Smith and coworkers (Tournier & Smith, 2003; Glass *et al.*, 2010),

Kneller and coworkers (Wood, Grudinin *et al.*, 2008; Kneller & Hinsen, 2009), Tobias and coworkers (Wood, Frolich *et al.*, 2008; Tobias *et al.*, 2009) and Kataoka and coworkers (Nakagawa *et al.*, 2008; Joti *et al.*, 2008).

It is now well accepted that time-averaged protein structures at high resolution, although necessary, are insufficient to fully understand structure–function relations and that knowledge of dynamics on different time scales is also required. Interest in molecular interactions and conformational changes on slow time scales has led to a major expansion of small-angle neutron and X-ray scattering studies (SAS) at neutron and synchrotron-radiation installations. The fast picosecond to nanosecond time scale on which neutron and MD calculations are sensitive is the time scale associated with the forces that operate in protein structures and interactions, and the implication of fast protein dynamics in protein function and activity has been established by various techniques, including neutron scattering. Dynamics has also been implicated in the evolution of proteins to acquire different functions and activities (Tokuriki & Tawfik, 2009) and in drug design (Rosenblum *et al.*, 2007).

There are few neutron facilities in the world with appropriate instrumentation to measure protein dynamics and it is hoped that experimental opportunities and efficiency will increase significantly with the development of new spallation sources (Teixeira *et al.*, 2008). The number of different proteins studied by neutron scattering will have to be multiplied considerably. Elastic window temperature scans provide the fastest neutron scattering spectrometric characterization of a protein sample, with a full scan taking about a day of measuring time. There are several examples that show that the parameters derived from elastic temperature scans are related to biological function and activity and provide a very powerful characterization of protein dynamics for biological applications. For the results to have an impact on fundamental biology and its applications, such as drug design, they will have to be made readily available and useful to the entire molecular-dynamics research community. In the present context of structural biology, I suggest that the time is ripe for neutron scatterers interested in dynamics to follow the example set by crystallographers and 'small-angle scatterers' and to move on from model systems to the wide variety of biological molecular systems and to make their results available in an appropriate format by depositing them in a Dynamics Data Bank.

2. A brief theoretical introduction to the elastic temperature scan

A more detailed treatment is given in Zaccai (2010).

2.1. Incoherent neutron scattering

The incoherent scattering analysed by neutron spectrometry as a function of the scattering-vector modulus, Q, contains information on single-particle motions. The scattering from a biological sample is dominated by the motions of ^{1}H nuclei, the incoherent cross-section of which is much larger than those of the other nuclei present (and of ^{2}H or D, which can therefore be used as a label). On the neutron experimental time scale, H atoms are very good indicators of internal and global dynamics because they are distributed fairly homogeneously in biological macromolecules and they move with the chemical groups with which they are associated (Gabel *et al.*, 2002).

2.2. Scattering as a function of energy transfer

The scattered intensity as a function of Q and energy transfer depends on the dynamic structure factor of the sample, which is a double Fourier transform of a length–time pair-correlation function. In the case of incoherent scattering, only the self-terms are considered, so that the pair-correlation is between a single particle at time zero and the same particle at time t.

The dynamic structure factor for single-particle motions shows three distinct regions as a function of energy transfer (Gabel *et al.*, 2002). The elastic peak at zero energy transfer corresponds to scattering from a particle as it moves while localized well inside the window defined by the length scale and time scale defined by the Q range and the energy resolution of the spectrometer, respectively. Quasi-elastic scattering (QENS) corresponds mainly to diffusion of the single particle progressively out of the window and the inelastic component corresponds to single-particle motions reflecting collective excitations of given frequencies. In the scattering pattern of a protein, the elastic peak is the strongest and the easiest to measure from reasonable sample mass in a reasonable time. In an elastic window scan, the elastic peak is integrated over an energy-transfer range well within the instrumental resolution and measured as a function of sample temperature.

2.3. Mean-square displacements (MSDs) from the elastic window

An MSD value for single-particle motions localized well inside the experimental length–time window can be calculated from the Q-dependence of the elastic window intensity. Each particle sweeps out a volume that scatters neutrons like a homogeneous large particle of the same shape. The observed signal is from a large number of nuclei in the sample moving in different directions, so that the corresponding large particles take up random orientations. The scattering is incoherent from particle to particle, but coherent for each single nucleus particle as it moves in time. Note that I have just described a scattering system that is strictly analogous to that of SAS from a dilute solution of non-interacting particles. The same Guinier approximation analysis can be applied to calculate an MSD value for the motion, analogous to the radius of gyration in SAS. The virtual solution of large particles is not monodisperse, because nuclei may move with different geometries. For a given population, however, the MSDs for internal protein motions are likely to be similar, validating a common Q-range analysis (Réat *et al.*, 1997). The application of the Guinier approximation is not model-dependent apart from the

Table 1
Length–time windows of three spectrometers at the Institut Laue–Langevin.

Spectrometer	Energy resolution (μeV)	Q range (Å^{-1})	Length window (MSD; Å^2)	Time window
IN16	0.9	∼0.02–∼2	∼1–10	≤∼1 ns
IN13	8	∼0.2–∼5	∼1	≤∼100 ps
IN6	50 (5.9 Å wavelength)	∼0.2–∼2	∼1	≤∼10 ps

assumption that the motions are well localized inside the experimental length–time window. The Guinier approximation is a good approximation for any type of motion, provided only that the length–time window and Q-range constraints are respected.

The analysis has to be adjusted appropriately when there are contributions to the scattering from diffusive motions out of the experimental length–time window. Gabel (2005) performed simulations and carefully analysed the effect on the elastic intensity of the presence of diffusion and the instrumental resolution.

2.4. The spectrometer as a motion filter

The experimental length–time window examined on a neutron spectrometer is defined by its Q range and energy resolution. I give the Institut Laue–Langevin (http://www.ill.eu) spectrometers as examples (Table 1), but the treatment of course applies to all spectrometers with corresponding parameters at other neutron facilities.

The trajectories expressed as the MSD of three typical atomic motion populations for a protein in solution involved in internal motions (vibrations and conformational sampling), water diffusion and global protein diffusion, respectively, are shown in Table 2 for the three time windows in Table 1.

The values in Table 1 are indicative in that the edges of the windows are not sharp. Note how with respect to elastic scattering the three instruments are sensitive to different motions. Table 2 shows that the choice of spectrometer acts like a motion filter to focus on only certain aspects of the sample dynamics, which will dominate to a good approximation. Internal motions are within the window for all three cases. The contribution of global protein diffusion is not negligible only for the IN16 window. Water diffusion is outside the window for IN13 and IN16 and will contribute significantly on IN6. The length–time window of IN13 is therefore uniquely suited to measure internal dynamics in solution, without 'pollution' of the scattering by contributions from global diffusion or water diffusion. Using the spectrometer parameters as a motion filter has made it possible to investigate the non-negligible solvent isotopic effect on internal protein dynamics by measurements in H_2O solution on IN13 (Tehei *et al.*, 2001). Previously, only measurements in D_2O were performed, assuming that the signal from diffusion in H_2O would swamp all other information. The results showed how the internal dynamics of a protein from an extreme halophile was correlated with its stability in solvents of different salt type and concentration and water isotopic composition.

Table 2
MSD of atoms during internal motion in a protein (Å^2), water diffusion and global protein diffusion (for a globular protein in solution with the molecular mass of haemoglobin).

Process	MSD in 10 ps	MSD in 100 ps	MSD in 1 ns
Atom in protein	∼1	∼1	∼1
Water diffusion (for $D \simeq 2.5 \times 10^{-5}$ cm^2 s^{-1})	∼2.5	∼25	∼250
Global protein diffusion (for $D \simeq 2.5 \times 10^{-7}$ cm^2 s^{-1})	—	∼0.1	∼1

2.5. MSD as a function of temperature

In classical mechanics, the energy of a harmonic oscillator is proportionally related to its mean-square amplitude by the force constant. Similarly, it has been pointed out by Zaccai (2000) that the local slope of the MSD *versus* temperature, T, has the dimensions of the reciprocal of a force constant. In a harmonic regime, the MSD *versus* T slope is indeed related to the mean force constant, $\langle k \rangle$, of the motion. However, it can be shown that even in the case of motion in a far from harmonic potential (such as, for example, sampling between conformational substates), in the low-Q limit the local MSD *versus* T slope corresponds to the reciprocal of an effective mean force constant written as $\langle k' \rangle$, which expresses well the resilience of the system for the motion (Zaccai, 2000, 2010; Bicout & Zaccai, 2001; Kneller, 2005.

3. MSD *versus T* behaviour is correlated with biological function and activity

In biological terms, the MSD is associated with flexibility and the $\langle k' \rangle$ value with structural resilience. Measurement of MSD and $\langle k' \rangle$ values has established the specificity of protein dynamics in the adaptation to extreme temperature of the proteome in live bacteria (Tehei *et al.*, 2004) and in homologous proteins from organisms adapted to different physiological temperatures (Tehei *et al.*, 2005). The effect of high NaCl and KCl concentrations and water-isotope environments on the dynamics of a halophilic protein have been shown to be strongly correlated with protein stability and suggested mechanisms for adaptation to a KCl-containing cytoplasm in an organism from the Dead Sea (Tehei *et al.*, 2001). Sacquin-Mora and coworkers have shown by neutron elastic temperature scans that point mutations have a significant effect on the global dynamics of the reaction-centre protein from the photosynthetic bacterium *Rhodobacter sphaeroides* and that the native protein is quantitatively more resilient and less flexible than two nonfunctional mutants (Sacquin-Mora *et al.*, 2007). Neutron experiments with specific H/D labelling have shown that the active core around the retinal binding site of bacteriorhodopsin, the light-activated proton pump in purple membranes of *Halobacterium salinarum*, is less flexible and more resilient than the protein overall, in accordance with its 'valve' function for proton translocation (Réat *et al.*, 1998; Wood, Lehnert *et al.*, 2008). Recent neutron work highlighted the difference in the dynamics of haemoglobin in hydrated powders, solutions of various concentrations and *in situ* in red

blood cells (Stadler *et al.*, 2009). The specific dynamic behaviour of haemoglobin is correlated with and may well contribute to body-temperature regulation in animals (Stadler *et al.*, 2008). These few examples illustrate how dynamics–function relationships can be measured, while underlining the importance of the exact composition and organization of the sample in a dynamics experiment if it is to be biologically relevant. A given protein is not just 'protein' and its hydration is not just 'water content'.

4. Complementarity with MD simulations

The trajectories (how far atoms move as a function of time) measured by neutron scattering correspond to values obtained in MD simulations, so that a direct comparison between calculation and experiment becomes possible. The experimental data validate the MD simulation, which consequently complements and extends the analysis from mean to site-specific dynamics. The paper by Wood, Grudinin *et al.* (2008) is an excellent example of the complementary approach. MSDs were measured as a function of temperature for bacteriorhodopsin samples in which certain amino-acid residues were hydrogenated in the otherwise fully deuterated protein, so that their motions dominated the scattering. One type of residue showed significantly large experimental MSD values. The MD simulation reproduced the result and identified the particular residue that was responsible for the large fluctuation.

5. Dynamics fingerprints: proposal for high throughput and a Dynamics Data Bank

The MSD and effective force constant can be considered as empirical parameters that are related to molecular flexibility and resilience, respectively, for a specified Q range (length scale), instrument resolution (time scale) and temperature range. The MSD and $\langle k' \rangle$ parameters provide a dynamic fingerprint of the sample on a given length and time scale.

Based on these considerations, the proposal for high-throughput experiments is to record the temperature-dependent MSD for a specified Q range and instrumental energy resolution, *i.e.* the MSD *versus* T data set measured from the elastic temperature scan for

(i) given sample conditions (powder hydration, solution, concentration, solvent composition . . .),

(ii) a given length–time window (Q range and energy resolution).

The MSD *versus* T data set provides experimental information on the dynamics of a system, which complements and can be extended by MD simulations and which, most significantly, correlates with biological function and activity.

An elastic temperature scan on the instruments at the ILL takes about 1 d for a sample of about 100 mg protein. The amount of material may seem to be prohibitive for certain applications, but there are many proteins that can be prepared in such amounts. Since incoherent scattering is examined, the sample need not be crystalline or even monodisperse, depending on the information required. The $\langle \ldots \rangle$ brackets in the MSD and $\langle k \rangle$ values mean exactly that: mean values measured for the entire population of single-particle motions in the sample and within the given length–time window, as for example in whole live cells. Experiments on whole cells do not pose sample-mass problems (Jasnin *et al.*, 2008) and they have proven to be extremely useful when the mean dynamics of the proteome was of interest, as in the case of the molecular adaptation of an organism to extreme temperature conditions (Tehei *et al.*, 2004).

In the light of the above considerations, a relatively high-throughput approach to measure protein dynamics can be envisaged. The data produced would be MSD *versus* T sets that would be deposited in a Dynamics Data Bank, together with the exact conditions under which they were measured [points (i) and (ii) above], to make them available to the MD community at large. It is important to point out that putting the approach into practice does not require the development of new instrumentation or methods. However, it does require a change in attitude of both the physics and biological communities interested in protein dynamics, and certainly better access to neutrons to encourage and support this type of experimental approach.

References

Bicout, D. J. & Zaccai, G. (2001). *Biophys. J.* **80**, 1115–1123.

Cordone, L., Ferrand, M., Vitrano, E. & Zaccai, G. (1999). *Biophys. J.* **76**, 1043–1047.

Cornicchi, E., Marconi, M., Onori, G. & Paciaroni, A. (2006). *Biophys. J.* **91**, 289–297.

Doster, W. (2010). *Biochim. Biophys. Acta*, **1804**, 3–14.

Doster, W., Cusack, S. & Petry, W. (1989). *Nature (London)*, **337**, 754–756.

Ferrand, M., Dianoux, A. J., Petry, W. & Zaccai, G. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 9668–9672.

Gabel, F. (2005). *Eur. Biophys. J.* **34**, 1–12.

Gabel, F. & Bellissent-Funel, M. C. (2007). *Biophys. J.* **92**, 4054–4063.

Gabel, F., Bicout, D., Lehnert, U., Tehei, M., Weik, M. & Zaccai, G. (2002). *Q. Rev. Biophys.* **35**, 327–367.

Glass, D. C., Krishnan, M., Nutt, D. R. & Smith, J. C. (2010). *J. Chem. Theory Comput.* **6**, 1390–1400.

Jasnin, M., Moulin, M., Haertlein, M., Zaccai, G. & Tehei, M. (2008). *Biophys. J.* **95**, 857–864.

Joti, Y., Nakagawa, H., Kataoka, M. & Kitao, A. (2008). *Biophys. J.* **94**, 4435–4443.

Kneller, G. R. (2005). *Phys. Chem. Chem. Phys.* **7**, 2641–2655.

Kneller, G. R. & Hinsen, K. (2009). *J. Chem. Phys.* **131**, 045104.

Nakagawa, H., Joti, Y., Kitao, A. & Kataoka, M. (2008). *Biophys. J.* **95**, 2916–2923.

Réat, V., Patzelt, H., Ferrand, M., Pfister, C., Oesterhelt, D. & Zaccai, G. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 4970–4975.

Réat, V., Zaccai, G., Ferrand, M. & Pfister, C. (1997). *Biological Macromolecular Dynamics*, edited by S. Cusack, H. Büttner, M. Ferrand, P. Langan & P. Timmins, pp. 117–122. Schenectady: Adenine Press.

Rosenblum, G., Van den Steen, P. E., Cohen, S. R., Grossmann, J. G., Frenkel, J., Sertchook, R., Slack, N., Strange, R. W., Opdenakker, G. & Sagi, I. (2007). *Structure*, **15**, 1227–1236.

Sacquin-Mora, S., Sebban, P., Derrien, V., Frick, B., Lavery, R. & Alba-Simionesco, C. (2007). *Biochemistry*, **46**, 14960–14968.

Stadler, A. M., Digel, I., Artmann, G. M., Embs, J. P., Zaccai, G. & Buldt, G. (2008). *Biophys. J.* **95**, 5449–5461.
Stadler, A. M., Digel, I., Embs, J. P., Unruh, T., Tehei, M., Zaccai, G., Buldt, G. & Artmann, G. M. (2009). *Biophys. J.* **96**, 5073–5081.
Tarek, M. & Tobias, D. J. (2000). *Biophys. J.* **79**, 3244–3257.
Tehei, M., Franzetti, B., Madern, D., Ginzburg, M., Ginzburg, B. Z., Giudici-Orticoni, M. T., Bruschi, M. & Zaccai, G. (2004). *EMBO Rep.* **5**, 66–70.
Tehei, M., Madern, D., Franzetti, B. & Zaccai, G. (2005). *J. Biol. Chem.* **280**, 40974–40979.
Tehei, M., Madern, D., Pfister, C. & Zaccai, G. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 14356–14361.
Teixeira, S. C. *et al.* (2008). *Chem. Phys.* **345**, 133–151.
Tobias, D. J., Sengupta, N. & Tarek, M. (2009). *Faraday Discuss.* **141**, 99–116.
Tokuriki, N. & Tawfik, D. S. (2009). *Science*, **324**, 203–207.
Tournier, A. L. & Smith, J. C. (2003). *Phys. Rev. Lett.* **91**, 208106.
Wood, K., Frolich, A., Paciaroni, A., Moulin, M., Hartlein, M., Zaccai, G., Tobias, D. J. & Weik, M. (2008). *J. Am. Chem. Soc.* **130**, 4586–4587.
Wood, K., Grudinin, S., Kessler, B., Weik, M., Johnson, M., Kneller, G. R., Oesterhelt, D. & Zaccai, G. (2008). *J. Mol. Biol.* **380**, 581–591.
Wood, K., Lehnert, U., Kessler, B., Zaccai, G. & Oesterhelt, D. (2008). *Biophys. J.* **95**, 194–202.
Zaccai, G. (2000). *Science*, **288**, 1604–1607.
Zaccai, G. (2010). *J. Non-Cryst. Solids*. doi:10.1016/j.jnoncrysol.2010.06.060.