Structure-Based Analysis of Protein Dynamics: Comparison of Theoretical Results for Hen Lysozyme With X-Ray Diffraction and NMR Relaxation Data

Turkan Haliloglu and Ivet Bahar*
Polymer Research Center and Chemical Engineering Department, Bogazici University, and TUBITAK Advanced Polymeric Materials Research Center, Istanbul, Turkey

ABSTRACT **An analytical approach based on Gaussian network model (GNM) is proposed for predicting the rotational dynamics of proteins. The method, previously shown to successfully reproduce X-ray crystallographic temperature factors for a series of proteins is extended here to predict bond torsional mobilities and reorientation of main chain amide groups probed by ^{15}N-H nuclear magnetic resonance (NMR) relaxation. The dynamics of hen egg-white lysozyme (HEWL) in the folded state is investigated using the proposed approach. Excellent agreement is observed between theoretical results and experimental (X-ray diffraction and NMR relaxation) data. The analysis reveals the important role of coupled rotations, or cross-correlations between dihedral angle librations, in defining the relaxation mechanism on a local scale. The crystal and solution structures exhibit some differences in their local motions, but their global motions are identical. Hinge residues mediating the cooperative movements of the α- and β-domains are identified, which comprise residues in helix C, Glu35 and Ser36 on the loop succeeding helix B, Ile55 and Leu56 at the turn between strands II and III. The central part of the β-domain long loop and the turn between strands I and II display an enhanced mobility. Finally, kinetically hot residues and key interactions are identified, which point at helix B and β-strand III as the structural elements underlying the stability of the tertiary structure. Proteins 1999;37:654–667.** © 1999 Wiley-Liss, Inc.

Key words: Gaussian network model; NMR order parameters; crystallographic temperature factors; coupled motions; hen egg-white lysozyme

INTRODUCTION

Although a unique conformation characterizes each protein in the native state under physiological conditions, it has become clear, early by simulations and experiments, that the folded state is not a static one, in a strict sense, but enjoys a multitude of microstates due to its vibrational degrees of freedom and the occasional jumps occurring between the bonds' rotational states. These motions are highly localized in general, so that the overall tertiary fold is preserved. Yet, a so-called dynamic equilibrium exists. The motions in the folded state are conveniently expressed as fluctuations about the mean (native state) positions. Such fluctuations are routinely observed in molecular dynamics (MD) simulations[1–3] and normal mode analyses,[4–6] or in experiments in the form of X-ray crystallographic temperature factors,[7–9] also referred to as the Debye-Waller or B-factors, and order parameters from nuclear magnetic resonance (NMR) relaxation.[10–12]

A simple analytical approach, referred to as *Gaussian network model* (GNM) has been recently proposed[13,14] as a simple tool for analyzing the preferred motions near the folded state, and understanding the dominant factors that control these motions. The remarkable agreement between the mean square (ms) fluctuations of residues predicted by the GNM and those revealed by X-ray crystallographic B-factors, suggests that the two ingredients of the theory — local packing density and geometry of tertiary contacts between residues — are the most important factors that dominate the observed behavior. Interestingly, in the theoretical model, no account of the specificities of inter-residue interactions, or anharmonicities of molecular potentials is taken. These properties underlie the choice of the unique native conformation for each sequence and cannot be ignored when studying protein folding/unfolding dynamics. Yet, in so far as the equilibrium dynamics of proteins near their the folded state is concerned, numerous applications[13–20] lend support to the applicability and utility of the GNM.

In the present study, the GNM approach will be extended to investigate the rotational and orientational behavior of individual bonds. The information provided on proteins' local motions by X-ray diffraction and NMR relaxation experiments, on the one hand, and by the theory (GNM), on the other, will be analyzed from a comparative viewpoint. These will be shown to contain complementary information.

Essentially, the two experimental techniques and the theory refer to different environmental conditions. X-ray crystallography characterizes the structure and fluctua-

TABLE I. Hen Lysozyme Structures Analyzed in the Present Study

PDB code	Structure type	Resolution (Å)	Reference
2lzt	X-ray (triclinic)	1.97	44
193l	X-ray (tetragonal)	1.33	45
1hwa	NMR	—	36

tions in the crystalline state; it is an invaluable tool for determining the mean coordinates in the folded state; but the deviations from these positions, expressed as temperature factors, should be interpreted with care, in that they may be biased by static disorder effects, conformational variability, and intermolecular contacts occurring in the crystalline form.[7–9] NMR relaxation, on the other hand, describes the behavior in solution. Proteins in solution undergo a wide variety of motions ranging from *local* atomic fluctuations and bond torsional librations to rotational jumps, cooperative helix-coil transitions, hinge bending movements of subdomains, and finally *global* unfolding. NMR spectroscopy has proven useful in providing information, both at the scale of individual bonds (spin relaxation experiments in proton, ^{13}C or ^{15}N-H NMR), and on the global folding/unfolding (two-dimensional NMR coupled with H/D exchange and rapid mixing techniques). Here we concentrate on the ^{15}N-H NMR relaxation data. Finally, GNM refers to the vibrational dynamics of a *single* molecule (or an isolated complex) in the absence of intermolecular effects and nonlinear contributions. The protein is modeled therein as a perfectly elastic network. The GNM results may thus be viewed as an *idealization*. They reflect the effect of intramolecular vibrational entropy, exclusively.[16] In this respect, GNM serves as a tool for discriminating entropic contributions in the observed dynamics.

In the following section, the physical quantities reflected by X-ray B-factors and NMR order parameters will be recapitulated, along with the extension of the GNM to presume these data. The application of the GNM to temperature factors has been illustrated in several studies, and will not be elaborated here. The extension to the analysis of internuclear reorientations is new, and will be presented in some details. Calculations will be performed for the hen egg-white lysozyme (HEWL) structures listed in Table I. Some correlations between the order parameters and the B-factors have been observed in previous studies of HEWL. See for example the comprehensive analysis of Dobson and collaborators.[21] Surface accessibility and thereby lack of intramolecular van der Waals contacts has been suggested therein to be a major determinant of mobility. Here, similarities and differences between the dynamics of the solution and crystal structures will be analyzed using the GNM. The mechanism of cooperative motions driven by the global modes in the folded state will be determined. The major determinant of the dynamics — in the subnanoseconds scale of motions presently studied — will emerge as the *geometry* of the three-dimensional structure, or the entropic effects, rather than the specific energetics.

DYNAMIC CHARACTERISTICS FROM X-RAY CRYSTALLOGRAPHY AND NMR RELAXATION

The X-ray crystallographic B-factors provide information on the mobilities (thermal motions) of atoms in the crystalline state. In the absence of nonlinear effects, the B-factors scale with the mean-square fluctuations of $<(\Delta R_i)^2>$ of atoms, as

$$B_i = 8\pi^2 <(\Delta R_i)^2> / 3 \qquad (1)$$

Here the subscript i refers to the serial index of the particular atom. $\Delta \mathbf{R}_i$ is the fluctuation vector from the mean position $<\mathbf{R}_i>$ of atom i in the crystal structure, and $(\Delta R_i)^2$ is the scalar product $(\Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_i)$. The B-factors thus provide an information on the *translational* mobility of individual atoms in the equilibrium state. In practice, atoms are subject to residue-specific, anharmonic potentials, and thus nonlinear effects operate.[22–24] Yet, the observed Debye-Waller factors of folded proteins can be approximated by unimodal quasi-harmonic motions.[25]

The *rotational* mobilities of the individual bonds, on the other hand, are probed by NMR relaxation experiments. ^{15}N-H NMR relaxation in particular, examine the time evolution of the amide bonds' reorientations. The *time-delayed orientational correlation function* measured in NMR is expressed by the second-order Legendre polynomial

$$M_2(t) = 3/2 <(\mathbf{m}_i(0) \cdot \mathbf{m}_i(t))^2> - 1/2$$
$$= 3/2 <\cos^2\Delta\alpha_i(t)> - 1/2 \quad (2)$$

Here $\mathbf{m}_i(t)$ is the unit vector along the examined internuclear bond (here the ith N-H bond), t is the time interval between two successive observations, $\Delta\alpha_i(t)$ is the angle between the vectors $\mathbf{m}i(0)$ and $\mathbf{m}_i(t)$, and the angular brackets denote the ensemble average over a multitude of snapshots taken with the fixed time interval t. $M_2(t)$ decays from 1 at $t = 0$ to the order parameter

$$S_i^2 = M_2(\infty) = 3/2 < \cos^2\Delta\alpha_i(\infty)> - 1/2 \qquad (3)$$

at long times. In the presence of an isotropic tumbling $<\cos^2\Delta\alpha_i(\infty)>$ assumes the value of 1/3 corresponding to free (random) reorientation, and $S_i^2 = 0$. On the other hand, if the motion is restricted to fluctuations in the neighborhood of the native conformation, $S_i^2 \neq 0$. The value of S_i^2 depends in this case on the square cosine of the amplitude of the angular fluctuations of the ith internuclear vector. The uppermost limit $S_i^2 = 1$ corresponds to fully constrained bonds.

The relaxation of the main chain amide bonds in ^{15}N-H NMR experiments is predominantly controlled by the torsional motions of the backbone bonds. These motions may be accomplished as coupled or isolated rotational jumps, and/or by librations (fluctuations in torsional angles). Librations may in turn be cooperative, or random. A torsion of $\Delta\alpha_i$ undergone by the backbone bond $\mathbf{l}_i$ rotates
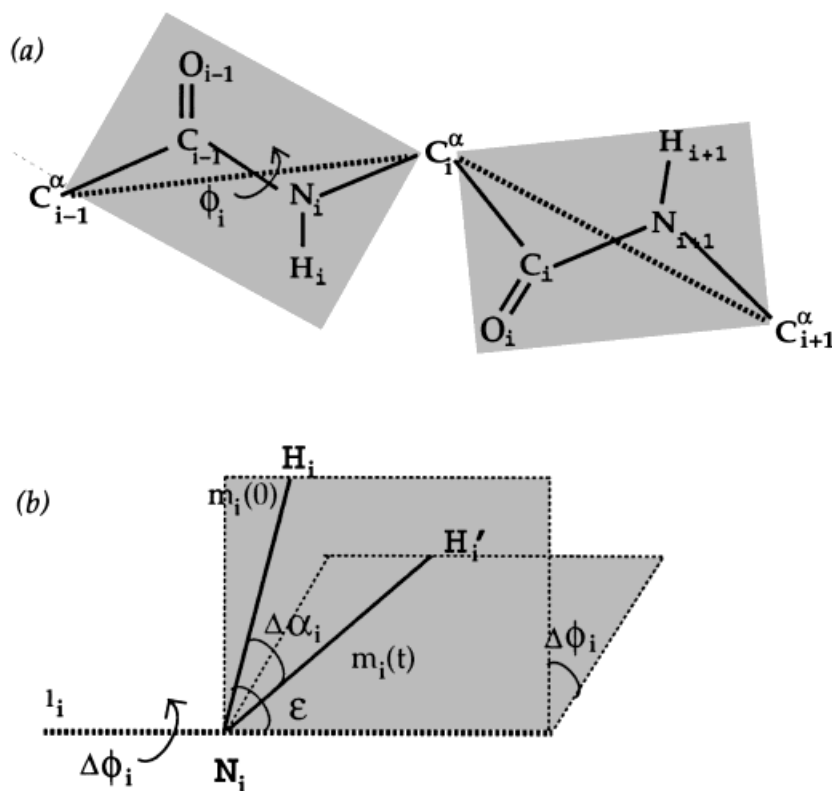
Fig. 1. **a:** Schematic representation of a protein segment, illustrating the virtual bond model of the backbone (dashed lines), and the position of the ith amide bond N_i-H in the plane comprising the three consecutive backbone bonds C^α_{i-1}—C_{i-1}, C_{i-1}—N_i and N_i—C^α_i. The ith virtual bond vector, $\mathbf{l}_i$, points from C^α_{i-1} to C^α_i. Its rotational angle is designated as ϕ_i. The ith amide bond vector is denoted as $\mathbf{m}_i$. PDB structures show that $\mathbf{m}_i$ makes an angle of $103° \pm 2°$ with $\mathbf{l}_i$. **b:** Diagram showing that a torsional rotation of $\Delta\phi_i$ undergone by the ith virtual bond induces an angular change of $\Delta\alpha_i$ in the orientation of bond $\mathbf{m}_j$. $\Delta\alpha_i$ and $\Delta\phi_i$ are related by Eq. (4), if the angle between $\mathbf{l}_i$ and $\mathbf{m}_i$ is fixed and equal to ϵ.

the side bond $\mathbf{m}_i$ affixed to the ith atom by an angle $\Delta\alpha_i = \Delta\phi_i$ provided that $\mathbf{l}_i$ and $\mathbf{m}_i$ are perpendicular to each other, and by $\Delta\alpha_i = 0$ when these ($\mathbf{l}_i$ and $\mathbf{m}_i$) are collinear. In the present virtual bond representation of protein backbone $\mathbf{m}_i$ is almost perpendicular to $\mathbf{l}_i$. Here l_i is the virtual bond connecting the α-carbons $i - 1$ and i, and ϕ_i is its torsional angle, and $\mathbf{m}_i$ is the ith N-H bond rigidly embedded in the plane defined by the peptide bond and amide group (see Fig. 1). In the more general case, geometric considerations lead to

$$\cos \Delta\alpha_i = 1 - \sin^2\epsilon \, (1 - \cos \Delta\phi_i) \quad (4)$$

if the angle between $\mathbf{m}_i$ and $\mathbf{l}_i$ is equal to ϵ and remains fixed during the motion. $<\cos^2\Delta\alpha_i>$ may then be written in terms of the *ms torsional autocorrelations* $<\Delta\phi_i^2>$ as

$$<\cos^2\Delta\alpha_i> \approx (1 - \sin^2\epsilon)^2$$
$$+ \sin^4\epsilon \, [1 - <\Delta\phi_i^2> + (5/9) < \Delta\phi_i^2>^2] \quad (5)$$

using the identity $<\cos\Delta\phi_i> = 0$, the serial expansion of $<\cos^2\Delta\phi_i>$ truncated beyond the terms fourth order in $\Delta\phi_i$, and the Gaussian approximation $<\Delta\phi_i^4> = (5/3) <\Delta\phi_i^2>^2$. The above expression for the reorientation of the internuclear vector $\mathbf{m}_i$ is based on the rotation of the backbone bond $\mathbf{l}_i$, only. For an accurate estimation of $<\cos^2\Delta\alpha_i>$, one should, however, consider the contribution of all bond torsions that might affect the reorientation $\Delta\alpha_i$ of $\mathbf{m}_i$. The nearest neighbors along the sequence have a dominant

effect, and the effect of the farther bonds is diminishingly small. The contribution of the bond $\mathbf{l}_j$ ($j \neq i$) is important to the extent that this bond torsion is correlated with that of $\mathbf{l}_i$. As shown below, the GNM lends itself to efficient calculation of the auto- and cross-correlations in bond torsions. The first and second nearest neighbors are found to exhibit the highest correlations. In accordance with this, the rotational motions of the bonds $\mathbf{l}_j$ ($|j - i| \geq 2$) will be taken into consideration when estimating S_i^2 (see the Appendix for more details).

As pointed out above, the B-factors and the order parameters primarily depend on $<\Delta R_i^2>$ and $<\Delta\phi_i^2>$, respectively. See Eqs. (1), (3), and (5). Thus, two different types of mobilities, translational and rotational, are reflected by the respective data. Usually, rotational motions are accompanied by translational motions, hence the observed correlation between the B-factors and order parameters. However, some regions may exhibit strongly coupled rotations that efficiently localize the motion, and leave the residue coordinates almost unchanged. Conversely, there may be translational motions that leave the reorientation of the internuclear vector unchanged. Such subtle differences between the two sets of data will be explored here with the help of the GNM.

Analytical Approach: Extension of GNM to Analyze Rotational Motions

In the GNM, the interactions between residues in close proximity are replaced by linear springs, in analogy with the elasticity theory of random polymer networks.[26] The

protein in the folded state is, therefore, assumed to be equivalent to a three-dimensional elastic network. The junctions of the network are identified with the α-carbons, and the springs by the single-parameter (γ) harmonic potentials between residues, following the original proposition of Tirion.[27] The junctions undergo Gaussian fluctuations about their mean positions. The equilibrium correlation between the fluctuations $\Delta \mathbf{R}_i$ and $\Delta \mathbf{R}_j$ of the α-carbons i and j is found from

$$<\Delta \mathbf{R}_i \bullet \Delta \mathbf{R}_j> = (3\ k_B T/\gamma)[\Gamma^{-1}]_{ij} \qquad (6)$$

where Γ^{-1} is the inverse of the Kirchhoff matrix Γ characteristic of the particular structure, the subscript ij designates the ijth element of the matrix in square brackets, k_B is the Boltzmann constant, T is the absolute temperature. Γ is an $n \times n$ matrix whose ijth off-diagonal element is -1 if α-carbons i and j are in contact ($|\mathbf{R}_i - \mathbf{R}_j| < 10$ Å), and zero otherwise.[13] The ith diagonal element of Γ is equal to the negative sum of the other elements in the same column. $<\Delta \mathbf{R}_i \bullet \Delta \mathbf{R}_j>$ may be decomposed into $n - 1$ terms, each representing the contribution of a mode k, i.e.,

$$<\Delta \mathbf{R}_i \bullet \Delta \mathbf{R}_j> = (3\ k_B T/\gamma)\left[\sum_k (1/\lambda_k)\ \mathbf{u}_k\ \mathbf{u}_k^T\right]_{ij} \qquad (7)$$

Here λ_k is the kth eigenvalue of Γ, and $\mathbf{u}_k$ is the kth eigenvector, and the summation is carried out in the range $2 \leq k \leq n$, omitting the first eigenvalue which is identically zero. The slowest, or largest amplitude, mode of motion in the folded state, also referred to as the *global* mode, is accounted for by the leading term $(1/\lambda_2)\mathbf{u}_2\mathbf{u}_2^T$.[14,16–18]

On the other hand, the fluctuations may be written as a linear combination of the changes in rotational angles $\Delta \phi_j$, as[28,29]

$$\Delta \mathbf{R}_i = \sum_{j=1}^{i-1} \mathbf{a}_{ij}\ \Delta \phi_j \qquad (8)$$

where $\mathbf{a}_{ij} \equiv \partial \Delta \mathbf{R}_i / \partial \Delta \phi_i$ is a vector for the passage from torsional angles to space coordinates, given by

$$\mathbf{a}_{ij} = \left[\prod_{k=0}^{j-1} \mathbf{T}_k\right] \mathbf{T}_j' [\mathbf{E}\ 0] \left[\prod_{k=j+1}^{i-1} \mathbf{G}_k\right] \begin{bmatrix} \boldsymbol{l} \\ 1 \end{bmatrix} \qquad (9)$$

Here $\mathbf{E}$ is the identity matrix of order 3, $\mathbf{T}_i$ is the transformation matrix commonly used in the statistical mechanics of polymers for expressing vectorial or tensorial quantities of the bond-based frame $i + 1$ in frame i,[30] $\mathbf{T}_i' \equiv \partial \mathbf{T}_i / \partial \phi_i$ is the matrix whose elements are found by taking the derivative of the elements of $\mathbf{T}_i$ with respect to ϕ_i. $\boldsymbol{l}$ is the bond vector $col(l\ 0\ 0)$ for C^α-C^α virtual bonds in the ith bond-based frame, and $\mathbf{G}_k$ is the 4×4 generator matrix given by[30]

$$\mathbf{G}_k = \begin{bmatrix} \mathbf{T}_k & \boldsymbol{l} \\ 0 & 1 \end{bmatrix} \qquad (10)$$

The starting point for deriving Eqs. (8) and (9) is to express the position vector of the ith α-carbon as $\mathbf{R}_i = [\mathbf{E} + \mathbf{T}_1 + \mathbf{T}_1\mathbf{T}_2 + \ldots + (\mathbf{T}_1\mathbf{T}_2 \ldots \mathbf{T}_{i-1})]\ \boldsymbol{l}$, and write $\Delta \mathbf{R}_i$ as a function of $\Delta \phi_k$ for $1 \leq k \leq i - 1$, after differentiating the transformation matrices with respect to the torsional angles. Using Eq. (8), the correlation $<\Delta \mathbf{R}_i \bullet \Delta \mathbf{R}_j>$ reads

$$<\Delta \mathbf{R}_1 \bullet \Delta \mathbf{R}_j> = \sum_{k=1}^{i-1} \sum_{l=1}^{j-1} [\mathbf{a}_{i,k} \bullet \mathbf{a}_{j,l}] <\Delta \phi_k \Delta \phi_l> \qquad (11)$$

In the present study, $<\Delta \mathbf{R}_i \bullet \Delta \mathbf{R}_j>$ are evaluated using the GNM and Eq. (11) is used in a recurrence scheme (see Appendix) to extract the torsional auto-and cross-correlations, $<\Delta \phi_k^2>$ and $<\Delta \phi_k \Delta \phi_l>$.

RESULTS AND DISCUSSION
Comparison of X-Ray Diffraction and Theoretical Results for HEWL

Figure 2 displays the comparison of the theoretical (solid curves) and experimental (dotted curves) results for the temperature factors of two HEWL crystal structures. Figure 2a refers to the triclinic crystal form, 2lzt (Table I); whereas Figure 2b refers to the tetragonal crystals, 193l (Table I). The correlation coefficients between experimental and theoretical results are 0.60 and 0.71, in the respective parts (Fig. 2a and 2b). A better agreement between theory and experiments is therefore observed for tetragonal crystals (Fig. 2b) compared to that for triclinic crystals (Fig. 2a). 193l is a higher resolution structure. The higher precision of the C^α-coordinates, which are used as input in the GNM calculations, along with the higher precision of the experimental B-factors, may have contributed to the closer agreement between theory and experiments.

Calculations performed for five other high resolution (2.0 Å or higher) lysozyme structures (PDB entries: 1at6, 1hel, 1lz1, 2lym, and 1aki) confirmed that comparable correlation coefficients are obtainable in other lysozymes, as well. The single-parameter, γ, of the theory [see Eq. (6)] was adjusted in all cases to match the absolute values of B-factors. The choice of γ for a given protein does not affect the correlation coefficient between experiments and theory because γ rescales all residue fluctuations as a whole, without affecting their relative values. γ varies in the range 0.26 ± 0.05 kcal/(mol. Å²) in the examined seven lysozyme structures, using an inter-residue interaction cut-off distance of 10 Å at 300°K.

The theoretical curves in Figure 2a and b obey similar qualitative features. The two HEWL structures 2lzt and 193l are superimposable with a root mean square deviation (RMSD) of 0.65 Å, in the α-carbon coordinates. This explains the similarity of the two theoretical curves which are calculated on the basis of α-carbon coordinates, exclusively.

The two sets of experimental data, on the other hand, exhibit substantial departure. Differences between the behavior of tetragonal and triclinic HEWL crystals were also seen in H/D exchange experiments. These were attrib-
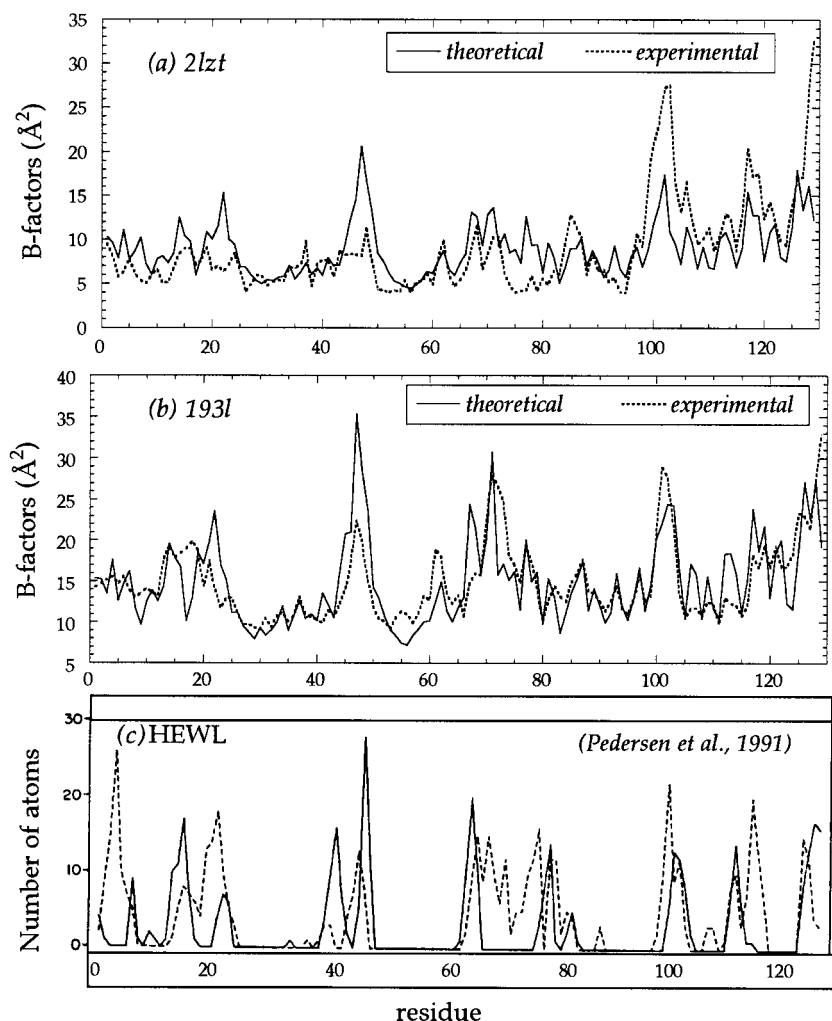
Fig. 2. **a:** Comparison of the X-ray crystallographic temperature factors for the triclinic crystals of HEWL[44] (dotted curve) with theoretical results calculated here using the GNM (solid curve). **b:** Comparison of the theoretical (solid) and experimental[45] (dotted) B-factors for the tetragonal crystal structure. The two structures have PDB codes 2lzt and 193l, respectively. Their α-carbon coordinates deviate by an RMS value of 0.65 Å, only. The theoretical curves in (a) and (b), calculated on the basis of α-carbon coordinates using the GNM, exhibit close similarities; whereas the experimental data differ mainly because of the different intermolecular contacts in the different crystal lattices illustrated in (c). **c:** Number of atoms making intermolecular contacts in the triclinic (dashed) and tetragonal (solid) crystals of HEWL, calculated by Pedersen et al.[31]

uted to the occurrence of different intermolecular contacts in the two crystalline forms.[31] Pedersen et al.[31] did a careful analysis of the intermolecular contacts in the two crystal forms. They identified the regions exhibiting contacts in excess of those in the isolated protein for both the triclinic and tetragonal crystals. The results of Pedersen et al. for the two respective crystal forms are displayed by the dashed and solid curves in Figure 2c. The peaks therein are the regions subject to intermolecular contacts associated with crystal packing. Their fluctuations are in principle dampened by these intermolecular interactions, whereas such effects are absent in the isolated molecule. The comparison of the curves in Figure 2c with those in Figure 2a and b reveals that there is a correlation between the regions subject to intermolecular constraints, and those observed in X-ray diffraction experiments to have fluctuations relatively more suppressed than theoretical predictions. Examples are the residues intervals 4–8, 20–24, 45–49, 68–83 in Figure 2a, 46–49, 66–69, 105–107, 112–115 in Figure 2b. There are a few other regions, such as the C-terminal part ($i \geq 100$) in the triclinic form of HEWL, or the residues 14–18, 80–81 in the tetragonal form, which do

not exhibit a suppression in mobility. We note that some fluctuations may be underestimated in the theory, because the single parameter γ is so adjusted to match the average fluctuation amplitude observed in experiments, ignoring that the latter may be lowered by crystal contacts.

The improved agreement between theory and experiments for the higher resolution structure, and the fact that a majority of the differences originates in intermolecular contacts — that are absent in physiological environment — benchmarks the applicability of GNM for evaluating the ms fluctuations $<\Delta \mathbf{R}_i^2>$. This type of agreement, and its further improvement upon consideration of higher resolution structure coordinates, also conform with our previous observations.[19]

Order Parameters: Comparison of NMR and Theoretical Results

The dynamics of HEWL have been examined by several NMR studies for the last two decades, ranging from proton relaxation measurements for detecting tyrosine ring flips,[32] to H/D exchange studies of local motion[31] or folding pathways.[33–35] Here we focus on the ^{15}N-NMR relaxation
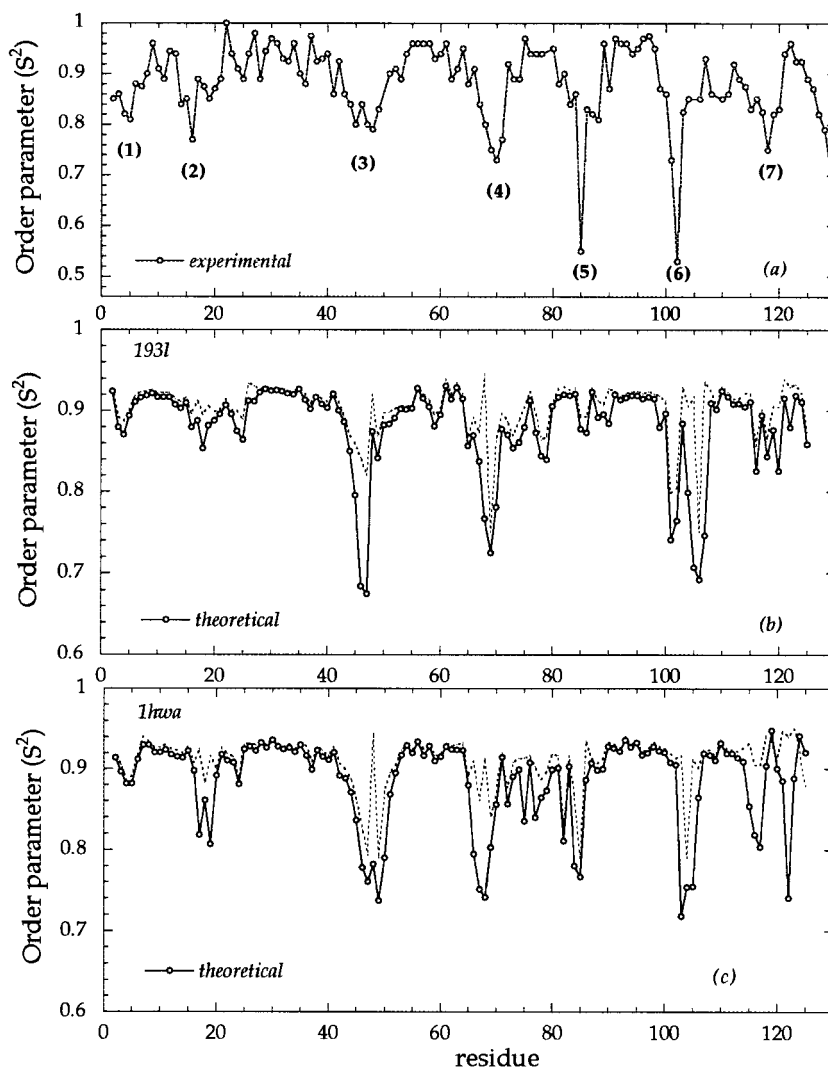
Fig. 3. **a:** Order parameters for main chain amide groups measured in ^{15}N-NMR relaxation experiments by Buck et al.[21] (dotted curve). See Figure 4 for the regions observed in NMR to be relatively disordered ($S^2 \leq 0.85$), labeled as 1–7 in the figure. **b and c:** Theoretical results obtained from Eqs. (5) and (7)–(10) for the tetragonal crystal (193l), and to the NMR structure (1hwa) determined by Smith et al.[36] The solid lines with the open results are calculated on the basis of auto- and cross-correlations between torsional motions; the dotted curves display the contribution of auto-correlations, exclusively.

data for main chain amide groups measured by Buck et al.[21] The experimental results are displayed in Figure 3a, and the results from our calculations in Figure 3b and 3c. The solid curve in Figure 3b is calculated using the tetragonal crystal structure (193l) coordinates, the one in Figure 3c using the solution structure determined[36] by NMR (1hwa), using the same force constants as used for the X-ray structure. The respective correlation coefficients are 0.65 and 0.60, after omitting a few outliers. The dotted curves in Figure 3b and 3c represent the theoretical results in the absence of coupling between adjacent bond rotations, i.e., on the basis of the rotational motions of bond $\mathbf{l}_i$, alone, without including the effect of the neighboring bonds' torsions on the reorientation of $\mathbf{m}_i$. The solid curves include the contributions of first and second neighbors.

The RMSD between the α-carbon coordinates of the NMR and crystal structures is 1.80 Å. This difference is large enough to affect the theoretical results, hence the differences in the theoretical curves shown in Figure 3b and 3c. Thus, the dynamic characteristics of the crystal

and solution structure, predicted by the theory, are not necessarily the same. Except for the C-terminal part, the behavior predicted for the solution structure conforms more closely with NMR experiments, which are indeed carried out in solution.

It is worth pointing out that the theoretical curves obtained here agree more closely with experimental data than do the curves from atomic molecular dynamics (MD) simulations.[37] This signals that the present approach incorporates the major determinants of the observed dynamics, namely the local packing density and the exact distribution, or geometry, of tertiary contacts. Furthermore, the most probable conformational fluctuations near the folded state are *analytically* calculated here, and therefore the results do not suffer from conformational sampling insufficiency or dependence on the duration of simulations, which may be incurred in atomic simulations. The limitation of the GNM, on the other hand, is its validity in the neighborhood of the native fold, exclusively; whereas partial unfolding, or disruption of secondary
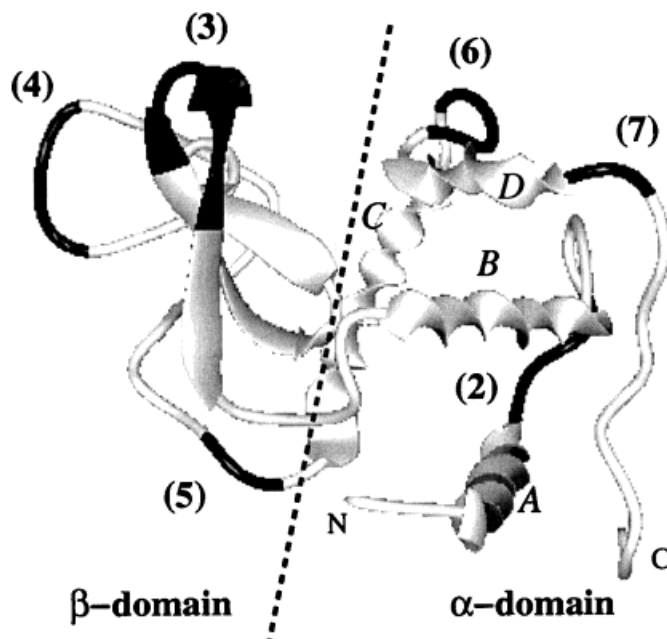
Figure 4. Ribbon diagram of HEWL. The β- and α-domains are separated by a dashed line for visual clarity. The β-domain comprises residues 36–84; the α-domain consists of two discontinuous segments 1–35 and 85–129, including four α-helices A-D. The regions observed in ^{15}N-H NMR relaxation to be relatively disordered are shown in black and indicated by the labels (2)–(7). See the labels in Figure 3a. Region (1), located at the N-terminus of helix A, is not labeled for clarity. The respective residues (shown in black) are 16–19 (2), 45–50 (3), 67–70 (4), 85–86 (5), 102–106 (6), and 116–117 (7).

structure such as unwinding of helices, can be examined by MD.

The profile of order parameters is pointed out to be predominantly determined by the variations in the longitudinal (T_1) and transverse (T_2) relaxation times.[21] In particular, the reciprocal relaxation times, $1/T_1$ and $1/T_2$, of individual residues, usually presented in the form of R-exchange parameters, are found to follow exactly the same qualitative pattern as the order parameters. This is consistent with the fact that $1/T_1$ and $1/T_2$ scale directly with the orientational correlation time τ in the extreme narrowing limit, and the fact that N-H bonds exhibiting small order parameter values within the experimental time window are generally those undergoing relatively fast reorientation processes.

Let us examine the results shown in Figure 3 in more details. There are seven regions exhibiting a distinct loss of orientation ($S^2 \leq 0.85$) within the time window of the NMR measurements, all of which are accounted for by the theory. These regions, shown in black in the ribbon diagram displayed in Figure 4, are (1) the N-terminal part of helix A; (2) the C-terminus of helix A, which is broadened in the theoretical curves to cover a portion of the loop between helices A and B; (3) the turn between strands I and II of the β-sheet domain, along with the strands termini; (4) the central part of the long loop between residues 61–78; (5) the C-terminal part of the 3_{10} helix 80–84; (6) the loop between helices C and D; and (7) the loop succeeding the helix D. On the other hand, regions having an enhanced stability on a local scale, such as the α-helices A (residues 4–15), B (24–36), C (89–99), and D

(109–115) are distinguished by their high S^2 values, both in the theory and experiments. The turn 2 (54–57) and the succeeding strand III (58–60) also exhibit reduced mobility. We note that the experiments generally exhibit smaller fluctuations in values at these regions, compared to theoretical results.

The 3_{10} helix between residues 80–84 is observed in experiments to be highly disordered, with a sharp minimum at the adjoining residue Ser85. This minimum is not obtainable by the theory when the crystal structure coordinates are used as input (see Fig. 3b). The solution structure, on the other hand, yields a sharp minimum at this precise location. The crystal and solution structures indicate a significant deviation near residues 82–86, as also pointed out by Buck et al.[21] The difference in *equilibrium* structure is shown here to be reflected on the *dynamics* as well. Bond 85 is indeed one of the three bonds of 1hwa, along with 104 and 127, which exhibit $<(\Delta\phi_i)^2>^{1/2}$ values of about 60°. The predicted fluctuations in rotational angles are displayed in Figure 5a, for the NMR (dashed) and crystal (solid) structures. Fluctuations of this size imply the tendency of the related bonds to undergo rotational transitions. This feature conforms with the results from MD simulations,[37] where frequent peptide bond transitions were observed for Ser85. In the case of the crystal structure 193l, bonds 69 and 106 are the most flexible ones.

Two other regions significantly deviating between the crystal and solution structures are the residue intervals 45–50 and 67–70. These coincide with the regions labeled 3 and 4 in Figures 3 and 4. These regions will be shown
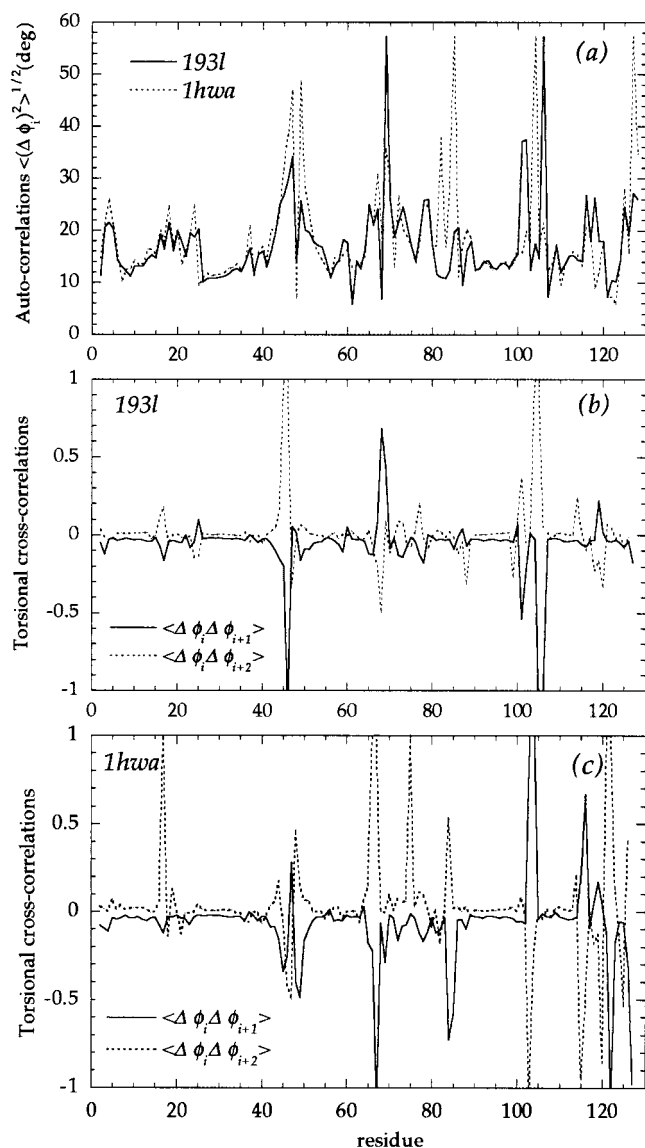
Fig. 5. a: Mean-square amplitude of torsional fluctuations, $<(\Delta\phi_i)^2>$, calculated for the virtual bonds $C_{i-1}^\alpha - C_i^\alpha$, as a function of bond index i. Results are presented for the crystal structure 193l (solid curve), and the NMR structure 1hwa (dotted curve). Note the enhanced torsional mobility of the bond 85 in the NMR structure, compared to its mobility in the crystal structure. **b:** Cross-correlations between the torsional motions of first (solid) and second (dotted) neighboring bonds as a function of bond index i, calculated for 193l. **c:** Same as (b), for the NMR structure 1hwa.

below to be engaged in large amplitude movements in the global mode of motion of HEWL. The differences in their precise equilibrium positions in the crystal or solution do not affect their dynamic characteristics in this case, as their mobilities are associated with the *global* relaxation of the molecule, rather than *local* conformational freedom.

Cooperativity Between Bond Torsional Motions

The theoretical results displayed in Figure 3 have been obtained by taking into consideration both the torsional *autocorrelations* $<\Delta\phi_i^2>$ of virtual C^α-C^α bonds, and the

cross-correlations $<\Delta\phi_i\Delta\phi_j>$ where $|j - i| \leq 2$ (see the Appendix). The contributions of the individual correlations may be clarified by separately examining the auto- and cross-correlations. Figure 5a illustrates the autocorrelations, and Figure 5b and 5c the cross-correlations for the indicated structures. The solid curves in Figure 5b and 5c represent the correlations $<\Delta\phi_i\Delta\phi_{i+1}>$ between adjacent bonds as a function of residue index i. The dashed curves display those between second neighbors, $<\Delta\phi_i\Delta\phi_{i+2}>$. These are found to affect a relatively small number of sites. Interestingly, a close correspondence emerges between the peaks observed in these curves and the regions indicated by NMR (and predicted by the theory) to be relatively disordered. Compare the minima in Figure 3 with the extrema of the curves in Figure 5b and 5c.

It should be recalled that NMR relaxation experiments probe the reorientation $\Delta\alpha_i$ of the N-H bonds, rather than the torsional angles $\Delta\phi_i$. Examination of PDB structures showed that the angle ϵ between $\mathbf{l}_i$ and $\mathbf{m}_i$ is in the range $102°–104°$, such that the reorientation $\Delta\alpha_i$ of $\mathbf{m}_i$ closely follows the torsional rotation $\Delta\phi_i$ of $\mathbf{l}_i$ [see Eq. (4)]. However, $\mathbf{m}_i$ may be reoriented in space even if $\Delta\phi_i = 0$, because of chain connectivity and the torsional mobility of other bonds along the chain. Our previous study indicated that torsional couplings are generally confined to 5–6 consecutive virtual bonds in folded structures.[38] This is equivalent to 15–18 real bonds. This range is sufficiently broad to capture most of the important correlations governing chain dynamics,[28,29] as also indicated by the present calculations. We note that in the original study of Go and Go,[4] the correlations in the fluctuations of dihedral angles in α-helices extend up to about eight residues.

Comparison of Translational and Rotational Motions

In an attempt to visualize the differences in the type of information conveyed by X-ray and NMR, we compared the normalized distributions for $<(\Delta\mathbf{R}_i)^2>$ and $<(\Delta\phi_i)^2>$ obtained by the theory. These two properties are not independent, i.e., residues exhibiting high $<(\Delta\mathbf{R}_i)^2>$ values, also enjoy an enhanced conformational freedom reflected by large $<(\Delta\phi_i)^2>$ values. Yet, rotational motions at certain bonds do not necessarily imply large amplitude displacements in the adjoining residues, particularly when these rotations occur in pairs in order to minimize the translation of the backbone. We observed such localization mechanisms in numerous studies of macromolecular motions.[28,29,38–40] Figure 5b and 5c reveal that such couplings are effective in selecting the relaxing regions.

Figure 6a shows the differences between the normalized ms fluctuations $<(\Delta\mathbf{R}_i)^2>$ and $<(\Delta\phi_i)^2>$ for 193l. Positive values refer to residues that exhibit relatively high displacements compared to their torsional motions; negative values indicate, on the contrary, the bonds that are relatively more susceptible for torsional motions. In the latter group, we notice bonds 69 and 104, which were also distinguished by their high amplitude torsional motions (see Figure 5a). In Figure 6b and 6c, on the other hand, we display in boxes the residues that undergo relatively large translational
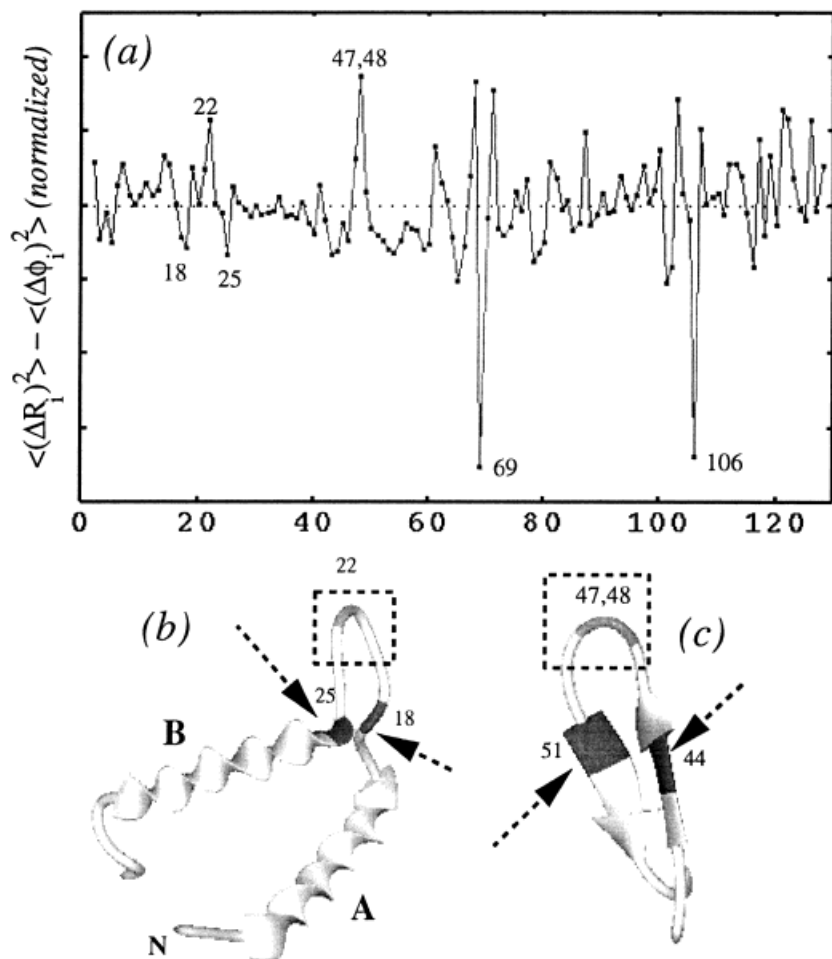
Fig. 6. **a:** Difference between the normalized $<(\Delta R_i)^2>$ and $<(\Delta\phi_i)^2>$ values, as a function of residue index i. Positive portions of the curve refer to residues whose translational mobilities, evidenced by $<(\Delta R_i)^2>$, are larger than their rotational mobilities, indicated by $<(\Delta\phi_i)^2>$; and negative regions refer to the opposite case, i.e., higher rotational freedom and more restrictive displacement in space. **b:** The loop between the helices A and B exhibits high mean-square fluctuations $<(\Delta R_i)^2>$, while bonds 18 and 25 accommodate the swinging motion of the loop by their enhanced torsional mobilities $<(\Delta\phi_i)^2>$. A similar behavior is observed in **(c)** for the turn between strands I and II of the β-sheet region. The relatively large scale displacement of the residues in the dashed box is ensured by the coupled rotations of the bonds shown by the arrows at the termini of the β-strands.

motions; whereas the residues indicated by the arrows tend to rotate rather than translate. These occur in couples in order to localize the motion of the segment (in box) enclosed between them. For example, in Figure 6b, the coupled rotations of bonds 18 and 25 apparently ensure the large scale swinging of the AB loop; whereas in Figure 6c, the translation of the turn residues is accommodated by the coupled torsions of bonds near the strands' termini.

Dominant Modes of Motion: Relevance to Function

Figure 7 displays the shapes of the dominant (slowest) modes of motion of HEWL in the folded state. These modes generally disclose the molecular mechanism relevant to biological function.[16–19] Figure 7a and 7b display the shapes for the first and second slowest modes, respectively. The respective eigenvalues are 1.25 and 2.55, i.e., the second slowest mode is about two times faster than the first. The peaks indicate the regions that exhibit the largest amplitude movements in these most cooperative (global) motions of the overall protein, whereas the minima indicate the regions acting as hinges.

The most mobile regions in the global motion of HEWL emerge in Figure 7a as (i) the central portion Gly67-Arg73 of the β-domain long loop succeeding strand III, and (ii)

residues Arg45-Thr51 comprising the turn between strands I and II, along with the two terminal residues of the strands. The high mobility of these two regions is in close agreement with the results from normal mode analysis performed at atomic level, where the largest displacement was found to occur for Thr47 and Gly71.[41] The group of residues Arg45-Thr51, shown in gray in the ribbon diagram in Figure 7b, is further distinguished as the only structural element active in the second slowest mode.

The most flexible regions in the global mode have been associated in previous GNM studies by substrate or ligand recognition sites. Interestingly, the binding epitope of lysozyme in two crystal forms of an anti-lysozyme Fab-lysozyme complex was pointed out by Davies and collaborators[42] to comprise two sites, one from Gln41 through Tyr53, and the second from Gly67 through Pro70, in exact agreement with the peaks presently identified (Fig. 7). Furthermore, the same regions are likely to be the first unfolding regions, due to their high intrinsic flexibility. The aptitude of these residues to easily depart from their equilibrium positions is also signaled by their low order parameter values. See regions labeled as 3 and 4 in Figures 3 and 4. Among the seven labeled regions exhibiting a loss of order on a *local* scale, only these two are
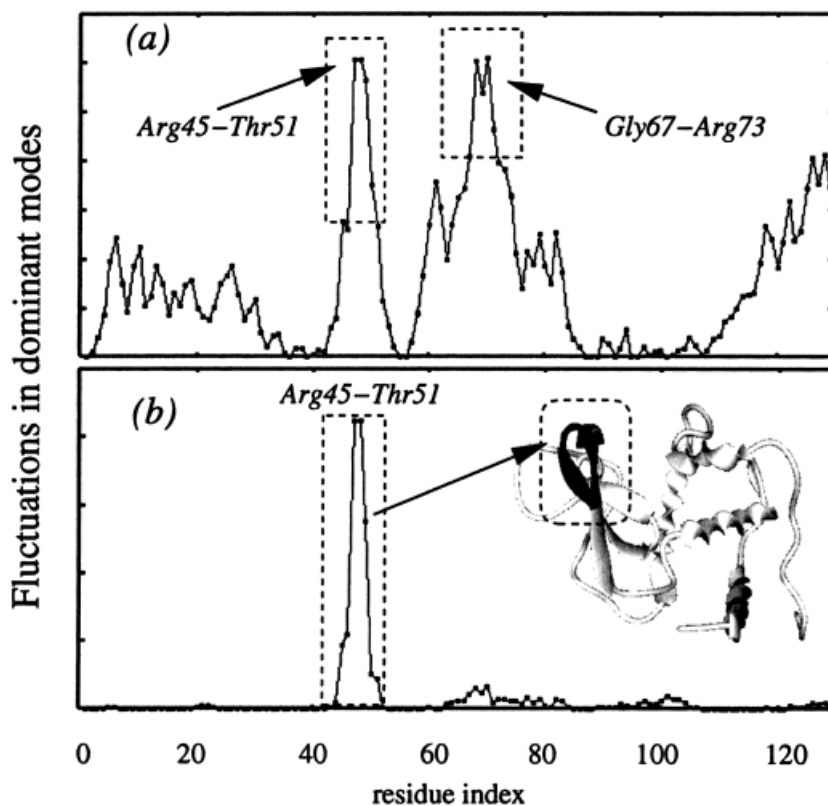
Fig. 7. Distribution of ms fluctuations in the dominant modes of motion. **a:** Displacements driven by the most cooperative (slowest/largest amplitude) mode of motion of HEWL in the folded state. Two regions of the β-domain exhibit the largest size movements: the central part (Gly67-Arg73) of the long loop (61–78) succeeding strand III, and the residues Arg45-Thr51 located at the turn between strands I and II, and at the termini of the two β-strands. **b:** the mean-square fluctuations in the next-slowest mode of motion. Therein the turn I and a few adjoining residues are moving, exclusively. This region is shown in gray in the enclosed ribbon diagram. Minima, on the contrary, indicate the regions acting as hinges in the global modes.

distinguished by the GNM to be involved in *global* motion. It is worth noting that the enhanced flexibilities of these regions are accurately predicted by the theory, in conformity with experiments, despite the fact that side chain–side chain interactions are not considered in the GNM. These results suggest that the dynamics is predominantly controlled by the architecture of the backbone, and by forces of entropic origin, in as much as no specific interaction is included in the model and all inter-residue interactions are controlled by uniform springlike potentials similarly to the elastic forces of entropic origin in random polymer networks.[16,26]

The lowermost minima in Figure 7a, on the other hand, are even more important. These play a critical role in biological activity by acting as *hinges* in the global motion. In conformity with previous studies,[5,6,41] the hinge bending sites are located at the connections or at the interfacial region between the α- and β-domains of the protein. An examination of the numerical results reveals three groups of residues forming the lowermost minima: [87–101], (35, 36, 39–42, and [55, 56]. The former includes several residues belonging to the helix C (between residues 89–99); the second lies at the connection between the α- and β-domains, and forms a small β-sheet with the N-terminal two residues; and the last group, [55, 56], is located at the turn between strands II and III. We note that the helix C was identified in the previous normal mode analysis of human lysozyme to be the bisector of the hinge mode of the protein, based on the finding that this region shows little motion in the hinge mode.[41] Furthermore, hinge points

such as Thr40, and the disulfide bridge between Cys76 and Cys94, were identified.[41] We note that Thr40 and Cys94 also lie among the group of hinge residues presently identified.

Kinetically Hot Residues

The existence of a correlation between the most tightly packed regions, which exhibit high frequency/small amplitude motions in the folded state, and the regions important for stabilizing the tertiary fold has been demonstrated and discussed in our previous work.[16,20] According to the GNM, the tightly packed residues occupy narrow wells in the free energy landscape, the curvature scaling directly with their coordination number. Consequently, they show little change in their coordinates in response to energy perturbations of a given size. Such regions resistant to structural change can be effective in maintaining the overall tertiary fold. In conformity with these studies, we focused on the shape of the high frequency motions of lysozyme. Figure 9 displays the results obtained from the superposition of the fastest two (dashed curve with filled circles) and fastest ten (solid curve with open circles) modes. The peaks are labeled for clarity. These, and particularly the highest, peaks are proposed here to be the kinetically hot residues of HEWL. Examination of Figure 3a shows that these residues remain highly ordered in NMR relaxation experiments. Their reduced mobility is consistent with their tight packing.

For comparative purposes, the regions engaged in the slowest and fastest modes are displayed in Figure 8a and
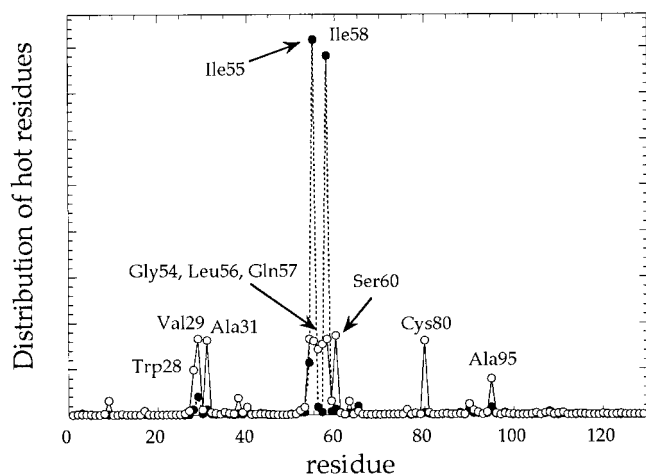
Fig. 8. Kinetically hot residues. The curve represents the shape of the highest frequency/smallest amplitude mode of motion, and the peaks are the residues engaged in these modes. Distributions are shown for the fastest two (dashed curve with filled circles) and ten (solid curve with open circles) modes.

Fig. 9. Ribbon diagrams illustrating the regions (shown in black) distinguished by **(a)** their high mobilities in the most cooperative, slowest/largest amplitude modes of motions, **(b)** their participation in the other extreme case of fastest/smallest amplitude of motions. The former group of residues are likely to easily depart from folded state coordinates; the residues in the second group, on the other hand, are expected to be intolerant to mutations.

8b, respectively. The regions shown in black in the ribbon diagram (Fig. 8a) are the most flexible residues in the global modes of the molecule (see Fig. 7). The most flexible residues of the overall protein are Thr47 and Asp48. On the other hand, the group of residues that undergo the highest frequency/smallest amplitude fluctuations (peaks in Fig. 8), are shown in black in Figure 8b. The former are likely to accommodate site mutations without an overall change in structure. The latter, on the other hand, might be disruptive upon mutating, being involved in an optimized network of interaction important for stability. In addition to the kinetically hot residues displayed in the ribbon diagram (Fig. 9b), the residues acting as hinges in the global modes might be intolerant to mutations. This sensitivity would be, this time, required for function.

Our structural analysis indicates the helix B as a structural element involved in important tertiary contacts. This helix includes the kinetically hot residues Trp28, Val29, Ala31, and the hinge-bending sites Glu35, Ser36 at the α–β interface. The interdomain contacts are predominantly achieved with the turn II and strand III residues (54–60) of the β-domain. Thus, the tertiary contacts between helix B in the α-domain, and turn II and strand III in the β-domain appear here to play a major role in stabilizing the tertiary structure. We note, furthermore, that the same residues Ile55 and Leu56 on turn II participate in the hinge-bending sites controlling the functional motions of the protein.

CONCLUSION
Major Determinant of Dynamics: Geometry as Opposed to Energetics

This study stipulates the utility of GNM, as a tool for understanding the major determinant(s) for short-time protein dynamics. GNM is a simple, but rigorous, approach that contains n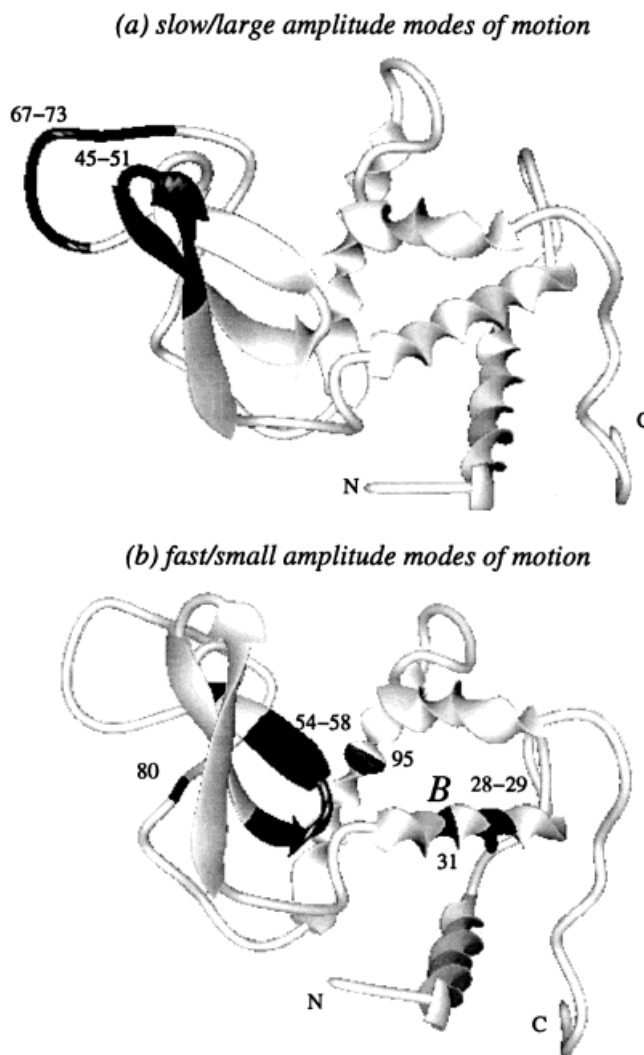o energetic interactions, or residue specificities. The only structural property considered therein is the *geometry* of the structure. The term "geometry" naturally embodies the local packing density and the unique *topology of nonbonded contacts* in the particular architecture. These are physical properties, entropic in nature. Analytical consideration of this information, however, yields surprisingly good agreements with experimental data.

In general, the discrepancy between theory and experiments can be attributed to either experimental or theoretical defects. On the theoretical side, the major simplification is that the specific interactions between side chains, which certainly lead to nonlinear effects and departures from simple harmonic oscillations, are not included. Also, no overall protein motion is included in the analysis.

Previous analysis of B-factors based on a rigid body model (of ten or more adjustable parameters describing the overall translation, rotation and screw motions of the protein) suggest that the observed fluctuations could be reflecting the distance of atoms from the centroid.[43] However, MD simulations, where rigid-body motions are removed, also exhibit the same type of increase in fluctuations with increasing distance from the centroid. The physically relevant quantity controlling the fluctuations is revealed in the present *single-parameter* theory as the local packing density, rather than the separation from the centroid. The smaller fluctuation amplitudes of interior residues is due to their higher coordination number; whereas the more loosely packed regions, such as the loops, undergo relatively larger amplitude motions.

It is interesting to note that the amide bonds belonging to α-helices are almost invariably predicted by the theory to strongly resist orientational changes. This resistance could be attributed to the high energy cost of dissociating hydrogen bonds between the amide and carbonyl groups of adjacent turns. However, hydrogen bonds are not explicitly included here in the theory. The observed resistance is simply a consequence of the tight packing of α-helical segments on a local scale. It is conceivable that such a tight packing is essentially achieved upon hydrogen bond formation, and it may be argued that the observed behavior is still indirectly controlled by hydrogen bonds. At any rate, a coarse-grained consideration of local packing density and geometry, without direct inclusion of specific interactions (hydrogen bonds) between functional groups, appears to be sufficient for accounting for the internal stability of the helical regions.

Torsional Librations versus Rotational Jumps: Does the Former Dominate the Order Parameters?

The present analysis does not directly take account of the jumps between rotational isomeric states. The sizes of residue fluctuations around their mean positions are determined in the GNM, in conformity with the basic assumption of harmonic potentials and thereby Gaussian fluctuations in inter-residue distances. And the accompanying changes in the torsional angles are evaluated from knowledge of residue mean square fluctuations. The angle changes compatible with fluctuations are thus expected to be of the type of librations, in general.

The RMS torsional autocorrelations $<(\Delta\phi_i^2)>^{1/2}$ calculated in the present study were indeed in the range $20° \pm 10°$, in general (see Fig. 5a). Yet, three bonds, 85, 104, and 127, yielded RMS changes in dihedral angles of about 60° for the solution structure (1hwa). These are likely to undergo rotational jumps because their torsional angle fluctuations are large enough to occasionally exceed the step size ($\approx 60°$) required for crossing a rotational saddle point. It is interesting to note that, in conformity with the theoretical predictions, the sharpest minima in the experimental order parameter curves were found at Ser85 and Asn103 (Fig. 3a).

The agreement between theoretical predictions and NMR data suggests that the present model-based on torsional librations and a few occasional isomeric rotations, usually occurring in couples, provides a realistic description of the process of relaxation probed in NMR relaxation experiments. It is interesting to note that MD simulations that exhibit relatively more frequent conformational transitions yielded S^2 values lower than those observed in experiments. The issues of raising the rotational barrier heights by 2–5 kJ/mol, or considering shorter (a few hundreds ps) portions of the MD trajectories, were then brought up so as to interpret the experimental data.[37] The present analysis resolves these uncertainties by determining unique, most probable values for bond rotational motions, irrespective of the length of simulations or the precise form of potential energy functions.

Local Relaxation Occurs Predominantly by Coupled Torsions

Our analysis shows that certain bonds enjoy relatively high *rotational* mobilities on a local scale despite being relatively constrained in space. Or, on the contrary certain segments can be displaced despite their rotational stiffness. One should not therefore expect a one-to-one correspondence between the $<(\Delta R_i)^2>$ and $<(\Delta\phi_i)^2>$, and the associated temperature factors and order parameters. A correlation between the experimental B-factors and ($1 - S^2$) values has been, however, observed in many studies.[21] The explanation lies in the fact that order parameters are modulated not only by $<(\Delta\phi_i)^2>$, but by two additional properties, (i) the relative orientation of the amide bond with respect to the backbone in the folded structure, accounted for by the angle ϵ_i, and (ii) the assistance of coupled, correlated rotations that effectively accommodate the motions and contribute to relaxation. And both of these two factors are closely related with the backbone architecture and connectivity, which also determine the $<(\Delta R_i)^2>$ values.

In particular, coupled rotations are shown here to have a substantial effect on the observed order parameters. These define the regions *effectively* engaged in relaxation. See for example the strong correlation between the peaks in Figure 5, indicative of the bonds undergoing highly correlated rotational motions, and the minima in Figure 4 revealing the disordered regions. Not surprisingly, the dynamic behavior of neighboring residues along the sequence is often similar.

In summary, the residues exhibiting low order parameters closely coincide with those capable of undergoing coupled rotations with their near neighbors along the sequence, and can thereby accommodate and efficiently localize a conformational disorder.

Implications about Global Dynamics and Key Residues of HEWL

Three groups of residues have been pointed out to control the global motion of the molecule: {87–98, 101}, {35, 36, 39–42}, and {55, 56}. These are all clustered near the axis separating the α- and β-domains of HEWL. The fact that a large number of C-helix residues are almost fixed in the global mode suggests that this helix essentially acts as

an axis of rotation, or a shaft, in the most cooperative motion of the protein. The β-turn residues [55, 56] are in close interaction with the B-helix residues [35, 36]. There are several interatomic contacts and/or hydrogen bonds between the side chain and/or backbone atoms of these two pairs of residues. This interdomain coupling suggests that a key role is played by these residues, or by the secondary structural elements comprising them.

We note that the β-sheet region, and particularly the turn between strands I and II, are highly mobile. Is the higher conformational freedom of β-sheet on a *local* scale, revealed in the present theoretical analysis of the dynamics of HEWL, an indication of its higher aptitude for unfolding? There is experimental evidence indicating the folding of the β-domain folding succeeds that of the α-domain. Conversely, the same domain may be expected to unfold more readily under denaturing conditions. The folding and unfolding pathways need not be the same, in general, due to the occurrence of multiple pathways associated with the ensemble of infinitely many configurations in the denatured state. However, it is very likely that the last step of folding, or the first step of unfolding, are uniquely defined, due to the uniqueness of the native state, itself. In HEWL, the two most flexible regions, Arg45-Thr51 and Gly67-Arg73 are most likely to unfold at an early step.

Finally, we should point out that the most important merit of the GNM is its simplicity and ability to extract information of the dynamics of biomolecular systems without requiring any knowledge of residue-specific empirical potentials or any computationally expensive calculations. The Fortran codes for doing the calculations presented here are accessible on the internet (http://klee.bme. boun.edu.tr).

ACKNOWLEDGMENT

We gratefully acknowledge useful discussions with Dr. B. Erman.

APPENDIX

We adopt the generalized coordinates {l_i, θ_i, ϕ_i} for describing the conformation of the ith residue. l_i is the length of the virtual bond vector $\mathbf{l}_i = \mathbf{R}_i - \mathbf{R}_{i-1}$ between C$^\alpha$ atoms $i - 1$ and i, θ_i is the angle between $\mathbf{l}_i$ and $\mathbf{l}_{i+1}$, and ϕ_i is the torsional angle of l_i. ϕ_i is calculated using $\phi_i = \cos^{-1}(\mathbf{n}_i \cdot \mathbf{n}_{i+1})$ where $\mathbf{n}_i$ is the unit vector normal to the plane formed by $\mathbf{l}_{i-1}$ and $\mathbf{l}_i$, i.e., $\mathbf{n}_i = \mathbf{l}_{i-1} \times \mathbf{l}_i / |\mathbf{l}_{i-1} \times \mathbf{l}_i|$. Bond torsions are considered as the only accessible degrees of freedom. The auto- and cross-correlations between torsional fluctuations are extracted from the correlations in residue fluctuations using Eqs (8)–(11) in a recurrence scheme the first steps of which are

$$<\Delta\phi_1\Delta\phi_1> = \frac{1}{\mathbf{a}_{21}^T\,\mathbf{a}_{21}}<\Delta\mathbf{R}_2\cdot\Delta\mathbf{R}_2> \quad (12)$$

$$<\Delta\phi_1\Delta\phi_2> =$$
$$\frac{1}{\mathbf{a}_{21}^T\,\mathbf{a}_{32}}[<\Delta\mathbf{R}_2\cdot\Delta\mathbf{R}_3> - \mathbf{a}_{21}^T\mathbf{a}_{31}<\Delta\phi_1\Delta\phi_1>] \quad (13)$$

$$<\Delta\phi_1\Delta\phi_3> = \frac{1}{\mathbf{a}_{21}^T\,\mathbf{a}_{43}}[<\Delta\mathbf{R}_2\cdot\Delta\mathbf{R}_4> - \mathbf{a}_{21}^T\mathbf{a}_{41}<\Delta\phi_1\Delta\phi_1>$$
$$- \mathbf{a}_{21}^T\mathbf{a}_{42}<\Delta\phi_1\Delta\phi_2>] \quad (14)$$

$$<\Delta\phi_2\Delta\phi_2> = \frac{1}{\mathbf{a}_{32}^T\,\mathbf{a}_{32}}[<\Delta\mathbf{R}_3\cdot\Delta\mathbf{R}_3> - \mathbf{a}_{31}^T\mathbf{a}_{31}<\Delta\phi_1\Delta\phi_1>$$
$$- \mathbf{a}_{32}^T\mathbf{a}_{31}<\Delta\phi_1\Delta\phi_2> - \mathbf{a}_{31}^T\mathbf{a}_{32}<\Delta\phi_2\Delta\phi_1>]. \quad (15)$$

The succeeding steps may be summarized by the general relationship

$$<\Delta\phi_i\Delta\phi_j> = (\mathbf{a}_{i+1,i}^T\mathbf{a}_{j+1,j})^{-1}\left[<\Delta\mathbf{R}_{i+1}\cdot\Delta\mathbf{R}_{j+1}>\right.$$
$$\left. - \sum_{k=1}^{i}\sum_{l=1}^{j}<\Delta\phi_k\Delta\phi_1>\mathbf{a}_{i+1,k}^T\mathbf{a}_{j+1,l}\right]. \quad (16)$$

The last double summation in Eq. (16) is performed over all k and l values in the range $1 \le k \le i$ and $1 \le l \le j$, except for the last term (i.e., $k = i$ and $l = j$).

The above torsional auto- and cross correlations were combined for calculating the order parameters. Let us consider, for example, the ith amide bond vector, $\mathbf{m}_i$, on the backbone, whose reorientation $\Delta\alpha_i$ is being probed by NMR. The loss in orientation imparted by $\Delta\phi_i$ and/or the associated S_i^2 is found from Eqs. (3) and (5), where ε is the angle between $\mathbf{m}_i$ and $\mathbf{l}_i$. However, there is an additional loss due to the rotational mobilities of the neighboring bonds along the backbone, and this contribution depends on the degree of correlation between neighboring bonds' torsional motions. The contribution of kth bond rotation $\Delta\phi_k$ to the reorientation of $\mathbf{m}_i$, is equally found from Eqs. (3) and (5), after replacing $\Delta\phi_i$ by $\Delta\phi_k$, and ε by the angle between $\mathbf{m}_i$ and $\mathbf{l}_k$. Inasmuch as this rotation affects S_i^2 to the extent that $\Delta\phi_i$ and $\Delta\phi_k$ are correlated, the effective contribution is found by scaling the results with the normalized cross-correlations $<\Delta\phi_i\,\Delta\phi_k>$. Cross-correlation up to second neighboring bonds ($|k - i| \le 2$, $k \ne i$) have been included in the present calculations.

REFERENCES

1. McCammon JA, Gelin BR, Karplus M. Dynamics of folded proteins. Nature 1977;267:585–590.
2. Elber R, Karplus M. Multiple conformational states of proteins: a molecular dynamics analysis of myoglobin. Science 1987;235:318–321.
3. McCammon JA, Harvey SC. Dynamics of proteins and nuclei acids. Cambridge, UK: Cambridge University Press; 1987.
4. Go M, Go N. Fluctuations of an α-helix. Biopolymers 1976;15:1119–1127.
5. Levitt M, Sander C, Stern PS. Protein normal mode dynamics. J Mol Biol 1985;181:423–447.
6. Gibrat J, Go N. Normal mode analysis of human lysozyme: study of the relative motion of the two domains and characterization of the harmonic motion. Proteins 1990;8:258–279.
7. Sternberg MJE, Grace DEP, Grace DC. Dynamic information from protein crystallography. An analysis of temperature factors from refinement of the hen egg-white lysozyme structure. J Mol Biol 1979;130:231–253.
8. Frauenfelder H, Petsko GA, Tsernoglou D. Temperature-dependent x-ray diffraction as a probe of protein structural dynamics. Nature (London)1979;280:558–563.

9. Artmiyuk PJ, Blake CCF, Grace DEP, Oatley SJ, Phillips DC, Sternberg MJE. Crystallographic studies of the dynamic properties of lysozyme. Nature (London) 1979;280:563–568.
10. Williams RJP. NMR studies of mobility within protein structure. Eur J Biochem 1989;183:479–497.
11. Wagner G. NMR relaxation and protein mobility. Curr Opin Struct Biol 1993;3:748–754.
12. Dobson CM, Karplus M. Internal motions of proteins: nuclear magnetic resonance measurements and dynamic simulations. Meth Enzymol. 1986;131:362.
13. Bahar I, Atilgan AR, Erman B. Direct evaluation of thermal fluctuations in proteins using a single parameter harmonic potential. Fold Design 1997;2:173–181.
14. Haliloglu T, Bahar I, Erman B. Gaussian dynamics of folded proteins. Phys Rev Lett 1997;79:3090–3093.
15. Bahar I, Wallquist A, Covell DG, Jernigan RL. Correlation between native-state hydrogen exchange and cooperative residue fluctuations from a simple model. Biochemistry 1998;37:1067–1075.
16. Bahar I, Atilgan AR, Demirel MC, Erman B. Vibrational dynamics of folded proteins: Significance of slow and fast modes in relation to function and stability. Phys Rev Lett 1998;80:2733–2736.
17. Bahar I, Jernigan RL. Vibrational dynamics of transfer RNAs: comparison of the free and synthetase bound forms. J Mol Biol 1998;281:871–884.
18. Bahar I, Erman B, Jernigan RL, Atilgan AR, Covell DG. Collective motions in HIV-1 reverse transcriptase: Examination of flexibility and enzyme function. J Mol Biol 1999;225:1023–1037.
19. Bahar I, Jernigan RL. Cooperative fluctuations and subunit communication in tryptophan synthase. Biochemistry 1999;38:3478–3490.
20. Demirel MC, Atilgan AR, Jernigan RL, Erman B, Bahar I. Identification of kinetically hot residues in proteins. Prot Sci 1998;7:2522–2532.
21. Buck M, Boyd J, Redfield C, MacKenzie DA, Jeenes DJ, Archer DB, Dobson CM. Structural determinants of protein dynamics: analysis of ^{15}N NMR relaxation measurements for main-chain and side-chain nuclei of hen egg white lysozyme. Biochemistry 1995;34:4041–4055.
22. Kuriyan J, Petsko GA, Levy RM, Karplus M. Effect of anisotropy and anharmonicity on protein crystallographic refinement. An evaluation by molecular dynamics. J Mol Biol 1986;190:227–254.
23. Ichiye T, Karplus M. Anisotropy and anharmonicity of atomic fluctuations in proteins: analysis of a molecular dynamics simulations. Proteins 1987;2:236–239.
24. Hayward S, Kitao A, Go N. Harmonicity and anharmonicity in protein dynamics: a normal mode analysis and principal component analysis. Proteins 1995;23:177–186.
25. Garcia AE, Krumhansl JA, Frauenfelder H. Variations on a theme by Debye and Waller: from simple crystals to proteins. Proteins 1997;29:153–160.
26. Flory PJ. Statistical thermodynamics of random networks. Proc Roy Soc Lond A 1976;351:351–380.
27. Tirion MM. Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. Phys Rev Lett 1996;77:1905–1908.
28. Bahar I, Erman B, Monerrie L. Kinematics of polymer chains with freely rotating bonds in a restrictive environment. 1. Theory. Macromolecules 1992;25:6309–6314.

29. Bahar I, Erman B, Monnerie L. Effect of molecular structure on local chain dynamics: analytical approaches and computational methods. Adv Polymer Sci 1994;116:145–206.
30. Flory PJ. Statistical mechanics of chain molecules New York: Interscience; 1969.Also reprinted by Hanser Publishers, Oxford, UK;1988. p432.
31. Pedersen TG, Sigurskjold BW, Andersen KV, et al. A nuclear magnetic resonance study of the hydrogen exchange behavior of lysozyme in crystals and solution. J Mol Biol 1991;218:413–426.
32. Campbell ID, Dobson CM, Williams RJP. Proton magnetic resonance studies of the tyrosine residues of hen lysozyme-assignment and detection of conformational mobility. Proc Roy Soc Lond B 1975;189:503–509.
33. Lu J, Dahlquist FW. Detection and characterization of an early intermediate of T4 lysozyme using pulsed hydrogen exchange and two-dimensional NMR. Biochemistry 1992;31:4749–4756.
34. Miranker A, Radford SE, Karplus M, Dobson CM. Demonstration by NMR of folding domains in lysozyme. Nature (London) 1991;349:633–636.
35. Radford SE, Dobson CM, Evans PA. The folding of hen lysozyme involves partially structured intermediates and multiple pathways. Nature (London) 1992;358:302–307.
36. Smith LJ, Sutcliffe MJ, Redfield J, Dobson, CM. Structure of hen lysozyme in solution. Biochemistry J Mol Biol 1993;229:930–944.
37. Smith LJ, Mark AE, Dobson C M, van Gunsteren WF. Comparison of MD simulations and NMR experiments for hen lysozyme. Analysis of local fluctuations, cooperative motions, and global changes. Biochemistry 1995;34:10918–10931.
38. Haliloglu T, Bahar I. Coarse-grained simulations of conformational dynamics of proteins. Application to apomyoglobin. Proteins 1998;31:271–281.
39. Baysal C, Atilgan AR, Erman B, Bahar I. Coupling between different modes in local chain dynamics: a modal correlation analysis. J Chem Soc Faraday Trans 1995;91:2483–2490.
40. Baysal C, Atilgan AR, Erman B, Bahar I. A molecular dynamics analysis of coupling between librational motions and isomeric jumps in chain molecules. Macromolecules 1996;29:2510–2514.
41. Brooks B, Karplus M. Normal modes for specific motions of macromolecules: application to the hinge-bending mode of lysozyme. Proc Natl Acad Sci USA 1985;82:4995–4999.
42. Sheriff S, Silverton EW, Padlan EA, et al. Three-dimensional structure of an antibody-antigen complex. Proc Natl Acad Sci USA 1987;84:8075–8079.
43. Kuriyan J, Weis WI. Rigid protein motion as a model for crystallographic temperature factors. Proc Natl Acad Sci USA 1991;88:2773–2777.
44. Ramanadham M, Sieker LC, Jensen LH. Refinement of triclinic lysozyme 2. The method of stereochemically restrained least-squares. Acta Crystallogr 1990;B46:63–69.
45. Vaney MC, Maignan S, Riess-Kaut M, Ducruix A. High resolution structure (1.33 Å) of a HEW lysozyme tetragonal crystal grown in APCF apparatus. Data and structural comparison with a crystal grown under microgravity from spaceHab-01 mission. Acta Crystallogr D 1996;52:505–517.