Motional displacements in proteins: dynamical heterogeneity, origin of wave vector dependent values by D. Vural, L. Hong, J. Smith and H.R. Glyde Phys. Rev. E 91, 52705 (2015)

Comment by W. Doster, Nov. 2017

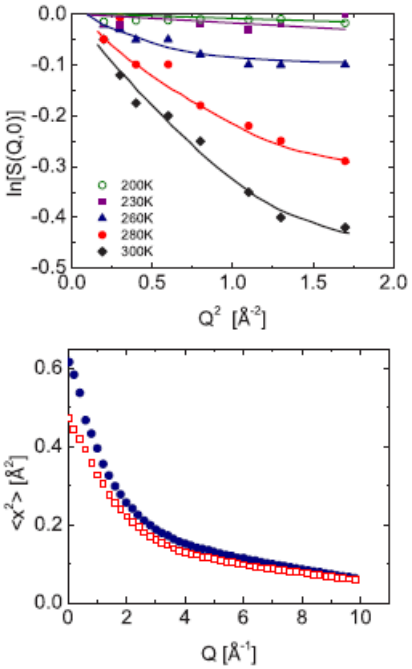


FIG. 1. (Color online) Top: Elastic component ($\omega = 0$) of the dynamical structure factor, $S(Q,\omega)$, as a function of wave vector Q of glutamate dehydrogenase observed by Daniel *et al.* [18] (also reproduced in Becker *et al.* [20]). The solid lines are a guide to the eye. Bottom: An MSD obtained from fitting to $S(Q,\omega)$ derived from MD simulation of lysozyme in Calandrini *et al.*, which shows a strong Q dependence [19]). The solid circles and open squares represent the MSD for $p = 0.1$ MPa and $p = 300$ MPa.

Elastic neutron structure factors of proteins are not linear when plotted on a log scale versus Q^2 as in fig. 1a). This means that the corresponding displacement distribution deviates from a Gaussian distribution. This was first demonstrated by Doster et al. in 1989 (Nature 337,754). Formally, as in fig. 1b, one could interpret this deviation by a Q-dependent mean square displacement MSD:

$S(Q, \omega = 0) = \exp(- Q^2 \langle x^2(Q)\rangle)$ (see eq. 1)

The "dynamical heterogeneity" concept assigns this observation to a distribution of Q-independent MSD.

$S_{DH}(Q,\omega = 0) \equiv 1/N \Sigma_i \exp(- Q^2 \langle x^2\rangle_i)$

equ. 8 with site specific MSD covering the hydrogen atoms with identical cross section.

1) The concept of dynamical heterogeneity

An MD simulation of hydrated lysozyme is analyzed aiming to demonstrate the relevance of "dynamical heterogeneity" (DH). Proteins are visualized as partially disordered systems by introducing a dominant role of a distribution of "conformational substates".(Frauenfelder et al. PNAS 111,12764, 2014) The DH concept deals with distributions of "mean square displacements" (MSD) of protein motions. MSD is not just the second moment of a displacement distribution, which requires extrapolation to zero Q. Instead one expands the neutron elastic structure factor S(Q,0) in terms of Gaussian distributions with distributed MSD (equ. 8).

The main goal is to identify the specific type of distribution , exponential, Weibull, Gamma… in this manuscript a bimodal distribution is deduced from simulations.

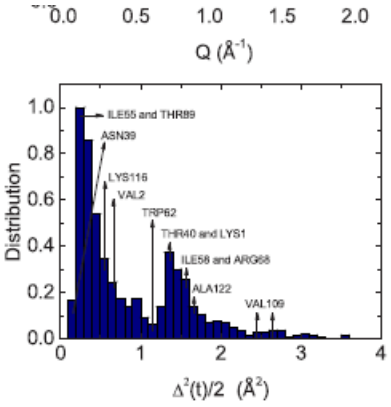


FIG. 9. (Color online) Top: The $\langle r^2 \rangle$ obtained from the fit of the model $I(Q,t)$ to the full $I_i(Q,t)$ (solid squares) and the Gaussian approximation $I_{iG}(Q,t)$ (solid circles) for individual H in ALA122 and THR40 that have large MSD and kurtosis. There is some difference in the MSD obtained from $I_i(Q,t)$ and $I_{iG}(Q,t)$ but the difference is small. Bottom: The distribution of the MSD of individual H in lysozyme at 300 K calculated using $\Delta_j^2(t)/2 = \langle (r_j(t) - r_j(0))^2 \rangle/2$ at time $t = 1$ ns from a 100-ns MD simulation.

Simulation Lysozyme

Doster et al. BBA 2005, derived by fitting **experimental** elastic scattering data of dry and hydrated myoglobin **a bi-Gaussian distribution of G(r, t = t_{res})**. The second moments of both sub-distributions were Q-independent and were assigned to two molecular processes,

(1) intrinsic non-Gaussian rotational transitions of mainly methyl groups
(2) Gaussian small scale displacements associated with interactions with hydration water.

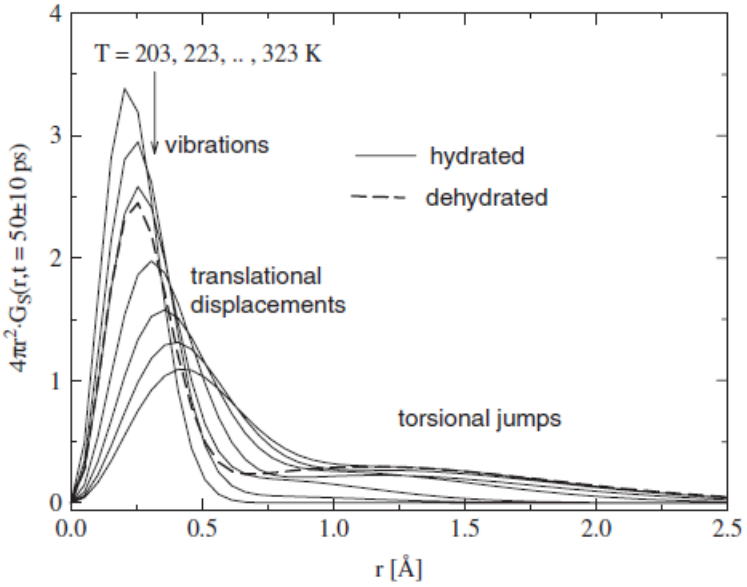


Fig. 2. Displacement distribution, $4\pi r^2 G(r, t = 50\,\text{ps})$, of hydrated (and dehydrated) myoglobin at fixed time with increasing temperature in steps of 20 degrees, derived from the fits in Fig. 1, dashed line: dehydrated myoglobin at 300 K.

Doster-Settles BBA 2005

The DH analysis by contrast is built on site heterogeneity and ignores intrinsic non-Gaussian origins. One is not interested in identifying the nature of molecular processes involved, such as rotational transitions of main chain or side chain atoms. This is quite surprising, since the elastic incoherent structure factor first of all reflects the variety of relevant molecular processes. Whether these processes have distributed parameters is a question of second order. To illustrate the discrepancy with conventional views, I focus on the comparison of experimental data with their simulation results.

2) **Intrinsic versus extrinsic origin of non-Gaussian structure factors**

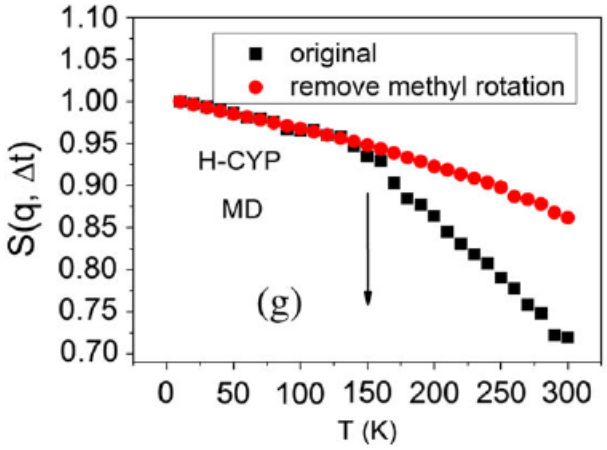
The first comprehensive analysis of non-Gaussian structure factors of proteins was performed in Doster, Settles BBA 1749, 173 (2005), "protein-water displacement distributions", which is not referenced here!!

Apart from spatial disorder, also intrinsic mechanisms exist like asymmetric motions or rotational jumps.

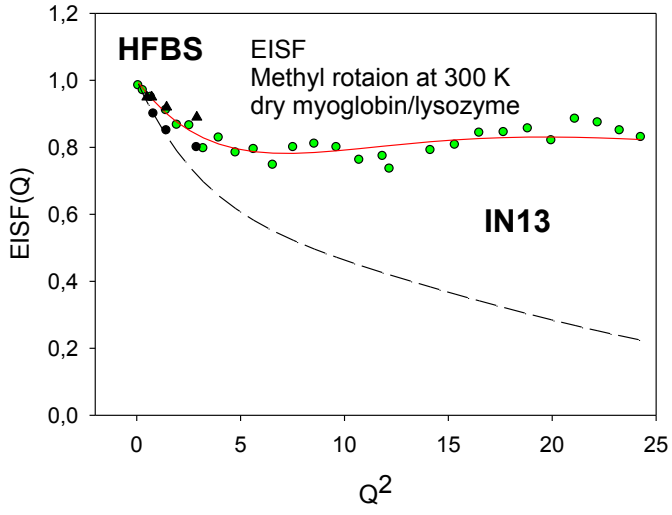
The extrapolation to zero Q for fixing MSD values was initially presented in1989, where the protein non-Gaussian structure factor was first observed . Two transition temperatures were reported, and the heterogeneity was assigned to two specific molecular processes, but no further disorder within each process was found. The proposed model structure factor assigned the Gauss-deviation fully to "rotational transitions" of side chains. Such spatially constrained motions lead to an intrinsic non-Gaussian structure factor for each atom. By contrast, already in 1991, Smith suggested site heterogeneity of proton environments as an alternative explanation (ref. Qu. Rev Biophys. 24, 227). On the contrary, Loncharich and Brooks at the same time reported in their simulation of hydrated myoglobin an important role of dihedral transitions (JMB, 215,430,1990).

The discrepancy between the two views persists up to date as evidenced by the present manuscript. An important goal of DH is therefore to prove that the local atoms obey Gaussian displacement distributions, suggesting that non-Gaussianity is a result of spatial averaging. To show this from simulations of lysozyme is the main topic here. They conclude that the fourth order cumulant , summed over all H in lysozyme is negligible, implying a local Gaussian distribution for each H. In Doster et al. BBA 2005 we have published non-Gaussianity factors which indicate asymmetric displacement distributions.

Since our" very detailed analysis" in 2005, it is well established that rotational transitions, specifically those associated with methyl groups, contribute strongly to the neutron elastic scattering function of proteins.



This fact is now fully accepted by MD theorists, but it was first demonstrated experimentally (Doster et al. BBA 2005). The figure above, taken from a recent paper of the Smith group, shows the elastic scattering function simulated for lysozyme at low hydration with and without taking into account methyl group rotational transitions (Liu et al. PRL 119, July 2017). This striking difference in the T-dependence affects also the Q-dependence, which must lead to an intrinsic non-Gaussian structure factor. This result is in my view inconsistent with the simulation results presented here. In Doster et al. BBA 2005,, the role of the intrinsic non-Gaussian structure factor has been examined, 90% of the Gauss deviation could be assigned to rotation, see figure below. Shown is also the Gaussian component and the data by Roh et al (BJ 2006 96, 2755) of lysozyme. The Roh data cover only a narrow Q-range of HFBS, but are mostly cited as main evidence of methyl group rotation.



Doster, Eur. Biophys.J.(2008) 37,591, fig. 5.

Peters and Kneller (ref. 29, JCP 2013 139, 165102) analyse experimental data with the DH model using a Gamma distribution. They relate their results to the real world of molecular motion without citing the source. They wrote:

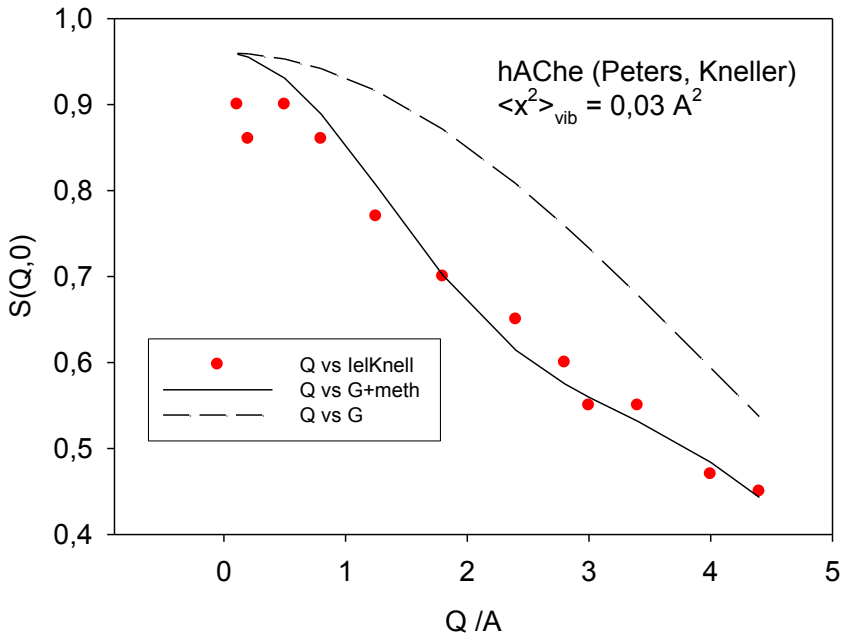
"*Applied to the neutron scattering data from human acetylcholinesterase, it reveals a strong increase of the motional heterogeneity at the two transition temperatures T = 150 K and T =*

220 K, respectively, which can be located with less ambiguity than with the Gaussian model. We find that the first transition is essentially characterized by a change in the form of the elastic scattering profile and the second by a homogeneous increase of all motional amplitudes. These results are in agreement with previous combined experimental and simulation studies of protein dynamics, which attribute the first transition to an onset of methyl rotations and the second to more unspecific diffusion processes involving large amplitude motions" (Peters, Kneller, JCP 2013)

I also performed fits of their data presented in fig. 2 with the rotational model introduced above: A Gaussian Debye-Waller factor combined with the methyl structure factor, and with the known 28% contribution of methyl groups + dihedral transitions to the cross section of proteins as input. Only a single parameter, the vibrational displacements was adjusted.
At 220 K other components are small.

$$EISF(Q) = \exp(-Q^2\langle x^2_{vib}\rangle) \cdot (1- 0{,}66\cdot f_{meth}(1-j_0(Qr\sqrt{3}))) \qquad \text{equ. 1}$$

$\langle x^2_{vib}\rangle = 0{,}03\ A^2$, 28 % $f_{met} = 0{,}25$, r = 1 A. T = 220 K Peters et al. fig. 2



Full line: equ. 1 adjusted to data, dashed: Gaussian component
Considering the error in the data of such experiments, a single adjustable parameter, the fit looks reasonable, suggesting that an alternative approach different from DH is possible.

Technical Remarks

Incorrect basic equation

Equ. 1 defines the "observed incoherent dynamic structure factor" incompletely:

$$S(Q,\omega=0) = AI_N(Q, t = \infty) \qquad \text{equ. 2}$$

The reason is explained in Doster et al. JCP 139, 45105, 2013: Even at zero frequency, quasi-elastic scattering contributes increasingly with temperature to the elastic line, such that the contributions from real elastic scattering seemingly decrease. This can lead to strange variations of the apparent mean square displacements (MSD) like an apparent decrease with temperature as shown by Peters and Kneller (2013). Moreover, the authors have observed such an up to now unexplained MSD decrease themselves (ref.3, 18). By contrast the right hand side of equ. 1, which is usually calculated by simulations, is not directly affected by zero frequency components, but depends on the simulation time, which is slightly less than infinity. It is sloppy to use infinity instead.

Tentative Conclusions:

This paper presents a too narrow explanation of non-Gaussian scattering functions of proteins. The relevant literature is not cited. Only promoting publications are mentioned. Alternative concepts are not even considered. Rotational transitions are ignored. A major part of the simulations is devoted to showing that the local displacement distribution is Gaussian. The non- Gaussian average structure factor was assigned to site heterogeneity. This seems to be inconsistent with experimental results and their own simulations. Formally one can always expand the elastic structure factor in terms of Gaussian distributions, as was done by us in 2005 deriving a bimodal distribution as the "minimal model". But this does not imply, that only spatial disorder, which certainly exists, contributes to the observed non-Gaussian structure factor. The Gaussian force constant model by Zaccai and Bicout, ref. 22,23 is cited in support of DH. These authors explain the increase of MSD at high temperature as a softening of harmonic force constants above a certain temperature. In Doster (concepts and misconceptions of the protein dynamical transition) Eur Bioph. J. 2008, 37 591 I argue that this type of molecular process is extremely unlikely. Rotational transitions between two or three states cannot be accounted for by softening of force constants. *Not to understand the methyl rotation in BR is the main reason why almost all conclusions in references [5], [22] and [23] in this publication are incorrect.*

Also the Japanese DH group, Mikio Kataoka and collaborators employ a bimodal distribution and claim that "DH dominantly contributes to non-Gaussian scattering. In their paper , Nakagawa et al 2006 Mat. Sci. Eng. A 356 and BBA 1804 (2010) 27 rotational transition are not considered. In 2010 they briefly write about our work:

*The non-Gaussianity was analyzed by rotational motions of side chains (in 1989,*which is still correct today). *On the other hand, the same authors interpreted (*in 2005, BBA) *the non-Gaussianity by a bimodal distribution (torsion +water-induced) with the same data (* the data are different, dehydrated myoglobin). *We analysed the non-Gaussianity with the DH model.* (in 2004, but assuming an exponential distribution). Kataoka continues: *The scattering data can be fitted with some different distribution models, Gaussian,*

exponential, bimodal, even two site jump. These distributions appear arbitrary without molecular interpretation. Why Kataoka and his DH coworkers refuse to include rotational transitions, remains a mystery. In our 2005 analysis we use a sum of two Gaussians as a formal expansion of the elastic structure factor to separate different kinds of motion. From this minimal treatment, we derive the MSD of the sub-distributions. In contrast to the DH approach, we do not exclude intrinsic non-Gaussianity. The DH concept ows a lot to Hans Frauenfelder and his distribution of conformational substates. Frauenfelder became so frustrated with the neutron community that, quite recently, he suggested a completely new scattering theory with inhomogeneous quasi-elastic spectra. This will be discussed in the next contributions. Since most workers in this bio-neutron scattering field are more religious than competent, I am certain, that most of them will adopt the HF model without critical review. Then DH is no longer needed.