J|A|C|S
ARTICLES
Published on Web 02/02/2006

Low-Temperature Protein Dynamics: A Simulation Analysis of Interprotein Vibrations and the Boson Peak at 150 K

Vandana Kurkal-Siebert* and Jeremy C. Smith

Contribution from the Interdisciplinary Center for Scientific Computing (IWR), University of Heidelberg, Im Neuenheimer Feld 368, D-69120 Heidelberg, Germany

Received September 12, 2005; E-mail: vandana.kurkal@iwr.uni-heidelberg.de

Abstract: An understanding of low-frequency, collective protein dynamics at low temperatures can furnish valuable information on functional protein energy landscapes, on the origins of the protein glass transition and on protein−protein interactions. Here, molecular dynamics (MD) simulations and normal-mode analyses are performed on various models of crystalline myoglobin in order to characterize intra- and interprotein vibrations at 150 K. Principal component analysis of the MD trajectories indicates that the Boson peak, a broad peak in the dynamic structure factor centered at about ∼2−2.5 meV, originates from $\sim 10^2$ collective, harmonic vibrations. An accurate description of the environment is found to be essential in reproducing the experimental Boson peak form and position. At lower energies other strong peaks are found in the calculated dynamic structure factor. Characterization of these peaks shows that they arise from harmonic vibrations of proteins relative to each other. These vibrations are likely to furnish valuable information on the physical nature of protein−protein interactions.

1. Introduction

There is a wide variety of vibrational motions in proteins, ranging from high-frequency ($\sim 10^{-15}$ s) localized oscillations to low-frequency ($\sim 10^{-11}$ s) collective modes.[1,2] The low-frequency modes may play an important role in protein function. For example, it has been recently demonstrated that ligand binding to an enzyme is accompanied by softening of low-frequency vibrations and that this contributes significantly to the binding free energy[3]. Moreover, normal-mode analyses have suggested that the conformational paths between different functional states are initiated by a few, delocalized, low-frequency modes in the frequency range below 2.5 meV.[4−6] Furthermore, ligand dissociation experiments in heme proteins have demonstrated coherent excitation of low-frequency vibrational modes in the 5−20 meV range, indicating a strong coupling of these modes with the binding reaction.[7,8]

A prominent feature observed in low-frequency inelastic neutron and Raman scattering spectra of glass-forming substances is the "Boson peak", defined as an excess in the vibrational density of states over the Debye level. This feature is also visible in proteins as a broad band occurring in an energy range of 1.5−3.5 meV[2,9−24] depending on temperature and hydration. Although a variety of studies has been performed, the origin of the Boson peak is not very well understood. Elucidating its origin is of fundamental importance in understanding the physical properties of glassy materials, biopolymers, and biological macromolecules.

The Boson peak appears to be connected to low-temperature anomalies in the specific heat and thermal conduction of glasses.[20,25−27] The intensity of the peak has been correlated with the strength of glass-forming substances.[25] Several theoretical models of the Boson peak dynamics have been proposed. One of these uses coupled classical oscillators with spatially

(1) Frauenfelder, H.; Sligar, S. G.; Wolynes, P. G. *Science* **1991**, *254*, 1598−1603.
(2) Frauenfelder, H.; Parak, F.; Young, R. D. *Annu. Rev. Biophys. Biophys. Chem.* **1988**, *17*, 451−479.
(3) Balog, E.; Becker, T.; Oetl, M.; Lechner, R.; Daniel, R.; Finney, J.; Smith, J. C. *Phys. Rev. Lett.* **2004**, *93*, 028103.
(4) Levitt, M.; Sander, C.; Stern, P. S. *J. Mol. Biol.* **1985**, *181*, 423−447.
(5) Marques, O.; Sanjouand, Y. H. *Proteins* **1995**, *23*, 557−560.
(6) Seno, Y.; Go, N. *J. Mol. Biol.* **1990**, *216*, 111−126.
(7) Groot, M. L.; Vos, M. H.; Schlichting, I.; van Mourik, F.; Joffre, M.; Lambry, J. C.; Martin, J. L. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1323−1328.
(8) Rosca, F.; Kumar, A. T. N.; Ionascu, D.; Ye, X.; Demidov, A. A.; Sjodin, T.; Wharton, D.; Barrick, D.; Saligar, S. G.; Tonetani, T.; Champion, P. M. *J. Phys. Chem. A* **2002**, *106*, 3540−3552.
(9) Smith, J. C. *Q. Rev. Biophys.* **1991**, *24*, 227−291.
(10) McCammon, J. A. *Rep. Prog. Phys.* **1984**, *47*, 1−46.
(11) Paciaroni, A.; Stroppolo, M. E.; Arcangeli, C.; Bizzarri, A. R.; Desideri, A.; Cannistraro, S. *Eur. Biophys. J.* **1999**, *28*, 447−456.
(12) Cusack, S.; Doster, W. *Biophys. J.* **1990**, *58*, 243−251.
(13) Diehl, M.; Doster, W.; Schober, H. *Biophys. J.* **1997**, *73*, 2726−2732.
(14) Cusack, S.; Doster, W.; Petry, W. *Phys. Rev. Lett.* **1990**, *65*, 1080−1083.
(15) Ferrand, M.; Dianoux, A. J.; Petry, W.; Zaccai, G. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 9668−9672.
(16) Fitter, J. *Biophys. J.* **1999**, *76*, 1034−1042.
(17) Brown, K.; Erfurth, S.; Small, E. W.; Peticolas, W. L. *Proc. Natl. Acad. Sci. U.S.A.* **1972**, *69*, 1467−1469.
(18) Richter, D. J. *J. Phys.: Condens. Matter* **1996**, *8*, 9177−9190.
(19) Engberg, D.; Wischnewski, A.; Buchenau, U.; Börjesson, L.; Sokolov, A. P.; Torell, L. M. *Phys. Rev. B* **1998**, *58*, 9087−9097.
(20) Frick, B.; Richter, D. *Science* **1995**, *267*, 1939−1945.
(21) Wutke, J.; Hernandez, J.; Li, G.; Coddens, G.; Cummins, H. Z.; Fujara, F.; Petry, W.; Sillescu, H. *Phys. Rev. Lett.* **1994**, *72*, 3052−3055.
(22) Hehlen, B.; Courtens, E.; Vacher, R.; Yamanaka, A.; Kataoka, K.; Inoue, K. *Phys. Rev. Lett.* **2000**, *84*, 5355−5358.
(23) Diehl, M.; Doster, W.; Petry, W.; Schober, H. *Biophys. J.* **1997**, *73*, 2726−2732.
(24) Smith, J. C.; Kuczera, K.; Tidor, B.; Karplus, M. *Physica B* **1989**, *156* and *157*, 437−443.
(25) Sokolov, A. P.; Rossler, E.; Kisliuk, A.; Quitmann, D. *Phys. Rev. Lett*, **1991**, *71*, 2062−2065.
(26) Angell, C. A. *Science* **1995**, *267*, 1924−1935.
(27) Ngai, K. L.; Sokolov, A.; Steffen, W. *J. Chem. Phys.* **1997**, *107*, 5268−5272.

10.1021/ja055962q CCC: $33.50

fluctuating nearest-neighbour force constants on a simple cubic lattice.[28] Another model, based on a soft potential, predicts that anharmonic localized potential wells are responsible for the Boson peak vibrational anomaly.[29] A mechanism based on the concept of interacting quasilocal oscillators has also been proposed.[30] Another, a "Two-Order-Parameter" model, proposes that the Boson peak arises from localized vibrational modes associated with long-lived locally favored structures that are intrinsic to the liquid state.[31]

Strict analogies between glasses or amorphous systems and proteins are difficult. Glasses are higly viscous liquids outside thermodynamic equilibrium and can be treated as a mestastable disordered crystal over relevant time scales due to their extremely slow solidification dynamics. Also, they display glass transitions which depend on their cooling rate. These features can be accounted for by the rugged potential energy landscape as in proteins. It is, however more justified to talk about "glasslike" properties of proteins.

It has been proposed that the low-frequency spectrum of a protein can be interpreted in terms of the vibrations of an elastic sphere,[32] which implies that the characteristic frequency of the Boson peak should vary with the reciprocal of the radius. This was observed not to be the case in proteins.[33] Far-infrared emission by Boson peak vibrations in globular proteins led to a description of the Boson peak as a manifestation of the elastic limit in the viscoelastic behavior of liquids.[34]

Extensive MD simulations as a function of temperature, carried out on both dry and hydrated protein systems reproduce the dynamic structure factor well, but with the position of the Boson peak shifted to a lower frequency compared to that of the experimental value.[35,36] A recent MD simulation suggests that the Boson peak arises from motions distributed throughout the protein,[37] and other work describes the Boson peak in terms of a hydration-related multiple minima protein energy landscape.[38]

Another important temperature-dependent phenomenon observed in proteins is the dynamical transition, involving an increase in the atomic fluctuations above the linear regime as a function of temperature. This dynamical transition, observed at ~180−220 K in proteins, is driven by solvent molecule translations[39,13] that activate a small number of collective motions in the protein.[40]

Here, we examine low-frequency modes in a protein crystal (myoglobin) at 150 K. At this temperature the majority of protein motions are expected to be harmonic. This leads to peaks in the amplitude-weighted density-of-states, i.e., the dynamic structure factor, that can be identified with specific motions in the system. The dynamics are calculated using molecular dynamics simulation and normal-mode analysis of a highly studied protein, myoglobin. The protein is simulated in a partially dehydrated crystalline state. This allows analysis of environment-dependent low-frequency dynamics in a system with realistic protein−protein contacts and with hydration levels similar to those in neutron scattering experiments that have been performed on the same protein.

The simulations provide evidence for considerable structuring of the low-frequency dynamic structure factor. The Boson peak is found to be environment-dependent and to consist of a large number of harmonic vibrations. Interestingly, other sharper peaks are also found at lower frequencies and are demonstrated to originate from harmonic interprotein vibrations. The results point the way to future research into the harmonic dynamics of protein networks and clusters at low temperatures.

2. Methods

2.1. Molecular Dynamics Simulations. Molecular dynamics simulations were performed of carboxymyoglobin at 150 K in a variety of environments. The starting structure 1A6G (monoclinic unit cell) was taken from the Protein Data Bank (http://www.rcsb.org), solved at a 1.15 Å resolution using X-ray crystallography.[41] MD simulations were performed with a primary box replicated with periodic boundary conditions. Three models were constructed of crystalline monoclinic carboxymyoglobin of unit cell dimensions $a \times b \times c$ of $63.80 \times 30.63 \times 34.42$ Å^3. The three models differed in the size of the primary simulation box. In one this comprised of one model unit cell, in the second, two unit cells ($2a \times b \times c$), and in the third, four unit cells ($2a \times 2b \times c$). A further model consisted of a single carboxy-myoglobin molecule in an orthorhombic primary box of dimensions $40.0 \times 50.0 \times 42.0$ Å^3.

The hydration of 0.35 h used here corresponds to the hydration of carboxy myoglobin used in the neutron scattering experiments of ref 23. Hydration was carried out by initially filling up all the spaces in the crystal with water, energy minimizing, heating to 300 K, equilibrating for 10 ps with velocity scaling in the *NVE* ensemble, and performing a constant pressure MD simulation at 300 K for 100 ps. Water molecules were then randomly removed until the required hydration was achieved. 26 chloride ions were added per unit cell leading to electrically neutral systems. The total number of atoms was 7174 atoms for the one-unit-cell simulation, 14 348 atoms for the two-unit-cell simulation, and 28 696 atoms for the four-unit-cell simulation. The simulation model with a single myoglobin molecule in an orthorhombic box contained 3605 atoms.

Calculations were performed with the CHARMM program[42] and the potential function parameter set CHARMM22.[43] Water molecules were modeled with the TIP3P potential.[44] Electrostatic interactions were computed using the particle mesh Ewald method[45] for which the direct sum cutoff was 16 Å and the reciprocal space structure factors were computed on a $64 \times 32 \times 32$ grid for $a \times b \times c$, $128 \times 32 \times 32$ grid for $2a \times b \times c$, $128 \times 64 \times 32$ grid for $2a \times 2b \times c$ of monoclinic crystal dimensions, and $64 \times 64 \times 64$ grid for orthorhombic crystal dimensions using sixth-degree B-splines.

The systems were energy minimized to a root-mean-square (RMS) force gradient of 10^{-3} kcal/mol/Å. Subsequently the systems were

(28) Schirmacher, W.; Diezemann, G.; Ganter, C. *Phys. Rev. Lett.* **1998**, *81*, 136−139.
(29) Buchenau, U.; Galperin, Y. M.; Gurevich, V. L.; Parshin, D. A.; Rasmos, M. A.; Schober, H. R. *Phys. Rev. B* **1992**, *46*, 2798−2808.
(30) Gurevich, V. L.; Parshin, D. A.; Schober, H. R. *JETP Lett.* **2002**, *76*, 650−654
(31) Tanaka, H. *J. Phys. Soc. Jpn*. **2001**, *70*, 1178−1182.
(32) Go, N. *Biopolymers* **1978**, *17*, 1373−1379.
(33) Painter, P.; Mosher, L.; Rhoades, C. *Biopolymers* **1982**, *21*, 1469−1472.
(34) Leyser, H.; Doster, W.; Diehl, M. *Phys. Rev. Lett.* **1999**, *82*, 2987−2990.
(35) Smith, J.; Kuczera, K.; Karplus, M. *Proc. Natl. Acad. Sci U.S.A.* **1990**, *87*, 1601−1605.
(36) Steinbach, P. J.; Brooks, B. R. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 9135−9139.
(37) Tarek, M.; Tobias, D. J. *J. Chem. Phys.* **2001**, *115*, 1607−1612.
(38) Joti, Y.; Kitao, A.; Go, N. *J. Am. Chem. Soc.* **2005**, *127*, 8705−8709.
(39) Tournier, A. L.; Xu, J.; Smith, J. C. *Biophys. J.* **2003**, *85*, 1871−1875.
(40) Tournier, A. L.; Smith, J. C. *Phys. Rev. Lett*. **2003**, *91*, 208106.
(41) Vojtechovsky, J.; Chu, K.; Brendzen, J.; Sweet, R. M.; Schlichting, I. *Biophys. J.* **1999**, *77*, 2153−2174.
(42) Brooks, B. R.; Kumar, A. T. N.; Ionascu, D.; Ye, X.; Demidov, A. A.; Sjodin, T.; Wharton, D.; Barrick, D.; Sligar, S. G.; Tonetani, T.; Chamipon, P. M. *J. Comput. Biol*. **1983**, *4*, 187−217.
(43) Mackerell, A. D.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(44) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(45) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577−8593.

uniformly heated to 150 K during 15 ps and equilibrated for 100 ps with velocity scaling in the *NVE* ensemble with $P = 1$ bar and $T =$ 150 K. Equilibration was continued for an additional 200 ps at constant temperature and pressure conditions without velocity rescaling. The temperature and pressure coupling were enforced with the Nóse−Hoover algorithm[46−48] using the temperature and pressure piston masses of 2000 kcal ps^2 and 500 au, respectively. Subsequently the NPT production runs were performed for 1 ns which is sufficiently long to fully sample the low-frequency vibrations. Coordinates were written out every 50 fs. Before analysis, all coordinate sets were superposed on a primary-box reference structure to remove overall unit cell translation and rotation.

2.2. Normal-Mode Analysis. Normal-mode analyses were performed with two molecules of MbCO (pdb entry 1A6G) in a monoclinic box replicated with periodic boundary conditions.[49−51] Normal modes were calculated using the CHARMM program with version 22 of an all-atom potential function[42] and parameters.[43] Protein molecules containing 5048 protein atoms and water molecules corresponding to 0.35 g/g of protein were energy minimized along with the images to an RMS gradient of 10^{-11} (kcal/mol Å).

The normal-mode analyses were performed on the energy minimized structures by diagonalization of the mass-weighted second-derivative matrices. The analysis resulted in 21 522 modes, 6 of which corresponded to the translation and rotation of the whole unit cell. The inelastic neutron scattering spectrum was calculated from these normal modes (details discussed in the next section).

2.3. Incoherent Neutron Scattering. 2.3.1. Neutron Scattering Intensities from MD Simulations. Neutron scattering experiments measure the dynamic structure factor, $S(\vec{q},\omega)$, with $\vec{q}$ and $E = \hbar\omega$ being the momentum and energy transfers, respectively. Since the incoherent scattering length of a hydrogen atom is an order of magnitude higher than the scattering lengths of all other atoms in the protein and water molecules, the coherent scattering can be assumed to be negligible and the total structure factor is $S(\vec{q},\omega) = S_{\mathrm{inc}}(\vec{q},\omega)$. $S(\vec{q},\omega)$ was computed by Fourier transforming the intermediate scattering function $I_{\mathrm{inc}}(\vec{q}, t)$ calculated from the MD trajectories:

$$S_{\mathrm{inc}}(\vec{q},\omega) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} I_{\mathrm{inc}}(\vec{q}, t)\, e^{-i\omega t}\, \mathrm{d}t \qquad (1)$$

with

$$I_{\mathrm{inc}}(\vec{q}, t) = \frac{1}{N}\sum_i b_{i(\mathrm{inc})}^2 \langle e^{i\vec{q}\vec{r}_i(t)} e^{i\vec{q}\vec{r}_i(0)}\rangle$$

Here, $\vec{r}_i(t)$ and $\vec{r}_i(0)$ are the positions of the atom i at time t and time $t = 0$, respectively, obtained from the MD simulations. The quantity $b_{i(\mathrm{inc})}$ is the incoherent scattering length of atom i. Since in proteins the hydrogen atoms are distributed throughout the molecule, the neutron scattering technique probes the global dynamics of the system. The intermediate scattering function and its Fourier transform, the dynamical structure factor $S_{\mathrm{inc}}(\vec{q},\omega)$, were calculated using the package nMOLDYN.[52] For each q value and each atom an orientational average over intermediate scattering functions for a fixed number of isotropically distributed vectors q_i was performed. The spectrum was smoothed by applying a Gaussian window of the form $W(m) = \exp[-{}^1/_2(\alpha(|m|/(N_t - 1)))^2]$ in the time domain with $m = -(N_t - 1)$ ..., $N_t - 1$. The widths in the time and frequency domains were $\sigma_t = \alpha/T$ and $\sigma_\nu = 1/(2\pi\sigma_t)$, respectively, and T is the total length of the simulation. α was chosen such that the width in the frequency domain corresponds to an instrumental full width at half-maximum (fwhm) of 60 μeV, i.e., that of ref 23.

2.3.2. Global Translation, Rotation, and Internal Contributions of the Individual Molecules to the Structure Factor. Complete separability of the translation, rotation, and internal motions is not possible as they are strongly coupled to each other.[53] To reduce the effect of this coupling on the calculated contributions, the translational contribution of the individual molecules to the total structure factor was calculated from the MD trajectories from which only the rotational contribution is removed. Then, from the resulting structure factor which contains both translational and internal contributions, the structure factor obtained for the internal contribution alone was subtracted. Thus, the coupling between the internal and rotational motions is removed. A similar procedure was adapted to calculate the rotational contribution. However, the coupling between the translational and the internal motion and that between the rotational and the internal motions are still present in the translational and rotational contributions, respectively. The contribution of only the internal motions of the individual molecules to the dynamic structure factor was obtained by calculating the structure factor from trajectories from which the global translation and rotation of the individual protein molecules were removed. The internal contribution has coupling with both the translational and rotational contributions.

2.3.3. Neutron Scattering Intensities from a Harmonic Model. Assumption of harmonic dynamics permits transformation of eq 1, and the result can be expanded in a power series over the normal modes of the protein given by[54]

$$S(\vec{q},\omega) = \sum_\alpha {b_\alpha}^2 \exp[-2W_\alpha(\vec{q})] \prod_\lambda [\sum_{n_\lambda} \exp(n_\lambda \hbar\omega_\lambda \beta/2)\, I_{n_\lambda}(X_{\lambda\alpha})]\, \delta[\omega - \sum_\lambda n_\lambda \omega_\lambda] \qquad (2)$$

where

$$X_{\lambda\alpha} = \frac{\hbar(\vec{q}\cdot\vec{e}_{\lambda\alpha})^2}{2m_\alpha\omega_\lambda \sinh(\hbar\omega_\lambda\beta/2)} \qquad (3)$$

$W_\alpha(\vec{q})$ is the exponent of the Debye−Waller factor, $\exp[-2W_\alpha(\vec{q})]$, for atom α and may also be expressed as a sum over the modes

$$2W_\alpha(\vec{q}) = \sum_\lambda \frac{\hbar(\vec{q}\cdot\vec{e}_{\lambda\alpha})^2}{m_\alpha\omega_\lambda}[2n(\omega_\lambda) + 1] = q^2\langle {u_{q\alpha}}^2\rangle \qquad (4)$$

In eqs 2−4, m_α is the atomic mass, λ labels the mode, n_λ is the number of quanta exchanged between the neutron and mode λ, and $n(\omega_\lambda)$ is the Bose occupancy. $\vec{e}_{\lambda\alpha}$ is the atomic eigenvector for atom α in mode λ, and ω_λ is the mode angular frequency. $\beta = 1/k_BT$ where k_B is Boltzmann's constant and T is the temperature. $\langle {u_{q\alpha}}^2\rangle$ is the projection of the mean-square displacement for the atom α on the scattering vector $\vec{q}$, and the brackets $\langle ...\rangle$ denote a thermal average.

$I_{n_\lambda}(X_{\lambda\alpha})$ is the n_λth-order modified Bessel function given by $I_{n_\lambda}(X_{\lambda\alpha}) = 1/n_\lambda!(X_{\lambda\alpha}/2)^{n_\lambda}$. Equation 2 is an exact quantum-mechanical expression for the scattered intensity. The case where all $n_\lambda = 0$ corresponds to elastic scattering. The case where the $\sum_\lambda n_\lambda = 1$ corresponds to single quantum processes called one-phonon scattering. Higher order terms represent other multiphonon processes. Substitution of normal-mode eigenvectors and eigenvalues in eq 2 allows the calculation of the incoherent inelastic neutron scattering in the harmonic approximation. The calculations performed here are within the one-phonon approximation limit which is valid for $q \rightarrow 0$, i.e.,

(46) Andersen, H. C. *J. Chem. Phys.* **1980**, *72*, 2384−2393.
(47) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695−1997.
(48) Nóse, S.; Klein, M. L. *Mol. Phys.* **1983**, *50*, 1055−1076.
(49) Krishnan, M.; Balasubramaniam, S. *Phys. Rev. B* **2003**, *68*, 064304−64310.
(50) Simdyankin, S. I.; Dzugutov, M.; Taraskin, S. N.; Elliott, S. R. *Phys. Rev. B* **2001**, *63*, 184301−184305.
(51) Fukui, K.; Sumpter, B. G.; Noid, D. W.; Yang, C.; Tuzun, R. E. *Phys. Chem. B* **2000**, *104*, 526−531.
(52) Kneller, G. R.; Keiner, V.; Kneller, M.; Schiller, M. *Comput. Phys. Comm.* **1995**, *91*, 191−214.
(53) Meinhold, L.; Smith, J. C. *Biophys. J.* **2005**, *88*, 2554−2563.
(54) Zemach, A. C.; Glauber, R. J. *Phys. Rev.* **1956**, *101*, 129−136.

$$S(\vec{q},\omega) = \sum_{\alpha=1}^{N}\sum_{\lambda=1}^{3N-6} b_\alpha^2 e^{\hbar\omega_\lambda\beta/2}\frac{e^{-2W_\alpha(\vec{q})}\hbar|\vec{q}\cdot\vec{e}_{\lambda a}|^2}{4m_\alpha\omega_\lambda \sinh\left(\frac{\beta\hbar\omega_\lambda}{2}\right)}\delta(\omega-\omega_\lambda) \quad (5)$$

Equation 5 represents the full quantum-mechanical scattering function for the one-phonon scattering.

2.4. Principal Component Analysis. Principal component analysis is a convenient method for representing the conformational space explored in an MD trajectory.[55,56] The set of principal components is the solution to the eigenvalue problem of the second-moment matrix, *A*, the elements of which are given by

$$A_{ij} = \sqrt{m_i m_j}\langle(\vec{r}_i(t) - \vec{r}_i^{\,m})(\vec{r}_j(t) - \vec{r}_j^{\,m})\rangle \quad (6)$$

where $\vec{r}_i$ and $\vec{r}_j$ are the positions, $\vec{r}_i^{\,m}$ and $\vec{r}_j^{\,m}$, the mean positions, and m_i and m_j, the masses of atom i and j; the average is taken over the time frames of the trajectory. The diagonalization of A yields the eigenvectors, w_k, i.e., the principal components and their associated eigenvalues, χ_k.

To determine the contributions of principal component modes, the trajectory, $\vec{h}_i^{\,k}(t)$ of a mode or a set of principal modes was calculated by projecting the MD trajectory of the internal motion onto the k th principal component, i.e.,

$$\vec{h}_i^k(t) = (\vec{r}_i(t) - \vec{r}_i^{\,m})\cdot w^k \quad (7)$$

where $\vec{r}_i(t)$ is the position vector, and $\vec{r}_i^{\,m}$, the mean position vector. The neutron scattering spectrum was then calculated from $\vec{r}_i^{\,\mathrm{new}}(t)$ given by

$$\vec{r}_i^{\,\mathrm{new}}(t) = (\vec{h}_i^k(t) + \vec{r}_i^{\,m}) \quad (8)$$

The contribution of the principal modes to neutron scattering spectra was quantified by defining

$$R = 1.0 - \langle(S(\vec{q},\omega)_{\mathrm{all}} - S(\vec{q},\omega)_{\mathrm{PCmodes}}/S(\vec{q},\omega)_{\mathrm{all}})\rangle_\omega \quad (9)$$

where $S(\vec{q},\omega)_{\mathrm{all}}$ and $S(\vec{q},\omega)_{\mathrm{PCmodes}}$ are the structure factors calculated from the MD trajectories and contributions from a set of modes, respectively.

3. Results and Discussions

3.1. Comparison of Experimental versus Calculated Spectrum. To avoid interference from the solvent hydrogens, experimental samples are usually H/D exchanged. Here, MD simulations of the crystal were performed with the water content (H_2O) of 0.35 g/g of protein. However, the neutron scattering spectrum is calculated using scattering cross-sections of D for all exchangeable hydrogen atoms. To test the validity of this approximation a further MD simulation was performed with the exchangeable hydrogens replaced by deuteriums (by doubling the hydrogen masses). The spectra calculated with these two different approaches are shown in Figure 1 and are very similar. The "Boson peak" is that at ~2 meV. Interestingly, an additional peak, at ~1 meV, is observed in the calculation. The neutron scattering spectra calculated from MD simulation of monoclinic unit cells of carboxymyoglobin with primary cells of varying dimensions and numbers of molecules are shown in Figure 2. The Boson-peak calculated from the primary cell containing eight molecules reproduces very well the experimentally observed Boson peak in ref 23.

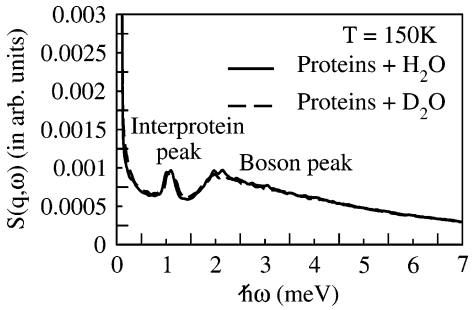

Figure 1. Structure factor $S(q,\omega)$ calculated at $q = 1.9$ Å^{-1} as a function of ω for one unit cell of protein containing two protein molecules hydrated to 0.35 h and using a resolution function of fwhm 60 μeV, i.e, that in ref 23. Solid and dashed lines are from the simulations carried out with H_2O and D_2O as the solvent, respectively. Simulations and experiments were performed at $T = 150$ K. The values are normalized such that $\sum b_{\mathrm{inc}}^2 = 1$ and elastic intensity is 1. The structure factors were calculated after removing the global translation and rotation of the unit cell.

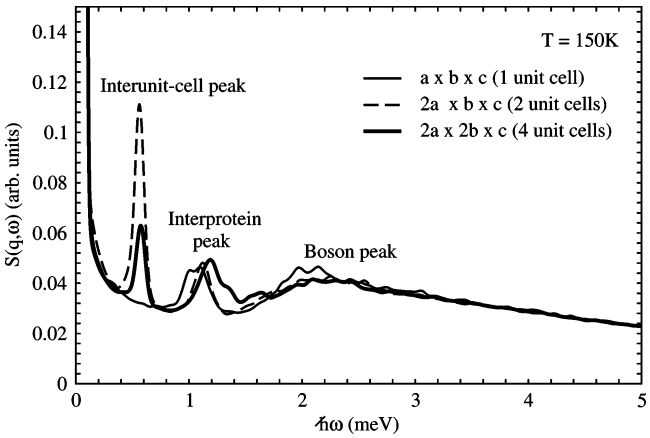

Figure 2. Structure factor $S(q,\omega)$ calculated at $q = 1.9$ Å^{-1} for MbCO crystal hydrated to 0.35 h, as a function of ω for primary cell of dimensions $a \times b \times c$ containing two molecules, $2a \times b \times c$ containing four molecules, and $2a \times 2b \times c$ containing eight molecules where $a \times b \times c$ is the unit cell of monoclinic crystal. The structure factor obtained with eight molecules in the primary cell is closer to the experimentally observed Boson peak.

The ~1 meV peak is not clearly visible, although it may correspond to a poorly resolved "shoulder" in the experimental spectrum of ref 23. However, the 100 K neutron scattering spectrum from plastocyanin[57] does display a peak at ~1 meV, which may originate from interprotein interactions.

3.2. Elucidation of the ~1 meV Peak. The presence of an ~1 meV peak in all simulation systems with varying dimensions of the primary cell and numbers of molecules shown in Figure 2 indicates that the peak is not an artifact due to the periodic boundary conditions. The crystal spectra calculated with two and four unit cells in the primary simulation box exhibit a further peak at ~0.6 meV. This suggests that the peaks observed at ~1 meV and ~0.6 meV could arise from the interaction between the proteins in a unit cell and between the unit cells, respectively.

To further examine the origin of the peak at ~1 meV, the neutron scattering spectrum, calculated from the simulation with a single unit cell of MbCO of dimensions $a \times b \times c$ as the primary image, was decomposed into whole-molecule translational and rotational and internal motional contributions. The resulting spectra are shown in Figure 3. The whole-molecule translation makes the largest contribution to the 1 meV peak.

(55) Karplus, M.; Kushick, J. N. *Macromolecules* **1981**, *14*, 325−332.
(56) Kitao, A.; Hayward, S.; Go, N. *Proteins* **1998**, *33*, 496−517.
(57) Bizzarri, A. R.; Paciaroni, A.; Arcangeli, C.; Cannistraro, S. *Eur. Biophys. J.* **2001**, *30*, 443−449.

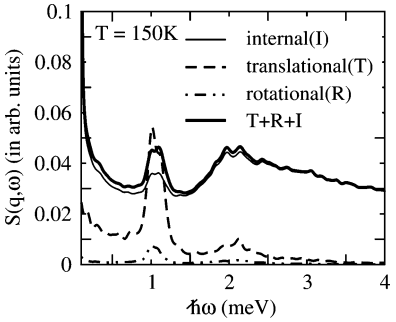

Figure 3. Contributions of rigid-body translational, rotational, and internal motions to the structure factor $S(q,\omega)$ calculated at $q = 1.9$ Å^{-1} as a function of ω for one unit cell of protein containing two molecules at 0.35 h.

The contribution of rotation is small. The structure factor from the internal contribution has an overall higher intensity at all the frequencies compared to the translational and rotational components due to quasielastic scattering but also contributes relatively little to the ~1 meV peak. The sum of the individual contributions do not add up to the total spectrum due to strong coupling between these motions such that their complete separability is not possible.[53]

The intermolecular interaction in protein crystals can be further probed by calculating $d(t)$, the distance between the center of mass of two proteins in the primary simulation cell as a function of time, and this is shown in Figure 4a. The corresponding autocorrelation function, $C(\tau) = \langle d(t)|d(t + \tau)\rangle$ is shown in Figure 4b. The oscillations observed in the autocorrelation function indicate intermolecular vibration. The period of oscillation, T in Figure 4b, is ~3.6 ps. Fourier transformation of the autocorrelation function generates the power spectrum, providing the associated frequencies; this is shown in Figure 4c. The frequency corresponding to $T \approx 3.6$ ps is 1.1 meV which therefore accounts for the ~1 meV peak observed in the structure factor calculated from the MD simulation of the one unit cell primary box. These observations indicate that the ~1 meV peak observed in the structure factor plot indeed arises from the intermolecular vibration between proteins in the unit cell. As can be seen from Figure 4c, center of mass motion contributes also to several vibrations below 0.5 meV which are not observed as individual peaks in the structure factor due to their being smeared out by the resolution function used.

Within the harmonic approximation it is possible to calculate the effective force constant k for the intermolecular interaction between the two protein molecules using the relationship $k = \omega_1^2\mu$, where ω_1 is the frequency of oscillation and μ is the reduced mass. The knowledge of effective force constant can then be used to estimate the frequency, ω_2 of intermolecular interaction in the case where the reduced mass is doubled (2μ), given by $\sqrt{\omega_1^2/2}$. From the structure factor calculated from the MD simulation with four unit cells with dimension $2a \times 2b \times c$ as the primary box, if we assume that the peak observed at 1.1 meV is due to the intermolecular interaction between proteins in the unit cell, then the frequency obtained by doubling the reduced mass corresponds to 0.77 meV. We observe a peak at the frequency of 0.57 meV for the same simulation system. The difference between the observed and the calculated frequency could be due to the assumption that the center of mass vibration between the molecules is harmonic. Thus, the 0.57 meV can be tentatively attributed to the intermolecular vibration between the unit cells.

The neutron scattering spectrum calculated from normal modes for one unit cell of MbCO crystal hydrated to 0.35 g/g of protein is shown in Figure 5. About 180 normal modes lie in the frequency range below 4 meV, where the main vibrational features are observed. Individual peaks are seen in the spectrum arising principally from peaks in the density of states. That these are not seen experimentally is likely due to anharmonicity and conformational heterogeneity.[58] However, the overall shape of the $S(\vec{q},\omega)$ envelope closely resembles that obtained using MD.

To investigate the origin of the vibrational features observed in the normal-mode analysis, the atomic trajectories were calculated as a superposition of the first 180 normal modes. The center of mass trajectories of the individual proteins were

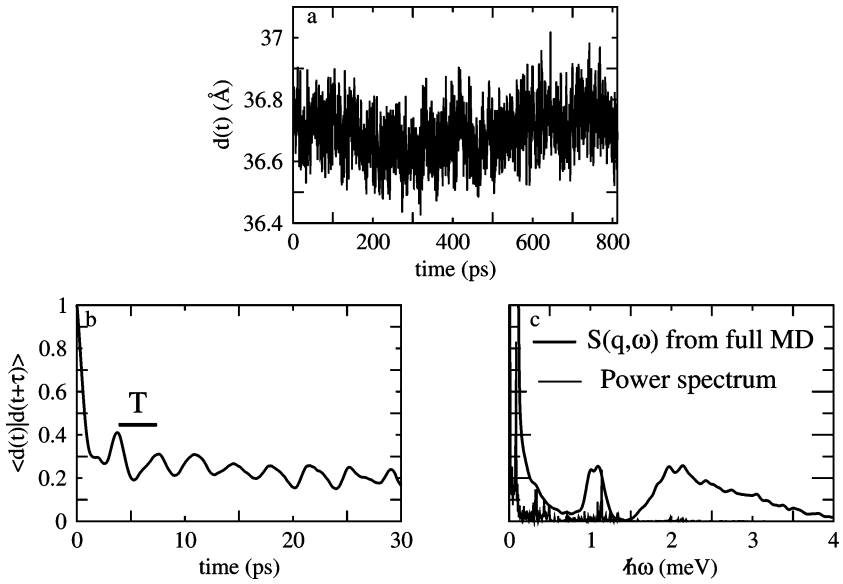

Figure 4. (a) Time series of the distance between the centers of masses of two proteins in the unit cell of hydrated MBCO crystal containing two protein molecules. (b) Autocorrelation function of the distance between the two proteins calculated in part a. (c) Power spectrum calculated by Fourier transforming the autocorrelation function in part b, shown together with $S(q,\omega)$ calculated from the full MD simulation.

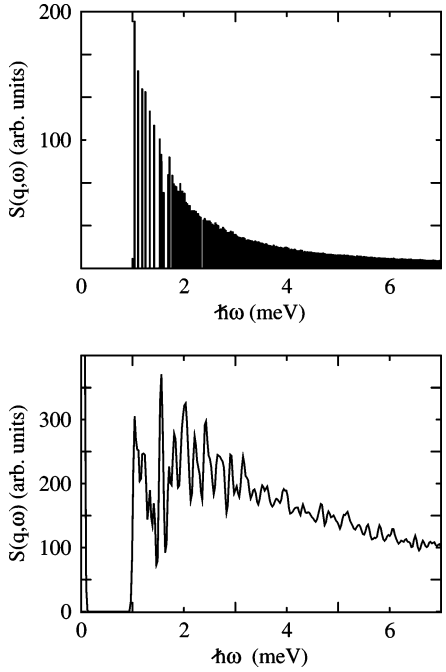

Figure 5. Inelastic neutron scattering spectrum calculated from the normal modes of one unit cell of hydrated monoclinic MbCO crystal containing two protein molecules. (Top) nonconvoluted neutron scattering spectra. (Bottom) Convoluted with the experimental resolution of fwhm = 60 μeV.

then extracted. The distance between the centers of masses is plotted as a function of time in Figure 6a. The autocorrelation function of the distance between the centers of masses and its Fourier transform, the power spectrum, are shown in Figure 6b and 6c. The oscillating behavior of the autocorrelation function obtained from the distances between the centers of masses as a function of time is an indication of intermolecular center-of-mass vibration. The period of oscillation in Figure 6b is ~3.6 ps which in turn corresponds to a frequency of ~1.1 meV. The power spectrum is found to display a peak at 1.0 meV.

MD simulation and normal-mode analysis thus both independently confirm that the ultralow-frequency peak at ~1 meV arises from the intermolecular vibration between the protein molecules.

Further evidence that this peak arises from intermolecular interaction is obtained from the absence of the peak in a simulation system containing only one protein in an orthorhombic box of water, as shown in Figure 7. The periodic

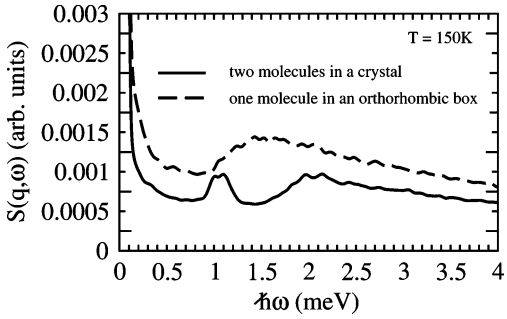

Figure 7. Structure factor $S(q,\omega)$ calculated at $q = 1.9$ Å^{-1} for a monoclinic crystal of MBCO containing two carboxymyoglobin molecules in a unit cell and one carboxymyoglobin molecule in an orthorhombic box. Both systems are hydrated to 0.35 g/g of protein.

boundary conditions exclude the relative motion of primary and image atoms and, thus, interprotein motions.

3.3. Origin of Boson Peak. In contrast to the 1 meV peak, the Boson peak motions are dominated by internal dynamics. The motions contributing to the Boson peak are now further investigated using principal component analysis. Scattering from the first *PCmodes* principal component modes, $S(\vec{q},\omega)_{\text{PCmodes}}$, was calculated by subtracting the dynamic structure factor, $S(q,\omega)$ calculated using all modes other than the first *PCmodes* from $S(q,\omega)$. The value R was calculated by scaling the intensity of the total spectrum, $S(\vec{q},\omega)$, such that it matches $S(\vec{q},\omega)_{\text{PCmodes}}$ at 1.3 meV.

The contributions of the principal component modes to the Boson peak and the R quantifying these contributions are displayed in Figure 8. More than 90% of the Boson peak arises from the first 200 principal component modes, and the proportion rises to 99.6% with an additional 300 modes. Using the

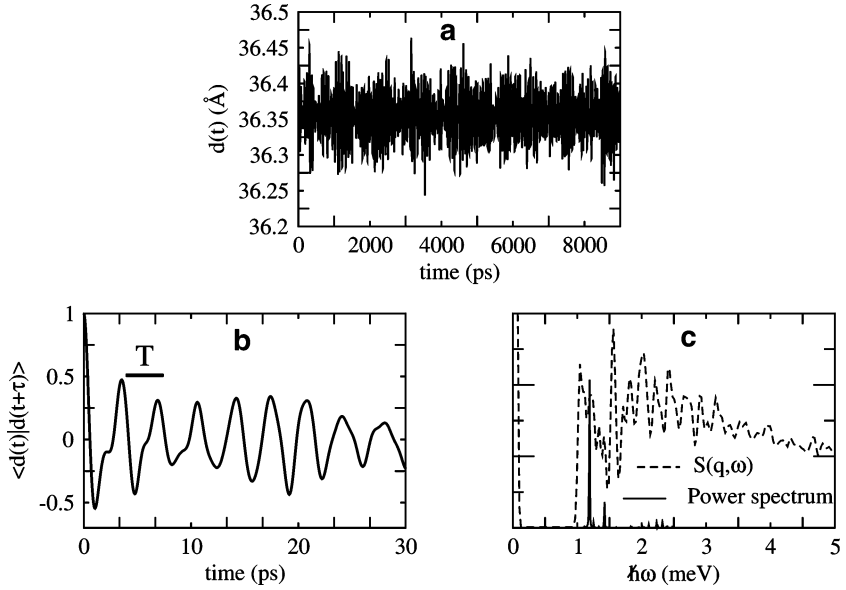

Figure 6. (a) The Distance between the centers of masses calculated from the normal modes of two proteins in the unit cell of hydrated MBCO crystal containing two protein molecules, as a function of time. (b) The autocorrelation function of the distance between the two proteins calculated in part a. (c) The power spectrum calculated from Fourier transforming the autocorrelation function in part b, together with $S(q,\omega)$ from Figure 5.

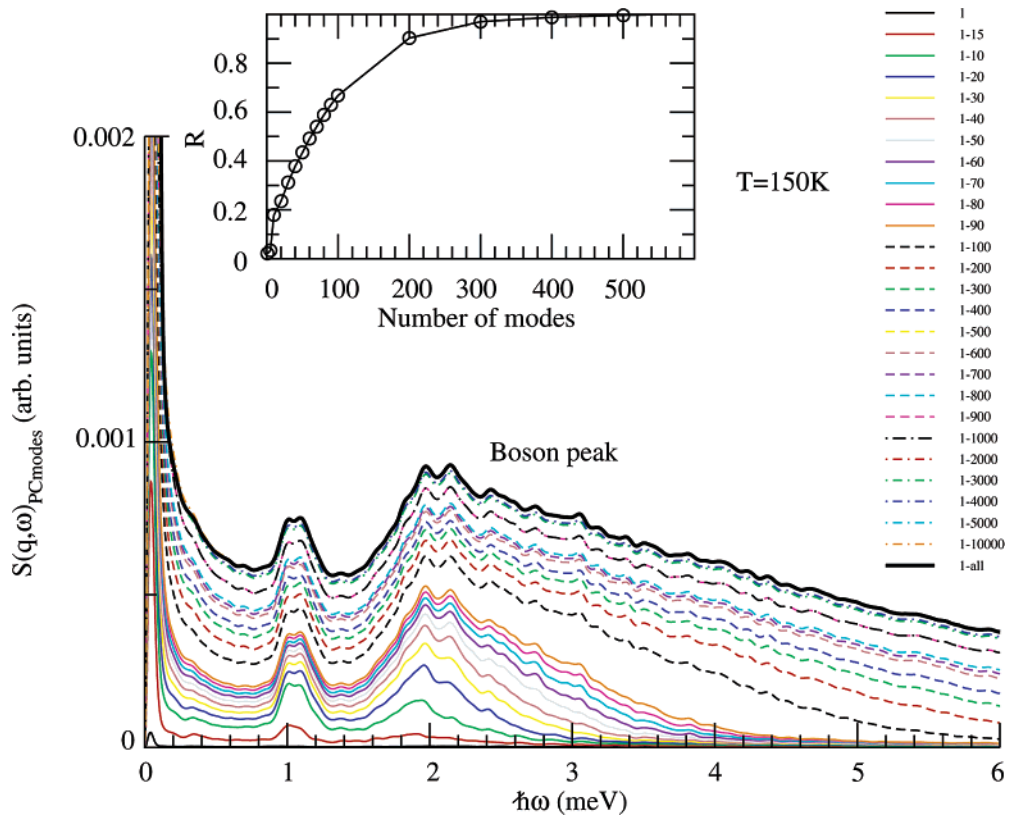

Figure 8. Contributions $S(q,\omega)_{\text{PCmodes}}$ from sets of principal component modes to the structure factor corresponding to the internal motions (global translation and rotation of the individual protein molecules removed) of protein atoms calculated for one unit cell of monoclinic MBCO crystal containing two protein molecules, at 150 K at 0.35 g/g of protein. Inset: Contributions from set of principal component modes $S(q,\omega)_{\text{PCmodes}}$ to the total structure factor $S(q,\omega)$ quantified by a quantity, R. The range E used for the calculation is $E = 1.3-4.4$ meV.

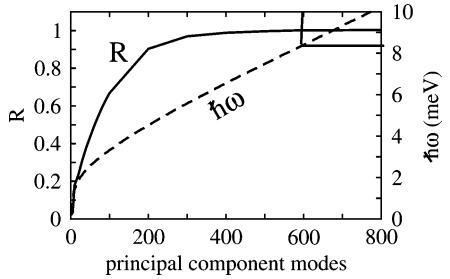

Figure 9. Contributions (R) from set of principal component modes to the structure factor $S(q,\omega)$ corresponding to the internal motions of protein atoms calculated for one unit cell of monoclinic MBCO crystal containing two protein molecules, at 150 K and 0.35 g/g of protein, along with the effective frequencies of the principal component modes. Solid lines in the upper right corner indicate the effective frequency corresponding to the 600th mode.

600 lowest-frequency modes accounts for 100% of the Boson peak. The low-frequency modes are collective motions in the protein, whereas higher frequency modes are more local.[4,59,60] Consequently, the Boson peak arises from collective motions.

The effective frequencies of the principal components are displayed in Figure 9. The contribution of the first five principal component modes to the Boson peak is <5%. Addition of five more modes which lie in the region of ~1−1.3 meV increases the contribution to 18%. Inclusion of modes up to the frequency ~8.7 meV covers the whole Boson peak.

The modes contributing to the Boson peak are now further investigated by calculating the effective free energy along the k th mode given by $\mu_k(x_k) = -k_B T \ln P_k(x_k)$. Furthermore, information on the anharmonicity can be obtained from $P_k(x_k)$, the probability distribution along the mode k.

For harmonic motion $P_k(x_k)$ is a Gaussian, and the standard error σ_k of a Gaussian fit to $P_k(x_k)$, obtained using the relation $\sigma_k^2 = 1000[P_k(x) - G_k(x)]^2$, is therefore zero for harmonic modes. Nonzero values of σ_k indicate anharmonicity.

The free energy profile along principal component mode 1 is shown in Figure 10a. This mode is clearly strongly anharmonic and exhibits multiminimum behavior. The first mode contributes the most to the total root-mean-square fluctuation, as shown in Figure 10b. However, the first mode does not contribute to the Boson peak. The free energy landscapes along the 2−5 largest-amplitude modes, none of which contribute to the Boson peak, are also anharmonic. A representative mode among them (principal component mode 5) is shown in Figure 10c.

It is well-known that the few very low-frequency principal component modes do not converge[61−63] due to the poor sampling. However, this poor sampling concerns only a small number of highly anharmonic modes. These highly anharmonic modes, which describe diffusive motion, contribute only to quasielastic scattering, at lower frequencies than those investigated here, and not to the Boson peak or the protein interaction vibrations. The PCA modes of importance in the present work are harmonic, or nearly harmonic, and their effective frequencies are high enough that they are well sampled during the simulation time. Indeed, the information present in these modes is very similar to that present in a normal-mode analysis.

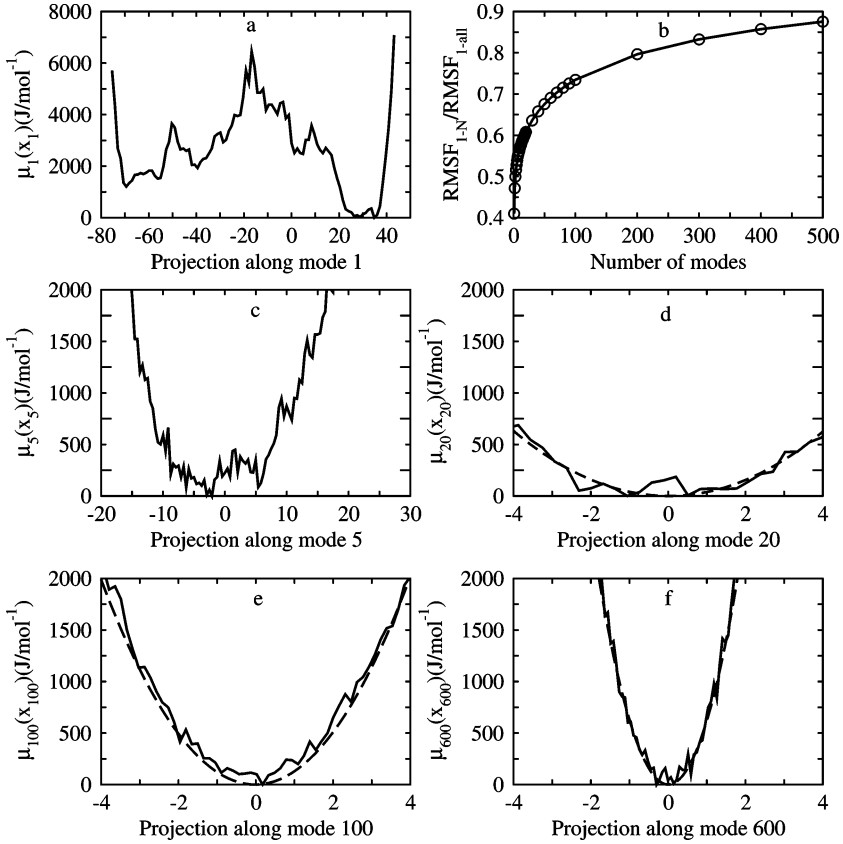

Figure 10. (a) Free energy profile along principal component mode 1. (b) Contributions from set of principal component modes to the mean square fluctuations for a crystal system containing one unit cell of monoclinic MBCO crystal at temperature 150 K at 0.35 g/g of protein. The mean square fluctuation from a set of 1 to N principal component modes, RMSF^{i}_{1-N}, of an atom i was calculated from different sets (same as in Figure 8) of modes, and their contribution to the total mean square fluctuations is obtained by the expression $\mathrm{RMSF}^{i}_{1-N}/\mathrm{RMSF}^{i}_{1-\mathrm{all}}$. (c) Free energy profile along principal component mode 5. (d−f) Free energy profiles along some of the principal component modes that contribute to the Boson peak.

The free energy landscapes obtained for some modes that do contribute to the Boson peak are shown in Figure 10d−f. These modes have σ_k values less than 0.5, indicating close to harmonic behavior.

The collective nature of the modes contributing to the Boson peak can be investigated by calculating the information entropy S, the exponential of which gives the number of atoms that a particular mode spans.[64] The information entropy is calculated using the relation $S = -\sum_{i=1}^{N} p_i \ln p_i$, where N is the total number of atoms and p_i is the projection of the coordinates onto the particular principal component mode. A plot of e^S as a function of mode number is shown in Figure 11. The ∼600 modes which contribute to the Boson peak each span more than 4000 atoms indicating their highly collective nature.

The present analysis indicates that protein−protein interactions do not directly contribute to the Boson peak. However, their presence is essential for reproduction of the Boson peak at the experimental frequency. Simulations with a single protein molecule shift the peak to lower frequencies. The presence of protein−protein interactions improves the description of the experimental (powder) environment while rigidifying the collective vibrations that give rise to the Boson peak, shifting them to higher frequencies. As can be seen from Figure 7, the intensity of the neutron scattering structure factor calculated from an MD simulation of a single protein in an orthorhombic box is higher than that obtained from the single unit-cell monoclinic crystal simulations. Also, the root-mean-square fluctuation of the protein in the orthorhombic box and in a crystal environment was found to be 0.38 Å and 0.35 Å, respectively. This indicates greater flexibility of the protein molecule in the absence of crystal contacts. The absence of intermolecular interactions softens the modes that give rise to the Boson peak, and thus

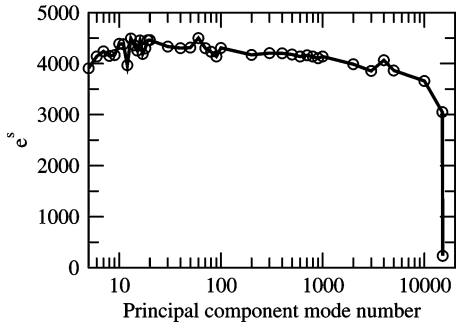

Figure 11. Value of e^s, a measure of the number of atoms that participate in a principal component mode as a function of principal component mode number.

(58) Lamy, A.; Souaille, M.; Smith, J. C. *Biopolymers* **1996**, *39*, 471−478.
(59) Go, N.; Noguti, T.; Nishikawa, T. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 3696−3700.
(60) Brooks, B. S.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 6571−6575.
(61) Hess, B. *Phys. Rev. E* **2000**, *62*, 8438−8448.
(62) Hess, B. *Phys. Rev. E* **2002**, *65*, 031910.
(63) Faraldo-Gómez, J. D.; Forrest, L. R.; Baaden, M.; Bond, P. J.; Domene, C.; Patargias, G.; Cuthbertson, J.; Sansom, M. S. P. *Proteins: Struct., Funct., Bioinf.* **2004**, *57*, 783−791.
(64) Yu, X.; Leitner, D. M. *J. Phys. Chem. B* **2003**, *107*, 1698−1707.

single molecule simulations hitherto performed have not been able to reproduce the experimental Boson peak frequency.

The experimentally observed Boson peak is smoother than the Boson peak calculated from MD simulations. However, as seen in Figure 2 (thickest line), increasing the number of molecules in the primary simulation box leads to a neutron scattering spectrum that is closer to the experimentally observed Boson peak. This observation further emphasizes the importance of an accurate description of the environment in describing low-frequency internal protein dynamics.

4. Concluding Remarks

Here, MD simulations and normal-mode analysis have been used to examine low-frequency motions in a protein at 150 K. An accurate description of the environment is found to be essential for reproducing the Boson peak (measured on powders) at the experimentally observed frequency. The Boson peak motions are identified to be collective and harmonic at the temperature investigated here. Earlier calculations[37] suggested that the Boson peak arises from motions distributed throughout the protein, consistent with our analysis here. The anharmonic, multiminimum modes that contribute most to the root-mean-square fluctuations, and thereby to the dynamical transition in proteins, do not contribute to the Boson peak. The inelastic neutron scattering spectrum obtained from normal-mode analysis also reproduces the Boson peak providing further evidence for the harmonic origin of the associated dynamics.

Of particular interest is the identification here of intermolecular vibrations at ultralow frequency (below 1 meV) in the dynamic structure factor also observed in the experimental spectrum of Plastocyanin.[57] The presence of an intermolecular vibration manifested as a clear peak in the low-frequency dynamic structure factor opens up the possibility of using neutron scattering spectroscopy as a tool for investigating interprotein vibrations. However, these interactions have very low frequencies and hence require instruments of very high energy resolution. *Polycrystalline* powders would be expected to have similar contacts as in the case of the crystal simulations and relatively sharp protein interaction vibrational lines. In contrast, *amorphous* powder may have a wider range of contact modes and thus a broader interaction peak. The extent of this broadening is of fundamental interest and may provide information on the relation between the properties of protein contact interfaces and interaction strengths. Investigating the temperature and hydration dependence of these ultralow-frequency modes should provide significant insight into the nature of protein−protein interactions. These experimental studies are in progress.

Acknowledgment. The authors acknowledge the financial support from the Deutsche Forschungsgemeinschaft. V.K. acknowledges support through the Olympia Morata fellowship and thanks Dr. K. Moritsgu, Dr. R. Siebert, and Dr. T. Becker for useful discussions and L. Meinhold for the program to project MD trajectories onto the principal component modes.

Supporting Information Available: Complete ref 43. This material is available free of charge via the Internet at http://pubs.acs.org.

JA055962Q