

## de Gennes Narrowing Describes the Relative Motion of Protein Domains

Liang Hong,<sup>1,2</sup> Nikolai Smolin,<sup>1,3</sup> and Jeremy C. Smith<sup>1,2,\*</sup>

<sup>1</sup>Center for Molecular Biophysics, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA

<sup>2</sup>Department of Biochemistry and Cellular and Molecular Biology, University of Tennessee, Knoxville, Tennessee 37996, USA

<sup>3</sup>Department of Cell and Molecular Physiology, Loyola University Chicago, Maywood, Illinois 60153, USA

(Received 5 February 2014; published 18 April 2014)

The relative motion of structural domains is essential for the biological function of many proteins. Here, by analyzing neutron scattering data and performing molecular dynamics simulations, we find that interdomain motion in several proteins obeys the principle of de Gennes narrowing, in which the wave vector dependence of the interdomain diffusion coefficient is inversely proportional to the interdomain structure factor. Thus, the rate of interdomain motion is inversely proportional to the probability distribution of the spatial configurations of domains.

DOI: 10.1103/PhysRevLett.112.158102

PACS numbers: 87.15.H-, 87.14.E-, 87.15.ap, 87.64.Bx

Interdomain motion is essential for the functioning of many multidomain proteins, modulating the entry of substrates into active sites [1], allosteric protein function [2], and enzymatic reactions [3]. In physiological conditions, these functional dynamic modes typically occur on a length scale of a few nm and time scales of ns –  $\mu$ s or longer [4–6]. Motions on these length and time scales are accessible to neutron spin echo spectroscopy (NSE), and this technique has recently been used to probe relative domain motions in several proteins [4,5,7,8]. NSE provides direct information on the key spatiotemporal property of domain motion—the distance dependence of the relative motion of protein domains—characterized by the wave vector ( $q$ ) dependence of the effective diffusion coefficient,  $D_{\text{eff}}(q)$  [4,5,7].

The present Letter demonstrates that  $D_{\text{eff}}(q)$  of the domain motion in several proteins is inversely proportional to the interdomain structure factor,  $S(q)$ .  $S(q)$  characterizes the equilibrium spatial arrangement of the constituent domains and can be determined using small angle neutron or x-ray scattering (SANS/SAXS). The inverse dependence is qualitatively described by de Gennes narrowing (DGN) theory [9,10], which posits an inverse relationship between the probability of a spatial configuration occurring and the fluctuation rate of the configuration, i.e., a less populated configuration fluctuates faster. Furthermore, molecular dynamics (MD) simulation of phosphoglycerate kinase (PGK) reveals that interdomain motion obeys overdamped Langevin dynamics with the velocity determined by the curvature of the underlying free energy profile, consistent with DGN.

We focus here on the dynamics of PGK but show that the principle also applies to other proteins. As seen in Fig. 1(a), PGK consists of two well-folded domains linked by a single helix. Hinge-bending motion between the two domains in PGK plays an important role in its function of transferring a phosphate group between ligands [3].

Details of the MD simulation protocol and the treatment of neutron scattering experimental data are described in the Supplemental Material [11].

The quantity measured in NSE experiments is the coherent intermediate scattering function,

$$I_{\text{coh}}(q, t) = \sum_i^Z \sum_j^Z b_i b_j \langle \exp[-i\vec{q} \cdot \vec{R}_i(0)] \exp[i\vec{q} \cdot \vec{R}_j(t)] \rangle, \quad (1)$$

where  $Z$  is the total number of atoms,  $b_j$  is the coherent scattering length, and  $\vec{R}_j$  is the position of atom  $j$ , while the brackets denote averaging over the time origin and the orientation of  $\vec{q}$ .  $D_{\text{eff}}(q)$  is obtained from the initial decay rate of the NSE-determined coherent intermediate scattering function and is defined as  $-(1/q^2) \lim_{t \rightarrow 0} (d/dt) \ln[I_{\text{coh}}(q, t)/I_{\text{coh}}(q, 0)]$ .

After taking into account the interprotein scattering and hydrodynamic interactions,  $D_{\text{eff}}(q)$  of a single protein molecule can be derived [4,5,7]; this is shown for PGK in Fig. S1, digitized from Ref. [4].  $D_{\text{eff}}(q)$ , thus obtained, contains both the protein internal motion, denoted as  $D_{\text{inter}}(q)$ , and the global rotation and translation of the whole macromolecule in water solution, denoted as  $D_{\text{glob}}(q)$ .  $D_{\text{glob}}(q)$  can be calculated using a hydrodynamic model based on the atomic-detail structure of the protein with the program HYDROPRO (see details in Ref. [4] and Fig. S1).  $D_{\text{eff}}(q) - D_{\text{glob}}(q)$  represents the protein internal dynamics only,  $D_{\text{inter}}(q)$ , plotted for PGK in Fig. 1(b).

The interdomain structure factor of the protein molecule,  $S(q)$  [Fig. 1(c)] is defined as,

$$S(q) = \frac{I(q)}{\sum F(q)}, \quad (2)$$

where  $I(q)$  is the form factor of the whole protein molecule and can be readily obtained from small angle scattering

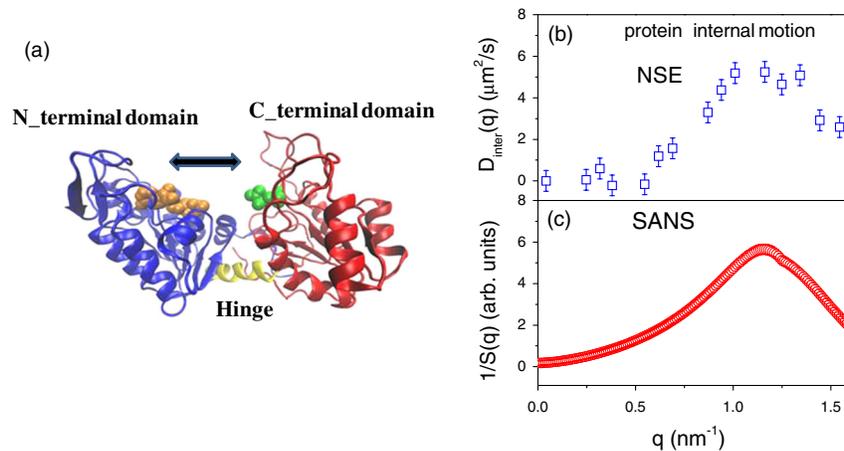


Fig. 1 (color online).  $D_{\text{inter}}(q)$  versus  $S(q)$  for PGK derived from neutron scattering experiments. (a) Structure of PGK consisting of two domains (left and right) of almost equal mass connected by a helix (middle). (b)  $D_{\text{inter}}(q)$  of PGK in aqueous solution (see the Supplemental Material [11] and Fig. S1). The value at the lowest  $q$  was measured by dynamic light scattering while all the other data were collected using NSE ([4]). (c) Interdomain structure factor. The form factor of the whole protein molecule was measured by SANS [4] and that for each single domain was calculated based on the crystal structure (PDB entry 3PGK [20]).

experiments on dilute protein solutions, and  $F(q)$  is the form factor of each single domain [21].  $S(q)$  describes the relative spatial arrangement of domains in the protein.

An intriguing finding here is that  $D_{\text{inter}}(q)$  exhibits an inverse  $q$  dependence on  $S(q)$  (Fig. 1). Moreover, a similar inverse  $q$  dependence between  $D_{\text{inter}}(q)$  and  $S(q)$  is also found for Taq polymerase and alcohol dehydrogenase, as shown in Figs. S2 and S3 (Supplemental Material [11]) using data from Refs. [4,7], suggesting that this relationship may be widely applicable.

To further examine  $D_{\text{inter}}(q)$  and  $S(q)$ , we performed an MD simulation of PGK.  $D_{\text{inter}}(q)$  derived from the MD trajectory is presented in Fig. 2(a), and the  $q$  dependence is indeed found to be the inverse of that of the corresponding, MD-derived  $S(q)$  [Fig. 2(b)]. To determine the contribution of interdomain motion to  $D_{\text{inter}}(q)$ , we post processed the MD trajectory so as to subtract the intradomain atomic fluctuations (see details in Fig. 2 legend). As seen in Fig. 2(a),  $D_{\text{inter}}(q)$  of the protein internal dynamics is dominated by the interdomain motion except in the high  $q$  region ( $q > 2.2 \text{ nm}^{-1}$ ), where the intradomain fluctuations start to contribute significantly. Hence, the internal dynamics in the range  $q \leq 2 \text{ nm}^{-1}$  as probed by NSE [Fig. 1(b)] is mostly large-scale, interdomain motion.

As shown in Figs. 1, 2, S2, and S3, both the experimental and MD simulation results show that  $D_{\text{inter}}(q)$  presents an inverse  $q$  dependence on  $S(q)$ . In Refs. [7,8], a formula containing  $S(q)$  in the denominator was used to fit the full NSE signal (i.e., including both the global and internal motion) on multidomain proteins. However, the physics behind the inverse relationship was not discussed. Moreover,  $D_{\text{eff}}(q)$  resulting from the global rotation of a multisubunit system can exhibit a very different  $q$  dependence than the internal dynamics, while the inverse

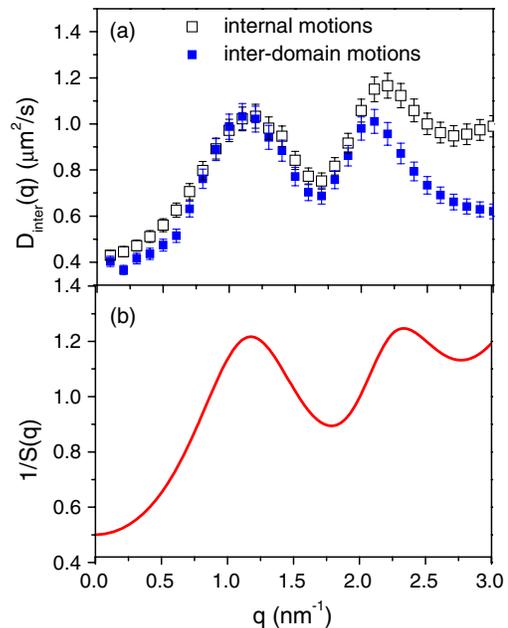


Fig. 2 (color online).  $D_{\text{inter}}(q)$  versus  $S(q)$  for PGK derived from MD simulations. (a) Effective diffusion coefficient,  $D_{\text{inter}}(q)$ . The empty squares denote the protein internal dynamics derived from the trajectory obtained by superimposing the MD trajectories of the protein molecule onto a reference structure.  $D_{\text{inter}}(q)$  with the intradomain fluctuations subtracted, i.e., representing the interdomain motion only, is plotted as solid squares and was obtained by further fitting the structure of the two domains [blue and red in Fig. 1(a)] in the reference state independently to the corresponding domains in each frame of the superimposed trajectory. (b) Interdomain structure factor.

relationship,  $D_{\text{eff}}(q) \sim 1/S(q)$ , is valid only for the latter. To illustrate such a difference, we construct two models. Model A: system consists of two identical spheres with 0.5 nm radius, each performing a three-dimensional random walk within a limited volume of 2 nm radius [see Fig. 3(a)]. Model B: in addition to the motion described in model A, the system performs an additional global rotation at a constant speed around an axis at the midpoint of the two spheres [see Fig. 3(b)]. Model A represents the internal motion in a multisubunit system, while model B corresponds to global rotation plus internal dynamics. The resulting  $D_{\text{eff}}(q)$  and  $S(q)$  are presented in Figs. 3(c) and 3(d). The  $q$  dependence of  $D_{\text{eff}}(q)$  calculated from model B differs significantly from that of model A [Fig. 3(c)] due to the addition of the global rotation, while  $1/S(q)$  corresponds exactly to  $D_{\text{eff}}(q)$  of model A [Fig. 3(d)]. Hence, the inverse relationship between  $D_{\text{eff}}(q)$  and  $1/S(q)$  is valid only for internal motion.

$D_{\text{eff}}(q) \sim 1/S(q)$  has been observed in many nonbiological systems, ranging from liquid argon [22] and deuterated methane [23] through to colloidal particles [24–26]. The inverse dependence can be qualitatively explained by de Gennes narrowing (DGN) [9], the underlying physics of which is schematically illustrated in Fig. S4. Assume that two particles, separated by a distance  $r$ , move relative to each other, and that the associated distance distribution,  $P(r)$ , peaks at  $r^*$  [Fig. S4(a)]. Fourier transform of  $P(r)$  from real to reciprocal space results in the interparticle structure factor,  $S(q)$  peaking at  $q^*$  ( $q^* \sim 1/r^*$ ) [Fig. S4(b)]. Accordingly, the underlying free energy profile exhibits a

minimum at  $r^*$  [Fig. S4(c)]. DGN predicts that the rate of the interparticle fluctuation is a minimum when the population of the configuration is a maximum, i.e., at  $r^*$  [9,10]. As  $D_{\text{inter}}(q)$  is an estimate of the interparticle fluctuation rate at the distance  $r$  ( $r \sim 1/q$ ) [9,10], it will exhibit a minimum at  $q^*$  ( $q^* \sim 1/r^*$ ) [Fig. S4(d)] where the fluctuation rate reaches a minimum, and, thus, resembles  $1/S(q)$ .

We now use the MD simulation on PGK to further examine the characteristics of the dynamics leading to the DGN. Figure 4(a) presents the distribution  $P(r)$ , of the distance,  $r$ , between the centers of mass of the two domains in PGK and the corresponding potential of mean force,  $U(r) = -RT \ln P(r)$ , where  $R$  is the gas constant. The peak in  $P(r)$ , and minimum in  $U(r)$ , are at  $q^* = 37.7 \text{ \AA}$ . The time-averaged interdomain velocity,  $\langle v(r) \rangle = \langle dr/dt|_{t=0} \rangle$ , shown in Fig. 4(b), exhibits a clear minimum at  $r^*$ , consistent with the DGN picture in which the interdomain fluctuation is slowest when the system assumes the most probable configuration. Furthermore, the shape of  $v(r)$  superimposes closely that of  $dU(r)/dr$  [Fig. 4(b)].

Figure 4(b) can be understood in the framework of Langevin dynamics. Assume that the protein interdomain motion follows the classical Langevin equation, i.e.,

$$m\ddot{r} + \zeta v(r) + \frac{\partial V(r)}{\partial r} = R(t), \quad (3)$$

in which  $m$  is the mass of the domain,  $\zeta$  is the friction coefficient,  $R(t)$  is the random force,  $\ddot{r}$  is the acceleration,

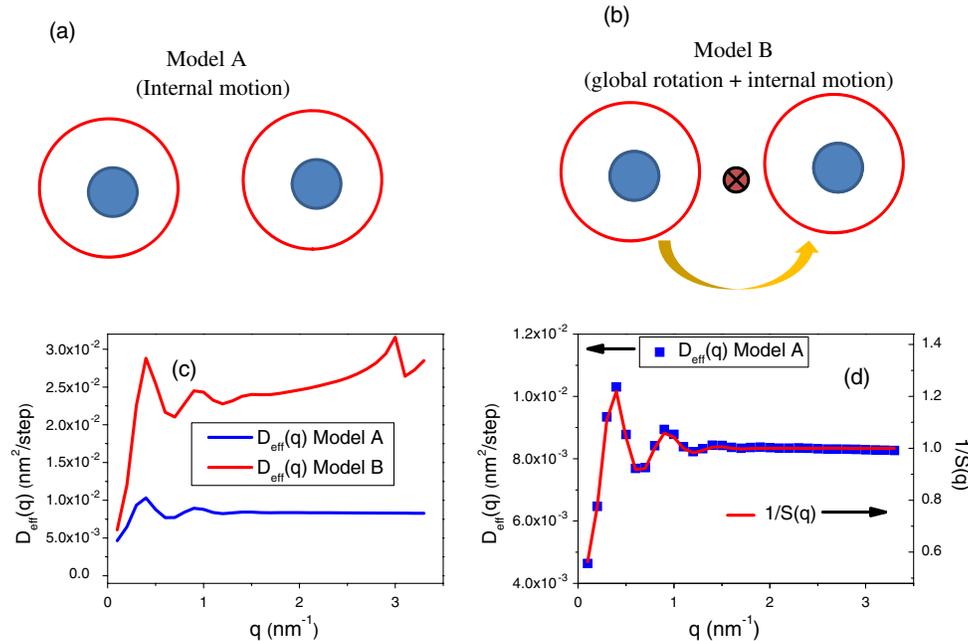


Fig. 3 (color online). Schematic illustration of (a) model A and (b) model B. (c)  $D_{\text{eff}}(q)$  calculated from models A and B. (d)  $D_{\text{eff}}(q)$  versus  $1/S(q)$  derived from model A.  $S(q)$  calculated from the two models are identical as global rotation does not alter the relative position of the two spheres and has, thus, no effect on the structure factor.

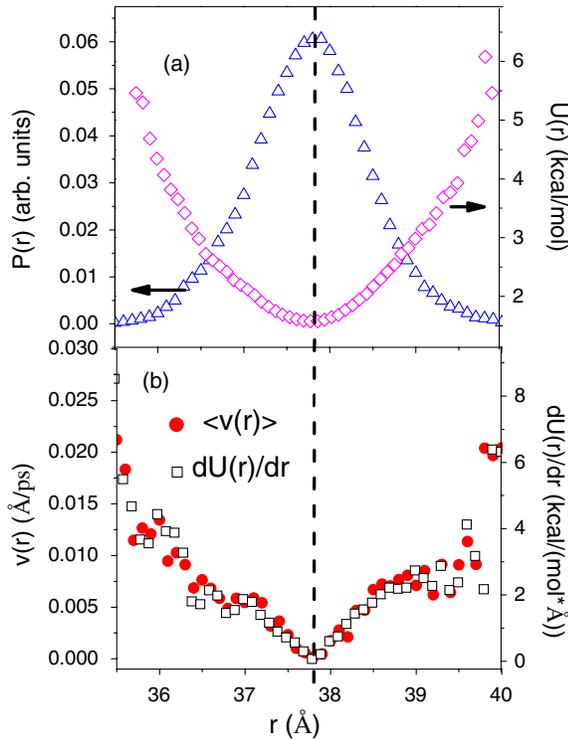


Fig. 4 (color online). Test of DGN using analysis of MD trajectories of PGK. (a) Distribution of interdomain distances and the corresponding potential of mean force. (b) Time-averaged velocity of the interdomain motion and the curvature of the potential of mean force. The position of  $r^*$  as noted in the text is marked by a dashed line. Figure 4(b) displays the absolute values.

and  $V(r)$  is the potential energy. By performing a time average of both sides of Eq. (3) and assuming overdamped dynamics [27], the acceleration and random force can be neglected and Eq. (3) reduces to

$$\langle v(r) \rangle = -\frac{1}{\zeta} \frac{\partial V(r)}{\partial r}, \quad (4)$$

corresponding exactly to Fig. 4(b).

Using Stokes' law,  $\zeta = 6\pi\eta a$ , where  $a$  is the radius of a single domain of PGK ( $\sim 2.5$  nm) and  $\eta$  is the viscosity,  $\eta$  for the protein interdomain motion can be derived based on Eq. (4) using the results of  $\partial U(r)/\partial r$  and  $\langle v(r) \rangle$  in Fig. 4(b), and is found to be  $5.16 \times 10^{-3}$  Ns/m<sup>2</sup>, four times larger than the value of bulk water at the temperature studied. This relatively high viscosity may arise from the fact that the protein molecule is hydrated by surface water molecules that are  $\sim 15\%$  more densely packed [28] with a self-diffusion coefficient of  $\sim 5$  to 10 times lower than that of bulk water [29,30].

Previously, the effective diffusion coefficient for internal motions in alcohol dehydrogenase [5] and PGK [4] measured by NSE was modeled using coarse-grained normal modes. The  $q$  dependence of  $D_{\text{inter}}$  derived from

overdamped low-frequency normal modes was found to resemble the shape of the NSE experimental data for these proteins [e.g., see Figs. S3(b) and S3(c)] [4,5]. The observation in the present Letter that protein interdomain motion obeys overdamped Langevin dynamics [Fig. 4(b)] furnishes a justification for the coarse-grained normal-mode approach, provided the correct potential and frictional damping are used [Eq. (3)]. The frequencies of the lowest-frequency normal modes in PGK were calculated to be  $\sim 5$  cm<sup>-1</sup> [31], consistent with the value estimated from the potential of mean force presented in Fig. 4(a) (a harmonic approximation,  $1/2kr^2$ , of  $U(r)$  yields a force constant,  $k = 1.2$  kcal/(mol\*Å<sup>2</sup>), corresponding to a harmonic vibrational frequency of  $\sim 6$  cm<sup>-1</sup>). Hence, the potential obtained from coarse-grained normal-mode approach is in quantitative agreement with that derived from the present all-atom MD simulation.

The present analysis demonstrates that protein interdomain motion follows the principle of de Gennes narrowing, i.e., that the initial decay rate of the coherent intermediate scattering function possesses an inverse  $q$  dependence on the interdomain structure factor. This can be simply understood as the domains moving slower with respect to each other when in a favored spatial arrangements. This interpretation is consistent with an MD analysis showing that for PGK the velocity of the overdamped interdomain motion follows the shape of the potential of mean force; i.e., the interdomain velocity is determined by the curvature of the protein energy landscape. de Gennes narrowing does not necessarily hold for time scales longer than those determining the initial decay rate of the coherent intermediate scattering function, and deviations from the narrowing principle will be of interest for future studies. Notwithstanding, the simple principle found in the present Letter may well be of general aid in our understanding of the relationship between the structure, dynamics, and energy landscapes of large-scale biological macromolecular systems.

We are grateful to Dr. Ralf Biehl and Dr. Dieter Richter for useful discussions and for furnishing the experimental SANS and NSE data of PGK. This research was funded in part from an EpsCOR grant from the U.S. Department of Energy. This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

\*smithjc@ornl.gov

- [1] A. Johs, I. M. Harwood, J. M. Parks, R. E. Nauss, J. C. Smith, L. Liang, and S. M. Miller, *J. Mol. Biol.* **413**, 639 (2011).
- [2] G. S. Martin, *Nat. Rev. Mol. Cell Biol.* **2**, 467 (2001).

- [3] R. D. Banks, C. C. F. Blake, P. R. Evans, R. Haser, D. W. Rice, G. W. Hardy, M. Merrett, and A. W. Phillips, *Nature (London)* **279**, 773 (1979).
- [4] R. Inoue, R. Biehl, T. Rosenkranz, J. Fitter, M. Monkenbusch, A. Radulescu, B. Farago, and D. Richter, *Biophys. J.* **99**, 2309 (2010).
- [5] R. Biehl, B. Hoffmann, M. Monkenbusch, P. Falus, S. Préost, R. Merkel, and D. Richter, *Phys. Rev. Lett.* **101**, 138102 (2008).
- [6] N. Smolin, R. Biehl, G. R. Kneller, D. Richter, and J. C. Smith, *Biophys. J.* **102**, 1108 (2012).
- [7] Z. M. Bu, R. Biehl, M. Monkenbusch, D. Richter, and D. J. E. Callaway, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 17646 (2005).
- [8] B. Farago, J. Li, G. Cornilescu, D. J. E. Callaway, and Z. Bu, *Biophys. J.* **99**, 3473 (2010).
- [9] P. G. de Gennes, *Physica (Amsterdam)* **25**, 825 (1959).
- [10] B. J. Ackerson, P. N. Pusey, and R. J. A. Tough, *J. Chem. Phys.* **76**, 1279 (1982).
- [11] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.112.158102> describes simulation protocol and detailed analysis of the experimental data, and includes Refs. [12–19].
- [12] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl, *J. Chem. Theory Comput.* **4**, 435 (2008).
- [13] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, *J. Comput. Chem.* **4**, 187 (1983).
- [14] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, *J. Chem. Phys.* **79**, 926 (1983).
- [15] T. Darden, D. York, and L. Pedersen, *J. Chem. Phys.* **98**, 10089 (1993).
- [16] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, *J. Chem. Phys.* **81**, 3684 (1984).
- [17] W. G. Hoover, *Phys. Rev. A* **31**, 1695 (1985).
- [18] M. Parrinello and A. Rahman, *J. Appl. Phys.* **52**, 7182 (1981).
- [19] Y. Kim, S. Hyun Eom, J. Wang, D.-S. Lee, S. Won Suh, and T. A. Steitz, *Nature (London)* **376**, 612 (1995).
- [20] H. C. Watson *et al.*, *EMBO J.* **1**, 1635 (1982).
- [21] D. I. Svergun and M. H. J. Koch, *Rep. Prog. Phys.* **66**, 1735 (2003).
- [22] K. Skold, *Phys. Rev. Lett.* **19**, 1023 (1967).
- [23] G. Venkatar, Ba. Dasannac, and K. R. Rao, *Phys. Lett.* **23**, 226 (1966).
- [24] J. C. Brown, P. N. Pusey, J. W. Goodwin, and R. H. Ottewill, *J. Phys. A* **8**, 664 (1975).
- [25] D. W. Schaefer and B. J. Berne, *Phys. Rev. Lett.* **32**, 1110 (1974).
- [26] J. K. Phalakornkul, A. Gast, R. Pecora, G. Nägele, A. Ferrante, B. Mandl-Steininger, and R. Klein, *Phys. Rev. E* **54**, 661 (1996).
- [27] W. Min, G. Luo, B. J. Cherayil, S. C. Kou, and X. S. Xie, *Phys. Rev. Lett.* **94**, 198302 (2005).
- [28] J. C. Smith and F. Merzel, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 5378 (2002).
- [29] J. D. Nickels *et al.*, *Biophys. J.* **103**, 1566 (2012).
- [30] F. Mallamace *et al.*, *J. Chem. Phys.* **127**, 045104 (2007).
- [31] C. Guilbert, F. Pecorari, D. Perahia, and L. Mouawad, *Chem. Phys.* **204**, 327 (1996).