

Characterization of Collective and Anisotropic Reorientational Protein Dynamics

S. F. Lienin and R. Brüschweiler*

Carlson School of Chemistry and Biochemistry, Clark University, Worcester, Massachusetts 01610
(Received 29 November 1999)

Intramolecular reorientational dynamics of proteins are described in terms of reorientational quasi-harmonic modes. These modes provide important insight into anisotropic and collective axial fluctuations of distinct molecular fragments, and they represent a highly compact description of intramolecular protein motions that are spectroscopically observable via nuclear spin relaxation. The method is applied to a molecular dynamics computer simulation of the protein ubiquitin.

PACS numbers: 87.15.He, 76.60.-k

Protein dynamics is generally both anisotropic and collective. Internal motional anisotropy is a consequence of the general lack of symmetry in the local atomic environment while the collectivity is mainly caused by the dense packing of proteins. Understanding anisotropy and collectivity of protein motions is of considerable interest, since these properties are important for protein function. For proteins in solution, nuclear magnetic resonance (NMR) relaxation of ^{13}C and ^{15}N nuclear spins is the major source of experimental information on dynamics reflecting reorientational motions of the lattice parts of spin interactions at nearly all nuclear positions [1]. The principal axes of the lattice part of relaxation-active spin interactions [2], such as chemical shielding anisotropies and dipolar interactions to directly bonded protons, often point along different directions and thus directly probe the anisotropy of reorientation. Suitable sets of relaxation experiments, including T_1 , T_2 , nuclear Overhauser effect, and cross-correlation experiments, allow the characterization of locally anisotropic motion of rigid fragments in terms of generalized order parameters [3,4] or by realistic analytical motional models, such as the three-dimensional Gaussian axial fluctuation model (3D GAF) [5,6].

The short-range nature of nuclear spin interactions imposes natural restrictions for directly monitoring collective motional modes that involve larger numbers of fragments in a correlated way. In a previous Letter [7] it was demonstrated how NMR relaxation data can be interpreted in terms of collective protein modes derived from harmonic [8] or quasi-harmonic analysis [9] by allowing adjustments of the amplitudes and directions of the lowest frequency modes to optimize agreement with experiment.

Quasi-harmonic analysis, also referred to as essential dynamics or principal component analysis, has now become a well-established tool to analyze correlated dynamics in structural ensembles [10]. Quasi-harmonic modes reflect both reorientational and translational internal dynamics. The latter are not directly observable in ^{13}C and ^{15}N relaxation data, but they substantially contribute to the total number of modes and thus increase the complexity of the model. For example, the description of the quasi-harmonic reorientations of N backbone peptide planes requires the explicit inclusion of at least three atoms per peptide plane,

leading to a Cartesian fluctuation (covariance) matrix with a dimension of at least $9N$. On the other hand, the reorientational motions of N peptide planes are, in principle, determined by only $3N$ modes.

We present a model that describes intramolecular reorientational dynamics in terms of reorientational quasi-harmonic modes that are extracted from an ensemble of protein structures, generated by a molecular dynamics (MD) or Monte Carlo computer simulation, or which are the result of NMR structure determination. While being conceptually simpler, such a reduced model still captures the dominant motions that affect nuclear spin relaxation. Similar to quasi-harmonic analysis [9,10], the modes are identified as the eigenvectors of a fluctuation matrix calculated from the structural ensemble.

Theory.—For the calculation of reorientational quasi-harmonic modes, each structure is centered and reoriented with respect to a reference structure to eliminate overall motion. Each structure is then subdivided into N different fragments, where each fragment consists of a group of atoms that form an internally relatively rigid entity, e.g., a peptide bond or an aromatic ring. The sum of all fragments does not need to yield the complete protein; i.e., it is possible to compute reorientational quasi-harmonic modes for a subsystem of the protein, for example, for the backbone only (*vide infra*). For each protein fragment $k = 1, \dots, N$, a coordinate system is defined by three orthogonal axes represented by the vectors $\mathbf{e}_{x,k}$, $\mathbf{e}_{y,k}$, and $\mathbf{e}_{z,k}$ of unit length. The axis directions are defined in terms of the local bonding geometry which ensures that they are rigidly attached to their fragments and synchronously reorient with them.

Next, a $3N$ -dimensional covariance matrix \mathbf{M} is computed from the inner products of the vectors $\mathbf{e}_{u,k}$ and $\mathbf{e}_{v,l}$,

$$\begin{aligned} M_{uv,kl} &= \langle (\mathbf{e}_{u,k}^T - \langle \mathbf{e}_{u,k}^T \rangle) (\mathbf{e}_{v,l} - \langle \mathbf{e}_{v,l} \rangle) \rangle \\ &= \langle \mathbf{e}_{u,k}^T \cdot \mathbf{e}_{v,l} \rangle - \langle \mathbf{e}_{u,k}^T \rangle \langle \mathbf{e}_{v,l} \rangle, \end{aligned} \quad (1)$$

where $k, l = 1, \dots, N$ represent the fragments and $u, v \in \{x, y, z\}$ label the axes. The angle brackets $\langle \dots \rangle$ signify an average over the set of structures. The reorientational motions of the N fragments are spanned by $3N$ orthonormal modes \tilde{Q}_j of dimension $3N$, which are solutions to the

eigenvalue problem

$$\mathbf{M}\vec{Q}_j = \lambda_j \vec{Q}_j, \quad j = 1, \dots, 3N. \quad (2)$$

The \vec{Q}_j are the reorientational eigenmodes, and their amplitudes are reflected in the eigenvalues λ_j . Each mode \vec{Q}_j can be subdivided into N distinct 3D vectors $\vec{q}_{j,k} = (Q_{j,3k-2}, Q_{j,3k-1}, Q_{j,3k})$, $k = 1, \dots, N$, where $Q_{j,m}$ is the component m of vector Q_j . Vector $\vec{q}_{j,k}$ is attached to fragment k and reoriented under the j th mode with a variance $\sigma_{j,k}^2 = \langle \vec{q}_{j,k}^T \cdot \vec{q}_{j,k} \rangle - \langle \vec{q}_{j,k}^T \rangle \langle \vec{q}_{j,k} \rangle = \lambda_j |\vec{q}_{j,k}|^2$. The normalization condition $|\vec{Q}_j|^2 = 1$ implies

$$\sum_{k=1}^N |\vec{q}_{j,k}|^2 = 1. \quad (3)$$

The net effect of all $3N$ reorientational modes on a single fragment k can be assessed from the 3×3 matrix $\mathbf{M}^{(k)}$ with elements

$$M_{uv}^{(k)} = \sum_{j=1}^{3N} \lambda_j (\vec{q}_{j,k})_u (\vec{q}_{j,k})_v, \quad u, v \in \{x, y, z\}. \quad (4)$$

The eigenvectors of $\mathbf{M}^{(k)}$ are the principal axes of reorientation \mathbf{e}_ξ with eigenvalues λ'_ξ ,

$$\mathbf{M}^{(k)} \mathbf{e}_\xi = \lambda'_\xi \mathbf{e}_\xi, \quad \xi \in \{\alpha, \beta, \gamma\}. \quad (5)$$

In the case of harmonic reorientations, the variances σ_α^2 , σ_β^2 , and σ_γ^2 of fluctuations about the axes \mathbf{e}_α , \mathbf{e}_β , and \mathbf{e}_γ are given by

$$\sigma_\alpha^2 = \frac{1}{2} \log \left(\frac{1 - \lambda'_\alpha}{(1 - \lambda'_\beta)(1 - \lambda'_\gamma)} \right) \quad (6)$$

and permutations in the indices α , β , and γ . This corresponds to the 3D GAF model [5] that was previously used to determine anisotropic peptide-plane dynamics in ubiquitin from NMR relaxation data [6]. It provides a local description of anisotropic reorientational motion ignoring correlation effects between the fragments. Equation (1) can thus be viewed as a generalization of the 3D GAF model to $3N$ dimensions, which we call the *collective axial fluctuation* model. Spin relaxation parameters can be readily calculated from $\mathbf{M}^{(k)}$ via σ_ξ^2 and \mathbf{e}_ξ [6], and thus eigenvalues λ_j and eigenvectors \vec{Q}_j can be adjusted to fit experimental data by procedures analogous to the ones described in Ref. [7].

Collectivity measure.—The reorientational modes \vec{Q}_j exhibit a variable degree of delocalization or collectivity with respect to the various protein fragments. A quantitative measure for the collectivity of each mode is the *mode collectivity* κ [11] which is related to information entropy:

$$\kappa_j = \frac{1}{N} \exp \left\{ - \sum_{k=1}^N |\vec{q}_{j,k}|^2 \log |\vec{q}_{j,k}|^2 \right\}, \quad (7)$$

where κ_j is the number between $1/N$ and 1 given by the ratio between the effective number of fragments that are collectively reoriented by the j th mode and the total

number of fragments. A small κ reflects local motion, while a large κ reflects a substantial degree of collectivity.

Correlated motions.—The reorientational motion of two fragments k and l is correlated if the same modes provide, on average, similar contributions. The amount of correlation can be expressed by the correlation coefficient,

$$r_{kl} = \text{cov}(s_k, s_l) / (\sigma_k \sigma_l) \quad (8)$$

between the sets s_k, s_l ,

$$\begin{aligned} s_k &= \{\lambda_1 |\vec{q}_{1,k}|^2, \dots, \lambda_{3N} |\vec{q}_{3N,k}|^2\}, \\ s_l &= \{\lambda_1 |\vec{q}_{1,l}|^2, \dots, \lambda_{3N} |\vec{q}_{3N,l}|^2\}, \end{aligned} \quad (9)$$

where the elements correspond to the variances described after Eq. (2). σ_k and σ_l are the standard deviations of the two sets s_k and s_l , and $\text{cov}(s_k, s_l) = \langle s_k s_l \rangle - \langle s_k \rangle \langle s_l \rangle$ is their covariance.

Application.—The method was applied to 1500 snapshots of a 1.5 ns MD trajectory (stored with a time increment of 1 ps) of the 76 amino-acid protein ubiquitin to study its backbone peptide-plane dynamics. The protein was embedded in a cubic box including 2909 explicit water molecules with the temperature set to 300 K. The simulation was carried out using the program CHARMM 24 [12] under periodic boundary conditions. More details about this simulation are given in Ref. [6]. For each snapshot, an orthogonal coordinate system was defined for each of the $N = 72$ nonproline peptide planes with axes $\mathbf{e}_{u,k}$, where $u = x, y, z$ and $k = 1, \dots, 72$. The matrix \mathbf{M} was then calculated according to Eq. (1) and diagonalized yielding $3 \times 72 = 216$ reorientational modes \vec{Q}_j with eigenvalues λ_j varying between 0.0003 and 1.94.

Figure 1 shows the local amplitudes $|\vec{q}_{j,k}|^2$ as a function of the mode number j and the peptide-plane number k , where the modes are sorted with respect to increasing eigenvalues. Different modes can exhibit a qualitatively different behavior as can be seen in the stack plot of Fig. 1(a). Four selected modes are displayed in more detail in Fig. 1(b). Some of the modes with the largest eigenvalues (e.g., mode 209) exhibit correlated large amplitude motion in backbone positions that are known from previous studies [6] to be considerably flexible, including several loop regions that connect regular secondary structures (β -strands and α -helix) indicated at the top of the figure. Some of the modes with small amplitudes, such as modes 1 and 4, exhibit a notably concerted behavior for extended uninterrupted parts of the backbone. Other modes involve many peptide planes scattered throughout the primary sequence, such as mode 172 which has the largest collectivity of all of the modes with $\kappa_{172} = 0.69$ involving about 50 of the 72 peptide planes. In Fig. 2 the collectivities κ_j of modes \vec{Q}_j are displayed as a function of the eigenvalues λ_j . Small amplitude reorientational modes have typically a smaller collectivity [exceptions include modes 1 and 4 of Fig. 1(b)]. The collectivity of larger amplitude modes grows on average and reaches a maximum around $\lambda_j = 0.04$ after which the collectivities decrease

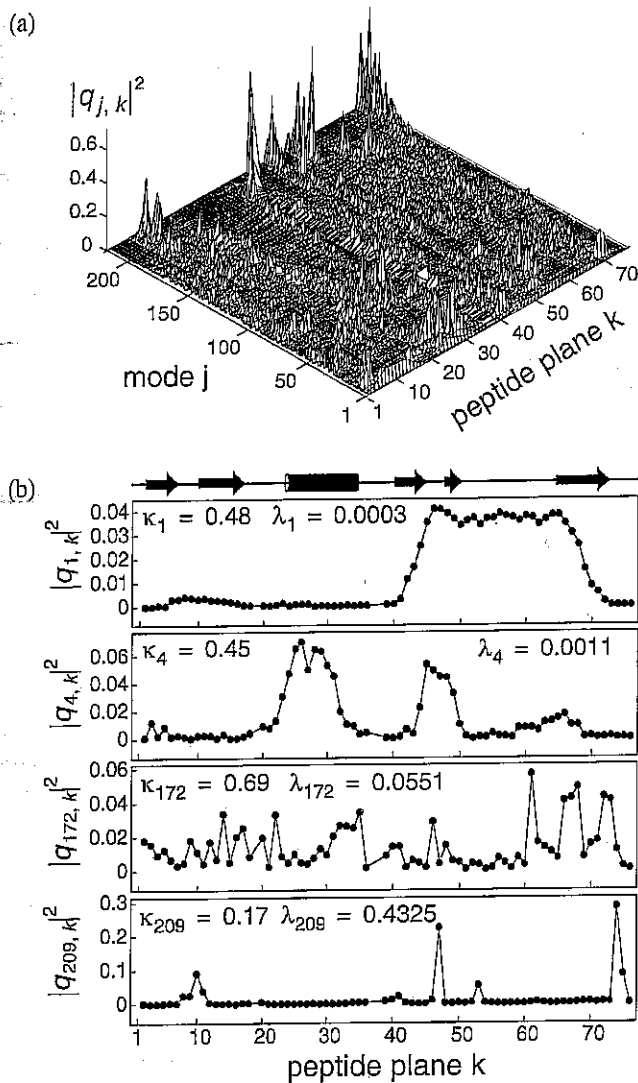


FIG. 1. (a) Reorientational mode amplitudes $|\bar{q}_{j,k}|^2$ of ubiquitin backbone peptide planes as a function of mode number $j = 1, \dots, 216$ and peptide-plane number k calculated from 1500 reoriented snapshots of a 1.5 ns MD trajectory of ubiquitin sorted with respect to increasing eigenvalues λ_j . (b) Selected cross sections through (a) for modes $j = 1, 4, 172$, and 209. At the top, the secondary structural elements of ubiquitin are indicated by arrows for β -strands and a cylinder for the α -helix.

again: The modes with largest λ_j possess the lowest collectiveities ($\kappa_j < 0.1$). These modes reflect backbone dihedral angle flips at various positions on slower time scales in the hundreds of ps. The distribution of κ is maximal at about $\kappa = 0.5$, while the λ distribution has a maximum close to zero (at $\lambda < 0.01$).

Reorientational correlation effects between pairs of peptide planes are displayed in Fig. 3 using the correlation coefficient r_{kl} of Eq. (8). A black square indicates that $|r_{kl}| > 0.5$. Since the mode eigenvalues λ_j explicitly enter the correlation coefficient [see Eq. (8)], correlations between larger amplitude modes are emphasized. As a consequence, mainly peptide groups that are strongly af-

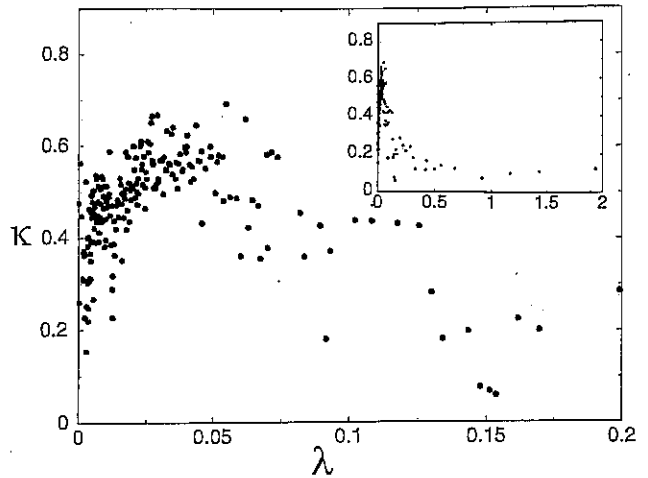


FIG. 2. Mode collectivity κ plotted as a function of the mode eigenvalue λ for the 216 backbone reorientational modes of ubiquitin calculated using Eq. (7). The inset shows the full range of eigenvalues.

ected by at least one large-amplitude mode of Fig. 1(a) are motionally significantly correlated to other peptide planes. There are exceptions, however, such as peptide planes 20 and 39–42, whose motion is rather restricted, but who show correlations to many of the more mobile protein regions. Interestingly, peptide plane 36, which is positioned between the α -helix and two adjacent proline residues at positions 37 and 38, shows large-amplitude reorientational dynamics that lacks significant correlations to other amino acids. Motional correlation effects inside the quite rigid α -helix and between β -strands do exist. These can be

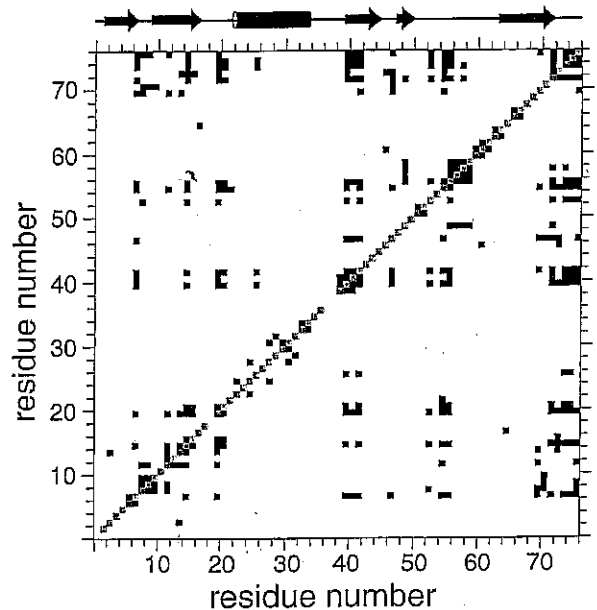


FIG. 3. Reorientational correlation matrix $|r_{kl}|$ between peptide planes of ubiquitin calculated using Eq. (8). A black square corresponds to a correlation coefficient $|r_{kl}| > 0.5$. The secondary structural elements are indicated at the top.

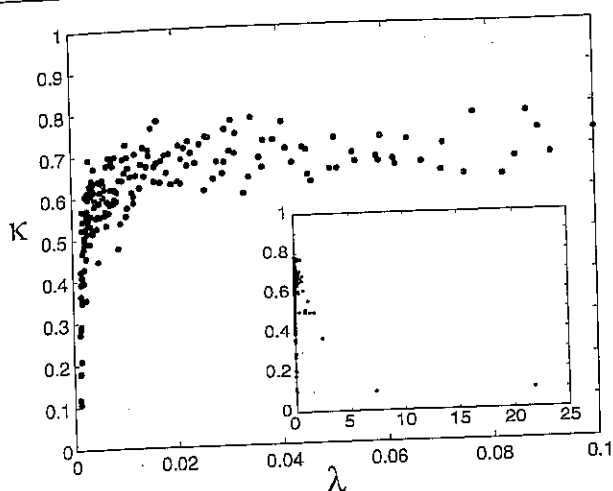


FIG. 4. Mode collectivity κ [see Eq. (7)] plotted as a function of the mode eigenvalue λ (in units of \AA^2) for the 216 quasiharmonic modes of the backbone nitrogen atoms of ubiquitin calculated from the 1.5 ns MD trajectory. Modes and eigenvalues were obtained by diagonalization of the matrix $M_{uv,kl} = \langle (r_{k,u} - \langle r_{k,u} \rangle)(r_{l,v} - \langle r_{l,v} \rangle) \rangle$, $u, v \in \{x, y, z\}$ where $r_{k,u}$ is the component u of the 3D position vector r_k of backbone nitrogen atom k . The inset shows the full range of eigenvalues.

better analyzed using correlation coefficients given by Eq. (8) but with sets s_k and s_l , where the λ_j in Eq. (9) are set to 1 (data not shown).

A hallmark of quasiharmonic analysis is the fact that a relatively small number of high-amplitude modes tend to dominate the behavior of the whole protein [9,10]. Here, 90% of reorientational motion is, for ten peptide planes, caused by the largest 8–30 modes, for 19 of the remaining peptide planes by the largest 56–111 modes, and for the remaining 43 peptide planes by 116–178 modes. Therefore interpretation of relaxation parameters in a reorientational subspace with a dimension that is substantially lower than $3N = 216$ is feasible for about 40% of the peptide planes that include all planes with high mobility.

We compared the reorientational mode collectivities (Fig. 2) to mode collectivities determined from a standard quasiharmonic analysis [9] applied to the nitrogen atoms of the 72 peptide planes for the same 1500 snapshots of ubiquitin. The collectivities are plotted as a function of the eigenvalues in Fig. 4. For small amplitude modes the collectivities show a similar behavior as for the reorientational modes, whereas for increasing amplitudes the quasiharmonic collectivities stay systematically higher than the reorientational ones (except for the three modes with largest amplitudes). Thus, projection of quasiharmonic modes on the reorientational subspace

leads to a decrease in collectivity, in particular for the large-amplitude modes, indicating that reorientational motions are systematically more local in character than translational motions. This is consistent with the picture that translational motions require the concerted displacement of large numbers of atoms due to the dense protein packing, while reorientational motions can be accommodated more locally. Furthermore, the differences in collectivity suggest that translational and reorientational low-frequency modes are to a significant extent decoupled from each other. The collective axial fluctuation model therefore represents a highly compact and self-contained description of intramolecular protein dynamics amenable to nuclear spin relaxation.

This work was supported by NSF Grant No. MCB-9904875. We thank Christoph Scheurer for computer support.

*To whom correspondence should be addressed.

Email address: bruschweiler@nmr.clarku.edu

- [1] L. E. Kay, *Nat. Struct. Biol.* **5**, 513 (1998).
- [2] A. Abragam, *Principles of Nuclear Magnetism* (Clarendon, Oxford, 1961).
- [3] G. Lipari and A. Szabo, *J. Am. Chem. Soc.* **104**, 4546 (1982).
- [4] M. W. F. Fischer, L. Zeng, Y. Pang, W. Hu, A. Majumdar, and E. R. P. Zuiderweg, *J. Am. Chem. Soc.* **119**, 12 629 (1997).
- [5] T. Bremi and R. Brüschweiler, *J. Am. Chem. Soc.* **119**, 6672 (1997).
- [6] S. F. Lienin, T. Bremi, B. Brutscher, R. Brüschweiler, and R. R. Ernst, *J. Am. Chem. Soc.* **120**, 9870 (1998).
- [7] R. Brüschweiler and D. A. Case, *Phys. Rev. Lett.* **72**, 940 (1994).
- [8] T. Noguti and N. Gō, *J. Phys. Soc. Jpn.* **52**, 3283 (1983); B. Brooks and M. Karplus, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 6571 (1983); M. Levitt, C. Sander, and P. S. Stern, *J. Mol. Biol.* **181**, 423 (1985).
- [9] R. M. Levy, M. Karplus, J. Kushick, and D. Perahia, *Macromolecules* **17**, 1370 (1984).
- [10] A. Amadei, A. B. Linssen, B. L. de Groot, D. M. van Alten, and H. J. Berendsen, *J. Biomol. Struct. Dyn.* **13**, 615 (1996); T. Haliloglu, I. Bahar, and B. Erman, *Phys. Rev. Lett.* **79**, 3090 (1997); K. Hinsen, *Proteins* **33**, 417 (1998); A. Kitao, S. Hayward, and N. Gō, *Proteins* **33**, 496 (1998); A. Abseher and M. Nilges, *J. Mol. Biol.* **279**, 911 (1998).
- [11] R. Brüschweiler, *J. Chem. Phys.* **102**, 3396 (1995).
- [12] R. B. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, *J. Comput. Chem.* **4**, 187 (1993); A. D. MacKerell, Jr. *et al.*, *J. Phys. Chem. B* **102**, 3586 (1998).