Lattice Dynamics of a Protein Crystal

Lars Meinhold,[1,*,†] Franci Merzel,[2] and Jeremy C. Smith[1,3]

[1]*Computational Molecular Biophysics, Interdisciplinary Center for Scientific Computing (IWR), University of Heidelberg, Im Neuenheimer Feld 368, D-69120 Heidelberg, Germany*

[2]*Laboratory for Molecular Modeling and NMR, National Institute of Chemistry, Hajdrihova 19, 1000 Ljubljana, Slovenia*

[3]*University of Tennessee/Oak Ridge National Laboratory Center for Molecular Biophysics, Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, Tennessee, 37831-6164, USA*

(Received 5 January 2007; published 25 September 2007)

All-atom lattice-dynamical calculations are reported for a crystalline protein, ribonuclease A. The sound velocities, density of states, heat capacity (C_V) and thermal diffuse scattering are all consistent with available experimental data. $C_V \propto T^{1.68}$ for $T < 35$ K, significantly deviating from a Debye solid. In Bragg peak vicinity, inelastic scattering of x rays by phonons is found to originate from acoustic mode scattering. The results suggest an approach to protein crystal physics combining all-atom lattice-dynamical calculations with experiments on next-generation neutron sources.

DOI: 10.1103/PhysRevLett.99.138101 PACS numbers: 87.15.−v, 63.20.Dj, 87.14.Ee, 87.64.Bx

X-ray protein crystallography has become a cornerstone of molecular biophysics and has been used to determine more than 33 500 protein crystal structures [1]. However, comprehension of protein function requires that structural coordinates be complemented by dynamical information. Crystalline protein systems are potentially rich sources of information on both protein:protein interactions and internal protein dynamics. Interprotein crystalline interactions give rise to lattice dynamics, which can be calculated using the harmonic approximation to the intermolecular potential function. Lattice dynamics reduces the translational crystal symmetry, giving rise to thermal diffuse scattering (TDS) in the diffraction pattern. This inelastic TDS provides direct information on correlated motions, which cannot be inferred from the elastic Bragg scattering [2–4].

The vast majority of computational protein biophysics has been performed using extensively developed atomic-detail potential functions. However, due to computational limitations, the large size of protein unit cells (containing the order of 10^3–10^5 atoms) has hitherto precluded lattice-dynamical calculations using these potential functions. In this Letter, the first all-atom lattice-dynamical calculation for a crystalline globular protein is reported.

Bovine pancreatic ribonuclease A (RNase) was chosen for the study due to its relatively small unit cell, the existence of a high-resolution diffraction analysis with all hydrogen atoms being resolved [5], and the existence of inelastic x-ray scattering data in the vicinity of Bragg peaks [6], which is likely to originate from scattering by lattice phonons. The vibrational density of states, sound velocities, and specific heat capacities are also derived and compared with experiment. In cases where experimental data for RNase are lacking comparison is made with other globular proteins that share similar physical and thermodynamical properties [7].

All calculations were performed on the monoclinic RNase crystal, with unit cell dimensions $a = 30.18$ Å, $b = 38.40$ Å, $c = 53.32$ Å, and $\beta = 105.85°$, composed of two protein molecules in the space-group symmetry $P2_1$, 902 water molecules and 16 chloride counter ions resulting in an electrically neutral system, in total comprising 6442 atoms. Atomic interactions were calculated using the CHARMM molecular mechanics force field and parameter set 22 [8,9] with the water molecules represented using the three-point transferable intermolecular potential function [10]. Electrostatic and van der Waals interactions were truncated at 13 Å using smooth switching (between 10–13 Å) and shifting functions, respectively. The system was energy-minimized using a scheme of alternating cycles of heating and minimization, reaching a final gradient $<4.2 \times 10^{-4}$ J mol^{-1} Å^{-1}.

In the harmonic approximation, the equations of motion for the atomic displacements from their equilibrium positions can be decoupled using the plane wave approach [11,12], yielding eigenmodes $\mathbf{u}(\mathbf{q})$ and associated eigenfrequencies $\boldsymbol{\omega}(\mathbf{q})$, commonly referred to as the dispersion relations. In Fig. 1 are shown $\boldsymbol{\omega}(\mathbf{q})$ for the RNase crystal along the principal crystal directions. Apart from the acoustic and lowest-frequency optic modes, the phonon dispersion curves show little dependence on the wave vector, indicating that most unit-cell, and thus protein, internal vibrations are independent of the crystal environment. In contrast, strong dispersion has been found for some molecular crystals [13–15]. However, the relative strength of crystal contacts in protein systems is weaker than, for example, the hydrogen bonds in these molecular crystals and, therefore, the higher-frequency optic modes for proteins may, in general, exhibit relatively little dispersion. For each acoustic branch the mixing of longitudinal (LA) and transverse (TA) character was computed using the projection $\xi(\mathbf{q}) = 1 - \langle \mathbf{u}'_\alpha \cdot \mathbf{q}' \rangle$ at $q_{hkl} = 0.1$, where $\langle \cdot \rangle$ denotes the average over all protein atoms and the prime indicates that the vectors are normalized to unity: these results are also included in Fig. 1. Along the principal crystal directions the mixing is less than 3%.

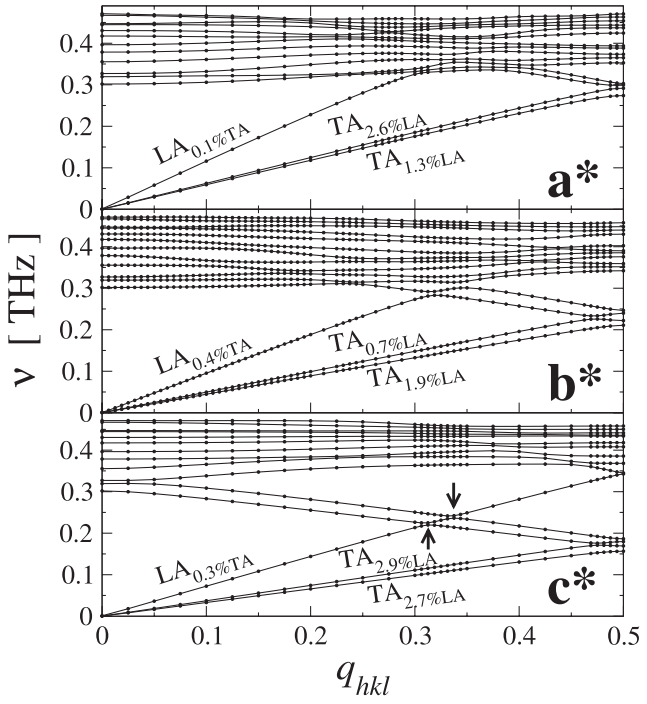

FIG. 1. Dispersion relations for the acoustic [predominantly transverse (TA) or longitudinal (LA); subscripts indicate TA-LA mixing at $q_{hkl} = 0.1$] and 12 lowest-frequency optical modes. Symbols are data points with connecting lines drawn for convenience. $q_{hkl} = q(2\pi/d_{hkl})^{-1}$, with d_{hkl} being the spacing of the real lattice. $\nu = \omega/2\pi$ is the true frequency. Arrows indicate anticrossings discussed in the text.

The arrows in Fig. 1 indicate selected regions of anticrossing in the dispersion relations. This phenomenon has been observed in a variety of systems, ranging from comparatively simple ionic crystals to more complex semiconductor heterostructures and molecular crystals like *L*-alanine [13,14] and naphthalene [15]. When two modes anticross, their associated eigenvectors can mix. As an example of this, at $q_{hkl} \approx 0.3$ the mode marked LA along $\mathbf{c}^*$ in Fig. 1 exchanges its LA character with the lowest-frequency optical mode, which inclines from above and subsequently transmits this LA character further to the optical mode with the next-higher frequency.

Figure 1 also shows that, along $\mathbf{a}^*$, $\mathbf{b}^*$, and $\mathbf{c}^*$, the TA branches are approximately linear over the whole Brillouin zone and, at the zone boundary, for each of these branches there exists an associated optic branch. This behavior would be expected for a system with a Brillouin zone twice as large [16]. Along $\mathbf{b}^*$ and $\mathbf{c}^*$ this zone doubling is due to the twofold screw axis. Wave vectors along $\mathbf{b}^*$ are insensitive to rotations around the screw axis and, therefore, the unit cell appears to have only half the actual repeat distance. For displacements perpendicular to the z axis the repeat distance is exactly one-half that of the unit cell. These findings demonstrate that the acoustic and lowest-frequency optic modes, which describe intermolecular vibrations within the whole unit cell, are governed by the space-group symmetric arrangement of the protein molecules, whereas the protein internal motions and the dynamics of the disordered solvent molecules are reflected in the higher-frequency optic modes. Furthermore, these results are complementary to studies of single proteins in solution for which the intramolecular dynamics has been found to be slaved to the solvent fluctuations [17].

From the acoustic branches in Fig. 1 the anisotropic velocities of sound v were determined using $v = \frac{\partial\omega}{\partial q} = \frac{2\pi\nu}{q}$ in the long-wavelength limit (Table I). The longitudinal v range from 3372.5 to 3713.5 ms^{-1} with the average over all directions being $3585 \pm 185\ \mathrm{ms}^{-1}$, approximately double the room-temperature experimental value of $v_{\mathrm{RNase}} = 1784 \pm 72\ \mathrm{ms}^{-1}$ [18]. This difference is due, in part, to the temperature dependence of v. Proteins in general [19,20], and RNase in particular [21], exhibit a dynamical transition with temperature resembling the glass transition. For other polymer systems exhibiting a glass transition, v has been found to decrease approximately linearly with increasing temperature [22–25], with slopes ranging from -4.3 to $-2.2\ \mathrm{ms}^{-1}\,\mathrm{K}^{-1}$ for polyethylene glycol (PEG) in CCl_4 [25] and poly(4-methyl-1-pentene) (P4MP1) [22], respectively, and in particular $-2.5\ \mathrm{ms}^{-1}\,\mathrm{K}^{-1}$ for crystals of lysozyme [hen egg-white lysozyme (HEWL) [24]]. Moreover, whereas an earlier measurement (using laser-generated ultrasound, the same technique as in Ref. [18]) at room temperature yielded $v_{\mathrm{HEWL}} = 1817\ \mathrm{ms}^{-1}$ [24], a more recent study using Brillouin scattering found a significantly larger $v_{\mathrm{HEWL}} = 2310 \pm 80\ \mathrm{ms}^{-1}$ [26]. Thus, combining the above considerations, the calculated sound velocity is not obviously inconsistent with the available experimental data.

The question arises as to whether aspects of the dispersion relations in Fig. 1 can be quantitatively reproduced using a simplified model, in which each protein molecule is represented by a rigid body of mass M. In this case, the RNase crystal space group can be represented by a one-site (mass $2M$, force constants $2G_{a,b}$) linear chain along $\mathbf{a}$ and $\mathbf{b}$, and a two-site (M, G_c) chain along $\mathbf{c}$, with G_i implicitly incorporating the solvent. G_i is related to the macroscopic elastic constant, $C_i = G_i l_i^2$ with $l_i = \{a, b, c\}$, and is given by $G_i = v_{L,i}^2 M/l_i^2$ [16]. With $M = 13691$ u and v_L from Table I, the protein:protein coupling force constants are $G_{a,b,c} = \{28.4, 20.7, 22.1\}\ \mathrm{kg\,s}^{-2}$. The comparatively larger G_a is likely to be due to the composition of the protein:protein interfaces, which involve a rather rigid α-helix:β-sheet structure along $\mathbf{a}$, but more flexible

TABLE I. Sound velocities, v, along $\mathbf{a}^*$, $\mathbf{b}^*$, and $\mathbf{c}^*$. The subscript to v indicates TA or LA modes and units are in $\mathrm{m\,s}^{-1}$.

	$\mathbf{a}^* = \|\|100\|\|$	$\mathbf{b}^* = \|\|010\|\|$	$\mathbf{c}^* = \|\|001\|\|$	Mean value
v_{T_1}	1722.9	1716.0	1695.3	1711 ± 14
v_{T_2}	1819.5	1908.8	1911.2	1880 ± 52
v_L	3372.5	3667.7	3713.5	3585 ± 185

loop:loop structures along **b** and **c**. The optical mode frequency at [000] can be estimated along $\mathbf{c}^*$ yielding $\nu = (\frac{G_c}{\pi^2 M})^{0.5} \approx 0.31$ THz which is in excellent agreement with the results of the atomic-detail model (Fig. 1).

The dispersion relations shown in Fig. 1 allow the average density of states $g(\nu)$ to be calculated. In general, $g(\nu)$ follows from averaging $\boldsymbol{\omega}(\mathbf{q})$ over the whole Brillouin zone. For practical application, however, this averaging must be restricted to a coarse grid in reciprocal space. Here it was found that, due to the low dispersion of the optical modes, it suffices to restrict this grid to the reciprocal lattice points shown in Fig. 1. $g(\nu)$ was then estimated by averaging over $\mathbf{a}^*$, $\mathbf{b}^*$, and $\mathbf{c}^*$ and is shown in Fig. 2. At low frequency, $g(\nu)$ exhibits a pronounced peak centered at 2.1 THz. Figure 2 also shows the experimental profile for dihydrofolate reductase (DHFR) which has been determined using inelastic neutron scattering [27]. The low-frequency peak in $g(\nu)$ is of closely similar absolute amplitude and position. It has previously been found that the position of this peak varies little with the particular protein [28]. Therefore, the calculated $g(\nu)$ is in agreement with the existing experimental data.

From $g(\nu)$, the temperature-dependent constant-volume heat capacity C_V can be calculated using

$$C_V(T) = \frac{h^2}{k_B T^2} \sum_i g(\nu_i) \nu_i^2 \frac{\exp\left(\frac{h\nu_i}{k_B T}\right)}{\left[\exp\left(\frac{h\nu_i}{k_B T}\right) - 1\right]^2} \Delta\nu, \quad (1)$$

where $\Delta\nu$ is the width of the frequency bins and i runs over all bins. The result is shown in Fig. 3. The inset to Fig. 3 shows that, at low temperature, C_V cannot be adequately described using the Debye T^3 law. Rather, $C_V \propto T^b$ with $b = 1.68$ for $T \leq 35$ K, which is in excellent agreement with the experimental range, $b = 1.60$–1.77, determined for other globular hydrated proteins [7].

Figure 3 also shows a comparison with experimental data for dry crystals of RNase [29] and RNase in aqueous solution [30]. The form of the calculated profile is very similar to that of dry RNase, with most of the average difference of 0.2 $\mathrm{Jg^{-1}\,K^{-1}}$ consistent with the experimental dehydration. The difference between the constant-volume and constant-pressure heat capacities, C_V and C_P, respectively, has been estimated for proteins to be $\approx$ 5% at 300 K and decreases rapidly with decreasing temperature [29]. Because of the onset of anharmonic motions at the dynamical transition temperature, $\approx$200 K, C_V can deviate from the harmonic description [Eq. (1)], as has been observed for hydrated crystals of lysozyme for which C_P was found to increase relative to the harmonic behavior in the range 150–220 K [31].

The satisfactory agreement with various experiments described above indicates that the present phonon model is a useful description of protein crystal vibrations. Phonons inelastically scatter x rays and our final calculations are dedicated to this so-called thermal diffuse scattering. TDS is centered at the Bragg peaks. Understanding TDS will provide dynamical information and also result in more accurate protein structural models.

For an inelastic scattering process, involving the emission or absorption of a single phonon, the scattering intensity can be written as [12,32]

$$I_{\mathrm{in}}(\mathbf{q}) = \frac{k}{k_0} \sum_j \frac{h}{2\nu_j} \left| \sum_\alpha f_\alpha e^{-W_\alpha + i(\mathbf{G}-\mathbf{q})\cdot\mathbf{r}_\alpha} \frac{(\mathbf{G}-\mathbf{q})\cdot\mathbf{u}_\alpha^j}{\sqrt{M_\alpha}} \right|^2 \times \coth\left(\frac{h\nu_j}{2k_B T}\right), \quad (2)$$

where k and k_0 are the magnitudes of the scattered and incident wave vectors, respectively, W_α is the Debye-Waller factor, **G** is the nearest reciprocal lattice point, and j is the mode index. Using Eq. (2), the scattering profile was calculated and is shown along $\mathbf{b}^*$ in Fig. 4. I_{in} is composed of a slowly varying nonmonotonic baseline, found to be due to inelastic scattering by the optic modes ($j > 3$), onto which intense peaks of varying height and width are superposed at regular intervals. The peaks are centered at the Bragg peak locations. In the vicinity of the

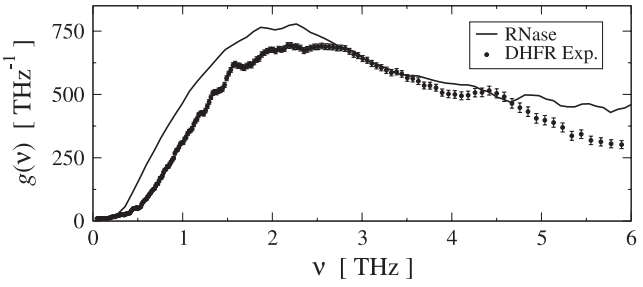

FIG. 2. Density of states $g(\nu)$ calculated from the phonon dispersion relation $\boldsymbol{\omega}(\mathbf{q})$ and averaged over $\mathbf{a}^*$, $\mathbf{b}^*$, and $\mathbf{c}^*$. Also shown is 120 K experimental data for hydrated (0.3 g water/g dry protein) powders of dihydrofolate reductase [27]. For comparison, the experimental data have been scaled by the ratio (3.05) of DHFR hydrogens (which dominate the experimental neutron scattering) to RNase protein atoms.

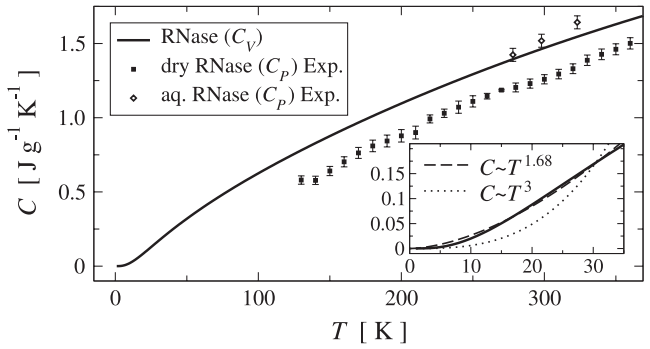

FIG. 3. Heat capacity C_V of the RNase crystal calculated using Eq. (1). Experimental heat capacities C_P for dry RNase crystals [29] and RNase in aqueous solution [30] are also shown. The inset shows the low-temperature regime in more detail. Dotted and dashed lines are least-squares fits of the Debye T^3-law and a T^b-function, respectively, to C_V^{RNase} for $T \leq 35$ K.

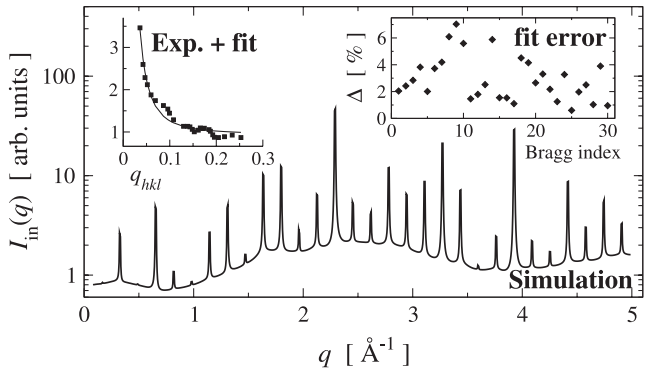

FIG. 4. Inelastic x-ray scattering $I_{\text{in}}(q)$ given by Eq. (2) versus q, along $\mathbf{b}^*$, calculated from all 19326 phonons, at 100 K. The q range is chosen to include the first 30 Bragg peaks, corresponding to 1.28 Å real-space resolution. The left inset shows experimental data for a single 6 Å-resolution reflection for RNase [6] and the least-squares fit $y(q) = A_0 + A_1 q^{-2}$. The right inset shows the root-mean-square deviation, $\Delta = \langle[(I_{\text{in}} - y)/I_{\text{in}}]^2\rangle^{0.5}$, for each Bragg peak.

Bragg peaks (small q) the frequency weighting in Eq. (2) obeys the relation $\coth(x) \propto x^{-1}$ for small x, hence $I_{\text{in}}^j \propto \nu_j^{-2}$ and therefore inelastic scattering from acoustic modes ($j \leq 3$, $\nu \propto q$) may be dominant. To determine whether the peaks in I_{in} are indeed due to purely acoustic scattering, the deviation from the theoretical q^{-2} dependence of I_{in} was calculated (see caption to Fig. 4 for details). The results are presented in an inset to Fig. 4 and show a maximal and average deviation of 7% and 3%, respectively. Therefore, the peaks in I_{in} are well approximated by the contribution from the acoustic mode scattering. This finding is consistent with the experimental observation, also shown in Fig. 4, that, for the same RNase crystal [6], the diffuse scattering intensity around a 6 Å resolution Bragg peak varies approximately as $I \propto q^{-2}$.

The lattice-dynamical calculations presented here extend the domain of application of a fundamental technique in theoretical solid-state physics to the study of crystalline proteins. In principle, the dispersion relations can be determined experimentally using inelastic neutron scattering, but the requirements of large protein crystals and high neutron flux have prevented such studies in the past. However, the coming online of high intensity neutron sources like the Spallation Neutron Source at Oak Ridge National Laboratory, and the satisfactory agreement with existing experimental data found here, suggest a combined theoretical and experimental lattice-dynamical approach that should yield a mine of information on protein crystals and the physics of interacting proteins.

*Present address: California Institute of Technology, 1200 East California Boulevard, Pasadena, CA 91125, USA.
†lars.meinhold@caltech.edu

[1] H. M. Berman *et al.*, Nucleic Acids Res. **28**, 235 (2000).
[2] J. Doucet and J. P. Benoit, Nature (London) **325**, 643 (1987).
[3] D. L. D. Caspar, J. Clarage, D. M. Salunke, and M. Clarage, Nature (London) **332**, 659 (1988).
[4] G. U. Nienhaus, J. Heinzl, E. Huenges, and F. Parak, Nature (London) **338**, 665 (1989).
[5] A. Wlodawer, L. A. Svensson, L. Sjolin, and G. L. Gilliland, Biochemistry **27**, 2705 (1988).
[6] I. D. Glover, G. W. Harris, J. R. Helliwell, and D. S. Moss, Acta Crystallogr. Sect. B **47**, 960 (1991).
[7] J. Edelman, Biopolymers **32**, 209 (1992).
[8] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, J. Comput. Chem. **4**, 187 (1983).
[9] A. D. MacKerell *et al.*, J. Phys. Chem. B **102**, 3586 (1998).
[10] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[11] H. Ibach and H. Lüth, *Festkörperphysik* (Springer, Heidelberg, 1995).
[12] F. Merzel, F. Fontaine-Vive, and M. R. Johnson, Comput. Phys. Commun. **177**, 530 (2007).
[13] D. Durand, M. J. Field, M. Quilichini, and J. C. Smith, Biopolymers **33**, 725 (1993).
[14] A. M. Micu, D. Durand, M. Quilichini, M. J. Field, and J. C. Smith, J. Phys. Chem. **99**, 5645 (1995).
[15] I. Natkaniec *et al.*, J. Phys. C **13**, 4265 (1980).
[16] M. T. Dove, *Introduction to Lattice Dynamics* (Cambridge University Press, Cambridge, U.K., 2005).
[17] P. W. Fenimore, H. Frauenfelder, B. H. McMahon, and F. G. Parak, Proc. Natl. Acad. Sci. U.S.A. **99**, 16 047 (2002).
[18] C. Edwards, S. B. Palmer, P. Emsley, J. R. Helliwell, I. D. Glover, G. W. Harris, and D. S. Moss, Acta Crystallogr. Sect. A **46**, 315 (1990).
[19] W. Doster, S. Cusack, and W. Petry, Nature (London) **337**, 754 (1989).
[20] D. Ringe and G. A. Petsko, Biophys. Chem. **105**, 667 (2003).
[21] B. F. Rasmussen, A. M. Stock, D. Ringe, and G. A. Petsko, Nature (London) **357**, 423 (1992).
[22] J. K. Krüger, L. Peetz, M. Pietralla, and H. G. Unruh, Colloid Polym. Sci. **259**, 215 (1981).
[23] M. Pietralla, R. M. Weeger, and D. B. Mergenthaler, Z. Phys. B **77**, 219 (1989).
[24] M. Tachibana, K. Kojima, R. Ikuyama, Y. Kobayashi, and M. Ataka, Chem. Phys. Lett. **332**, 259 (2000).
[25] M. Pochylski, F. Aliotta, Z. Blaszczak, and J. Gapinski, J. Phys. Chem. B **109**, 4181 (2005).
[26] S. Speziale, F. Jiang, C. L. Caylor, S. Kriminski, C. Zha, R. E. Thorne, and T. S. Duffy, Biophys. J. **85**, 3202 (2003).
[27] E. Balog, T. Becker, M. Oettl, R. Lechner, R. Daniel, J. L. Finney, and J. C. Smith, Phys. Rev. Lett. **93**, 028103 (2004).
[28] M. Nöllmann and P. Etchegoin, Phys. Rev. E **60**, 4593 (1999).
[29] G. Zhang, S. Gerdes, and B. Wunderlich, Macromol. Chem. Phys. **197**, 3791 (1996).
[30] P. L. Privalov and G. I. Makhatadze, J. Mol. Biol. **213**, 385 (1990).
[31] Y. Miyazaki, T. Matsuo, and H. Suga, J. Phys. Chem. B **104**, 8044 (2000).
[32] B. Donovan and J. F. Angress, *Lattice Vibrations* (Chapman and Hall, London, 1971).