

Anisotropy of Fluctuation Dynamics of Proteins with an Elastic Network Model

A. R. Atilgan,* S. R. Durell,[†] R. L. Jernigan,[†] M. C. Demirel,* O. Keskin,[†] and I. Bahar*[†]

*School of Engineering and Polymer Research Center, Bogazici University, and TUBITAK Advanced Polymeric Materials Research Center, Bebek 80815, Istanbul, Turkey; and [†]Molecular Structure Section, Laboratory of Experimental and Computational Biology, Division of Basic Sciences, National Cancer Institute and Science Applications International Corp., National Institutes of Health, Bethesda and Frederick, Maryland 20892-5677 USA

ABSTRACT Fluctuations about the native conformation of proteins have proven to be suitably reproduced with a simple elastic network model, which has shown excellent agreement with a number of different properties for a wide variety of proteins. This scalar model simply investigates the magnitudes of motion of individual residues in the structure. To use the elastic model approach further for developing the details of protein mechanisms, it becomes essential to expand this model to include the added details of the directions of individual residue fluctuations. In this paper a new tool is presented for this purpose and applied to the retinol-binding protein, which indicates enhanced flexibility in the region of entry to the ligand binding site and for the portion of the protein binding to its carrier protein.

INTRODUCTION

It is well known that proteins in the folded state are not rigid, but instead can sample numerous conformations in the neighborhood of their native conformation (Frauenfelder et al., 1991). These can generally be viewed as fluctuations near equilibrium positions, called here *equilibrium fluctuations*. These fluctuations occur in addition to two other more specific classes of conformational transitions: 1) local conformational isomeric jumps occasionally occurring near native state coordinates, particularly for side chains having rotatable bonds; and 2) large-scale changes: some proteins can also have two or more equilibrium states relevant to their function (Damaschun et al., 1999; Frauenfelder and McMahon, 1998). Nevertheless, apart from these two categories, the most characteristic general fluctuations are usually small in magnitude, not exceeding several Ångströms, and lie in the subnanosecond frequency range.

Details of molecular motions in the folded state can, in principle, be elucidated by molecular dynamics (MD) simulations and normal mode analyses (NMA) using all-atom empirical potentials developed for proteins (McCammon and Harvey, 1987; Brooks et al., 1988; Kitao and Go, 1999). However, the use of atomic approaches becomes computationally inefficient with increasing size of the system and can even obscure the larger scale changes. However, recent studies reveal the success of coarse-grained protein models and simplified force fields for describing the vibrational dynamics of simple models; these are particularly appropriate for describing the collective motions of complex systems or of the largest proteins (Bahar et al., 1999; Bahar and Jernigan, 1999; Hinsen et al., 1999), composed of more than

several thousand residues. Such systems cannot usually be investigated using conventional atomic models and potentials. Simple models and efficient computational methods (Amadei et al., 1993; van Aalten et al., 1997a) become increasingly important with the pressing need to systematically elucidate the dynamics of large numbers of proteins and their complexes within the scope of structural and functional genomics.

In a series of recent papers (Bahar et al., 1997, 1998a, 1999; Haliloglu et al., 1997; Bahar and Jernigan, 1998), we have shown that the fluctuation dynamics of proteins can be modeled as those of elastic networks; the nodes are the residues, and the linkers are the inter-residue potentials stabilizing the folded conformation (Fig. 1). In this model, shortly referred to as the Gaussian network model (GNM) of proteins, residues are assumed to undergo Gaussian-distributed fluctuations about their mean positions, being coupled by harmonic potentials. No distinction is made between different types of amino acids, so that a generic force constant γ is adopted for the interaction potential between all pairs of residues sufficiently close. Experimental values of such generic force constants may become available soon from new neutron scattering experiments (Zaccai, 2000), which would permit a more direct way to evaluate these constants.

Results from GNM calculations were found to be in excellent agreement with x-ray crystallographic temperature factors (also called Debye-Waller or B-factors) (Bahar et al., 1997, 1998a, 1999; Haliloglu et al., 1997; Bahar and Jernigan, 1998, 1999; Demirel et al., 1998; Jernigan et al., 1999; Keskin et al., 2000), the H/D exchange free energies of amide protons with solvent for various proteins near native state conditions (Bahar et al., 1998b), and the order parameters from NMR-relaxation measurements (Haliloglu and Bahar, 1999). The success of the GNM has been attributed to the following features: 1) on a coarse-grained scale molecular motions can be approximated by normal fluctu-

Received for publication 30 June 2000 and in final form 23 October 2000.

Address reprint requests to Dr. Robert Jernigan, NCI, NIH, MSS, LECB, DBS, LECB, Rm. B-116, Bldg. 12B, Bethesda, MD 20892-5677. Tel.: 301-496-4783; Fax: 301-402-4724; E-mail: robert.jernigan@nih.gov.

© 2001 by the Biophysical Society

0006-3495/01/01/505/11 \$2.00

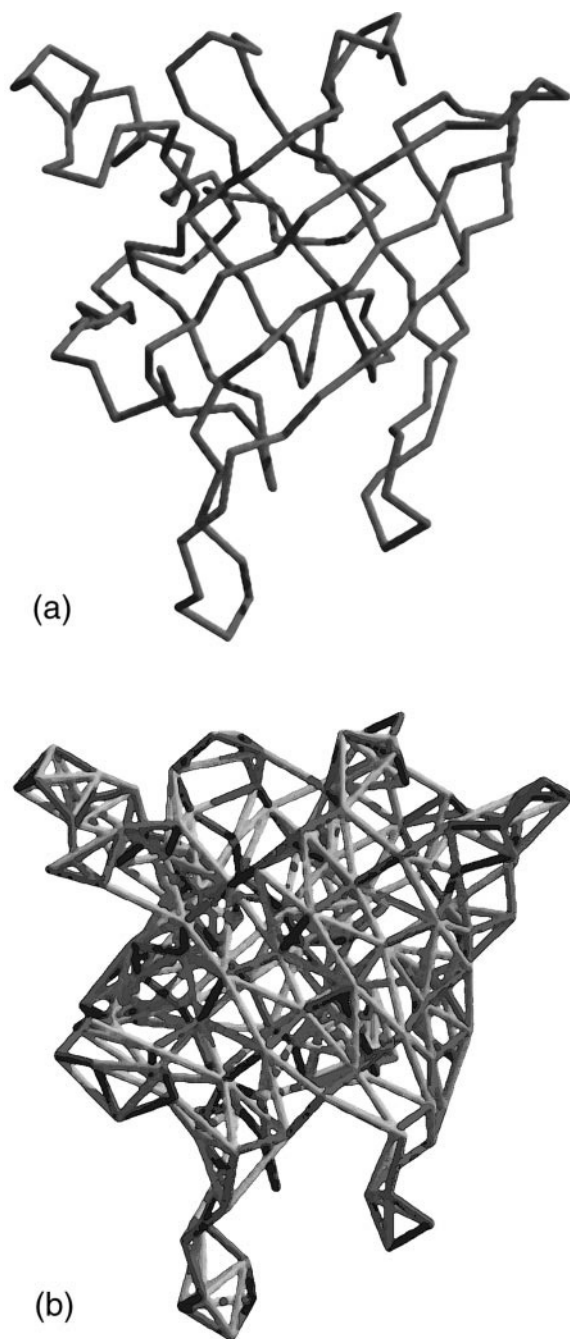


FIGURE 1 Representation of protein structure as an elastic network in the GNM. The backbone trace is shown in dark lines in (A) and (B). The gray rods in (b) connect the pairs of α -carbons located within a distance $s_{ij} \leq r_c$. The figure displays the model for retinol-binding protein (RBP) crystal structure (Zanotti et al., 1998). A network structure with the spring connections is shown for a relatively short cutoff distance ($r_c = 6 \text{ \AA}$).

ations with rescaled force constants, in conformity with the central limit theorem applied to superimposed probability density functions, even if the motions of individual atoms depart from harmonicity; and 2) GNM yields an analytical solution, devoid of sampling inaccuracies found in most other types of simulations.

Although the above-listed experimental data and theoretical results all relate only to the magnitude of the fluctuations, the directions are also important. GNM has been amply verified to provide information on the magnitudes of fluctuations. All fluctuations are implicitly assumed to be isotropic in the GNM; no directional preferences or 3-D character of motion is accounted for. The molecule is viewed as a collection of N sites, one for each residue, resulting in an ensemble of $N - 1$ independent modes, instead of the $3N - 6$ modes that would be obtained in a 3-D description. However, in reality the fluctuations are in general anisotropic (Kuriyan et al., 1986; Ichiye and Karplus, 1987), and it is important to assess the directions of collective motions, as these can be directly relevant to biological function and mechanisms. It is not indeed possible to acquire an understanding of the mechanism of motion unless the fluctuation vectors, in addition to their magnitudes, are elucidated. An extension of the GNM, called the anisotropic network model (ANM), is presented herein to address this issue. A recent comparison of the results from MD simulations with the GNM and its extension for treating the anisotropy of fluctuations lends support to the development of the ANM (Doruker et al., 2000).

For self-consistency and development in context, a recapitulation of the GNM is presented first. This will be followed by two formulations of the ANM, based on different approaches, which will be shown to be mathematically equivalent. An application to a β -barrel protein, retinol binding protein (RBP), will be illustrated in Results and Discussion. A discussion of the utility and limitations of the ANM will be presented in the Conclusion.

THEORY

Kirchhoff matrix of contacts

The GNM description of proteins bears a close resemblance to the Rouse chain model (Rouse, 1953; Doi and Edwards, 1986) of polymers in which the molecule is represented as a collection of beads and springs. But, whereas in the Rouse chain only sequentially adjacent beads are connected by springs, in the GNM sequentially distant pairs in close contact in addition to neighbors along the sequence are coupled. The classical Rouse matrix

$$\Gamma = \begin{bmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & \dots & \dots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 1 \end{bmatrix} \quad (1)$$

is therefore replaced by a Kirchhoff matrix of contacts,

$$\Gamma_{ij} = \begin{cases} -1 & s_{ij} \leq r_c \\ 0 & s_{ij} > r_c \end{cases} \quad \Gamma_{ii} = - \sum_{k,k \neq i}^N \Gamma_{ik} \quad (2)$$

where the non-zero off-diagonal elements refer to residue pairs i and j that are connected via springs—their separation s_{ij} being shorter than a cutoff value r_c for inter-residue interactions—and the diagonal elements are found from the negative sum of the off-diagonal terms in the same row (or column). The diagonal terms thus represent the coordination number, i.e., the number of individual residues found within a coordination sphere of radius r_c . See Fig. 2 for a schematic representation of the equilibrium (\mathbf{R}_i^0) and instantaneous (\mathbf{R}_i) position vectors of the residues, and the corresponding separation vectors \mathbf{s}_{ij} and \mathbf{s}_{ij}^0 .

Potentials and forces in the GNM

The Kirchhoff matrix is conveniently used in Flory (1976)

$$V = (\gamma/2) \Delta \mathbf{R}^T \Gamma \Delta \mathbf{R} \quad (3)$$

for evaluating the overall conformational potential of the structure. Here $\Delta \mathbf{R}$ is the N -dimensional vector whose elements are the fluctuation vectors $\Delta \mathbf{R}_i$ of the individual residues ($1 \leq i \leq N$). Note that the above matrix multiplication yields a summation over all fluctuations squared $(\Delta s_{ij})^2 + (s_{ij} - s_{ij}^0)^2 = (\Delta \mathbf{R}_i - \Delta \mathbf{R}_j)^2$ for all pairs $s_{ij} \leq r_c$, multiplied by $(\gamma/2)$. This summation represents the overall potential of the system. It can also be viewed as the Hamiltonian whose integration yields the configurational partition function Z_N .

Let M represent the total number of springs composing the network. Clearly, M is found from the summation of the coordination number of individual residues divided by 2, i.e.,

$$M = \frac{1}{2} \sum_i \Gamma_{ii} \quad (4)$$

where $1 \leq i \leq N$. In the GNM, the M -dimensional vector $[\Delta \mathbf{s}]_{M \times 1}$ of the instantaneous fluctuations of the individual springs is associated with M forces, which may be arranged in an M -dimensional array as

$$[\mathbf{f}]_{M \times 1} = \gamma [\mathbf{I}]_{M \times M} [\Delta \mathbf{s}]_{M \times 1} \quad (5)$$

Here $[\mathbf{I}]_{M \times M}$ is the identity matrix of order M . The subscripts denote the size of the matrices (or vectors) enclosed in brackets. Note that the forces and fluctuations in Eq. 5 are along the directions of the corresponding springs, hence their representation as arrays of dimension M . At equilib-

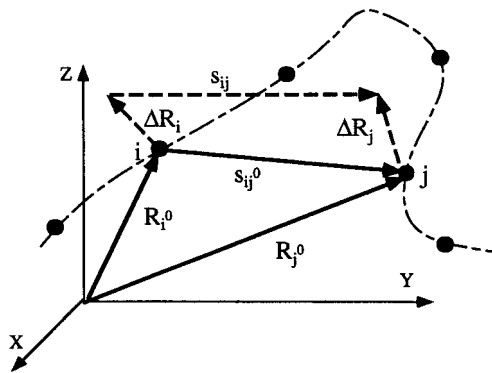


FIGURE 2 Schematic representation of the fluctuations $\Delta \mathbf{R}_i$ and $\Delta \mathbf{R}_j$ in the position vectors of residue sites i and j . The equilibrium position vectors with respect to the frame XYZ are denoted as \mathbf{R}_i^0 and \mathbf{R}_j^0 , and their instantaneous values are \mathbf{R}_i and \mathbf{R}_j . \mathbf{s}_{ij}^0 and \mathbf{s}_{ij} are the equilibrium and instantaneous separation vectors between sites i and j . The change in the separation with respect to the equilibrium coordinates is $\mathbf{s}_{ij} - \mathbf{s}_{ij}^0 = \Delta \mathbf{R}_j - \Delta \mathbf{R}_i$.

rium, the force on each residue must be zero, so that

$$\gamma \Gamma \Delta \mathbf{R} = 0 \quad (6)$$

Here again, we recall that $\Delta \mathbf{R}$ is the N -dimensional array of the individual fluctuation vectors in the positions of the N residues. Clearly, the constant γ can be eliminated from Eq. 6, to obtain $\Gamma \Delta \mathbf{R} = \mathbf{0}$.

Fluctuations

The cross-correlations between residue fluctuations are found from

$$\begin{aligned} \langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j \rangle &= (1/Z_N) \int (\Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j) \exp\{-V/k_B T\} d\{\Delta \mathbf{R}\} \\ &= (3k_B T / \gamma) [\Gamma^{-1}]_{ij} \end{aligned} \quad (7)$$

where $[\Gamma^{-1}]_{ij}$ represents the ij th element of the inverse of Γ . The mean-square (ms) fluctuations of individual residues can be readily found from Eq. 7, taking $i = j$, i.e.,

$$\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_i \rangle = \langle (\Delta R_i)^2 \rangle = (3k_B T / \gamma) [\Gamma^{-1}]_{ii} \quad (8)$$

$\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j \rangle$ can be expressed as a sum over the contributions $[\Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j]_k$ of the individual modes, in an expansion using the eigenvalues λ_k and eigenvectors \mathbf{u}_k of Γ in

$$\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j \rangle = \sum_k [\Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j]_k = (3k_B T / \gamma) \sum_k [\lambda_k^{-1} \mathbf{u}_k \mathbf{u}_k^T]_{ij} \quad (9)$$

Here, the summation is performed over all $(1 \leq k \leq N - 1)$ non-zero eigenvalues of Γ . It is clear that Eqs. 7–9 permit us to calculate the amplitudes of fluctuations for individual residues without providing information regarding their absolute orientations or directions. Directional preferences, which are important for molecular behavior and can affect the molecular mechanisms of biological processes, are elucidated by the new ANM introduced next.

Anisotropic motions from force balance

Let us consider a central residue i , subject to interactions with m neighbors located within r_c . Under native state conditions, the sum of forces on residue i along the X -, Y -, and Z -directions must each be equal to zero, i.e.,

$$\begin{aligned} \sum_j f_{ij} \cos \alpha_{ij}^X &= \sum_j f_{ij} (X_j - X_i) / s_{ij} = 0 \\ \sum_j f_{ij} \cos \alpha_{ij}^Y &= \sum_j f_{ij} (Y_j - Y_i) / s_{ij} = 0 \\ \sum_j f_{ij} \cos \alpha_{ij}^Z &= \sum_j f_{ij} (Z_j - Z_i) / s_{ij} = 0 \end{aligned} \quad (10)$$

where the summations are performed over all near neighbors of residue i (Γ_{ii} of them), f_{ij} is the force on site i due to its interaction with site j , α_{ij}^X is the angle between the X axis and the line of action of f_{ij} , which also coincides with the direction of the spring between sites i and j . X_i , Y_i , and Z_i are the components of \mathbf{R}_i . These three equations can be cast in a matrix form as $\mathbf{B} \cdot \mathbf{f}_i = \mathbf{0}$, where \mathbf{B} is the $3 \times \Gamma_{ii}$ matrix of cosines, and \mathbf{f}_i is the Γ_{ii} -dimensional vector of the magnitudes of the forces exerted on site i . This force balance can be generalized to the complete set of N sites and M interactions as

$$[\mathbf{B}]_{3N \times M} [\mathbf{f}]_{M \times 1} = [\mathbf{0}]_{3N \times 1} \quad (11)$$

We note that M , the total number of springs in the network, is given by $M = zN/2$ on the average, where z is the mean coordination number of residues. For $r_c = 7.0$ Å, $z \leq 7.5$, such that the $3N$ equalities given in Eq. 11 are fewer than the number of unknowns (M). This set of equations is thus underdetermined.

Determination of the unknowns (forces) requires invoking two additional properties (Demirel et al., 1999): 1) the constitutive equation, relating the forces to deformations $\Delta \mathbf{s}$, and 2) the kinematics or geometry of deformation relating the fluctuations of the springs to those of the nodes. In the linear approximation, the former is given by

$$[\mathbf{f}]_{M \times 1} = [\mathbf{K}]_{M \times M} [\Delta \mathbf{s}]_{M \times 1} \quad (12)$$

where $[\mathbf{K}]_{M \times M}$ is a diagonal matrix whose i th element represents the force constant referring to the i th residue pair. In the GNM, $[\mathbf{K}]_{M \times M}$ is simply given by $[\mathbf{K}]_{M \times M} = \gamma [\mathbf{I}]_{M \times M}$ (see Eq. 7).

The kinematics relationships between the individual deformations and the positions of the residues are, however, given by

$$[\mathbf{B}]_{M \times 3N}^T [\Delta \mathbf{R}]_{3N \times 1} = [\Delta \mathbf{s}]_{M \times 1} \quad (13)$$

where $[\mathbf{B}]^T$ is the transpose of \mathbf{B} . Substituting Eqs. 12 and 13 into Eq. 11, we obtain

$$[\mathbf{B}]_{3N \times M} [\mathbf{K}]_{M \times M} [\mathbf{B}]_{M \times 3N}^T [\Delta \mathbf{R}]_{3N \times 1} = [\mathbf{0}]_{3N \times 1} \quad (14)$$

For the simple case $[\mathbf{K}]_{M \times M} = \gamma [\mathbf{I}]_{M \times M}$, Eq. 14 reduces to

$$\gamma [\mathbf{B}]_{3N \times M} [\mathbf{B}]_{M \times 3N}^T [\Delta \mathbf{R}]_{3N \times 1} = [\mathbf{0}]_{3N \times 1} \quad (15)$$

which upon comparison with Eq. 6 stipulates the equivalence of the N -dimensional Kirchhoff matrix Γ of GNM to the $3N$ -dimensional matrix $[\mathbf{B}][\mathbf{B}]^T$ of the ANM.

Anisotropic fluctuations from the second derivative of harmonic potentials

The $3N$ -dimensional counterpart $[\mathbf{B}][\mathbf{B}]^T$ of Γ can be alternatively derived from the second derivative of the overall potential (Eq. 3) using conventional methods of NMA (Go et al., 1983; Brooks and Karplus, 1983). To this aim, let us first consider a single spring between residues i and j , subject to the harmonic potential

$$\begin{aligned} V &= (1/2)\gamma(s_{ij} - s_{ij}^0)^2 \\ &= (1/2)\gamma[(X_j - X_i)^2 + (Y_j - Y_i)^2 \\ &\quad + (Z_j - Z_i)^2]^{1/2} - s_{ij}^0)^2 \end{aligned} \quad (16)$$

The first and second derivatives of V with respect to the components of \mathbf{R}_i are

$$\begin{aligned} \partial V / \partial X_i &= -\partial V / \partial X_j \\ &= -\gamma(X_j - X_i)(1 - s_{ij}^0/s_{ij}) \end{aligned} \quad (17)$$

$$\begin{aligned} \partial^2 V / \partial X_i^2 &= \partial^2 V / \partial X_j^2 \\ &= \gamma(1 + s_{ij}^0(X_j - X_i)^2/s_{ij}^3 - s_{ij}^0/s_{ij}) \end{aligned} \quad (18)$$

Similar expressions hold for the Y - and Z -components of \mathbf{R}_i , such that in compact notation $\nabla_i V = -\gamma[\mathbf{s}_{ij}(1 - s_{ij}^0/s_{ij})]$ and $\nabla_i^2 V = 3\gamma$. At equilibrium,

$s_{ij} = s_{ij}^0$, and Eqs. 17 and 18 reduce to

$$\partial V / \partial X_i = 0 \quad (19)$$

$$\partial^2 V / \partial X_i^2 = \gamma(X_j - X_i)^2/s_{ij}^2 \quad (20)$$

Likewise, the second cross-derivatives become

$$\begin{aligned} \partial^2 V / \partial X_i \partial Y_j &= -\partial^2 V / \partial X_j \partial Y_i \\ &= -\gamma(X_j - X_i)(Y_j - Y_i)/s_{ij}^2 \end{aligned} \quad (21)$$

In the case of Γ_{ii} neighbors surrounding residue i , Eqs. 20 and 21 are replaced by

$$\partial^2 V / \partial X_i^2 = \gamma \sum_j (X_j - X_i)^2/s_{ij}^2 \quad (22)$$

$$\partial^2 V / \partial X_i \partial Y_i = \gamma \sum_j (X_j - X_i)(Y_j - Y_i)/s_{ij}^2 \quad (23)$$

where the summations are carried out over all neighbors ($j = 1, \Gamma_{ii}$) of residue i .

In the general case of N residues connected by M linkers, the second derivatives of the overall potential are organized in the $3N \times 3N$ Hessian matrix \mathcal{H} . \mathcal{H} is composed of $N \times N$ super-elements of size 3×3 , i.e.,

$$\mathcal{H} = \begin{bmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} & \cdots & \mathbf{H}_{1N} \\ \mathbf{H}_{21} & & & \mathbf{H}_{2N} \\ \vdots & & & \vdots \\ \mathbf{H}_{N1} & & & \mathbf{H}_{NN} \end{bmatrix} \quad (24)$$

The ij th super-element ($i \neq j$) \mathbf{H}_{ij} of \mathcal{H} is

$$\mathbf{H}_{ij} = \begin{bmatrix} \partial^2 V / \partial X_i \partial X_j & \partial^2 V / \partial X_i \partial Y_j & \partial^2 V / \partial X_i \partial Z_j \\ \partial^2 V / \partial Y_i \partial X_j & \partial^2 V / \partial Y_i \partial Y_j & \partial^2 V / \partial Y_i \partial Z_j \\ \partial^2 V / \partial Z_i \partial X_j & \partial^2 V / \partial Z_i \partial Y_j & \partial^2 V / \partial Z_i \partial Z_j \end{bmatrix} \quad (25)$$

Equation 21 gives the elements of \mathbf{H}_{ij} for the ANM. The elements of the diagonal super-elements \mathbf{H}_{ii} , however, are given by Eqs. 22 (diagonal) and 23 (off-diagonal). The elements of \mathcal{H} are conveniently calculated for the ANM using the α -carbon position vectors of databank structures.

We note that \mathcal{H} and $\gamma[\mathbf{B}][\mathbf{B}]^T$ are equal to each other. As a simple verification consider, for example, the first diagonal element \mathbf{H}_{11} of \mathcal{H} . \mathbf{H}_{11} should be equal to the “11” element $[\gamma\mathbf{B}\mathbf{B}^T]_{11}$ of $\gamma[\mathbf{B}][\mathbf{B}]^T$, which can be found from Eq. 10 as

$$[\gamma\mathbf{B}\mathbf{B}^T]_{11} = \gamma \sum_j \cos^2 \alpha_{ij}^x = \gamma \sum_j (X_j - X_i)^2/s_{ij}^2 \quad (26)$$

This is identical to the first diagonal element of \mathcal{H} (see Eq. 22). It can be similarly shown that all elements of $\gamma\mathbf{B}\mathbf{B}^T$ and \mathcal{H} are identical. Therefore, the counterpart of the Kirchhoff matrix Γ of the GNM is simply $(1/\gamma)\mathcal{H}$ in the ANM. Its decomposition yields $3N - 6$ non-zero eigenvalues, and $3N - 6$ eigenvectors that reflect the respective frequencies and shapes of the individual modes.

The inverse of \mathcal{H} is composed of $N \times N$ super-elements, each of which scales with the 3×3 matrix of correlations between the components of pairs of fluctuation vectors. The ij th off-diagonal super-element of \mathbf{H}^{-1} , for example, refers to the cross-correlations between the three different components of $\Delta \mathbf{R}_i$ and $\Delta \mathbf{R}_j$; whereas the i th super-element of \mathbf{H}^{-1} describes the self-correlations between the components of $\Delta \mathbf{R}_i$.

RESULTS AND DISCUSSION

Calculations were performed for a retinol binding protein (pig plasma retinol binding protein, referred to here as RBP) whose crystal structure has been determined at 1.65 Å resolution (Zanotti et al., 1998) (PDB code: 1aqb). RBP is a β -barrel protein of 183 residues. It is composed of eight antiparallel β -strands and an α -helix near the C-terminus. It belongs to the super-family of lipocalins, β -class proteins that bind hydrophobic ligands in their interior (Murzin et al., 1995). RBP transports lipid alcohol vitamin A (retinol).

Cutoff distance and force constant for inter-residue interactions in the ANM

There is a single parameter in the theory, the force constant γ . This parameter is a measure of the strength of intramolecular potentials that stabilize the native fold. In the GNM studies, γ has been estimated for each protein by comparing the theoretically predicted mean-square fluctuations of α -carbons with those indicated by the x-ray crystallographic B-factors, using $\langle (\Delta \mathbf{R}_i)^2 \rangle = 3B_i/(8\pi^2)$. A force constant of 1.0 ± 0.5 kcal/(mol \cdot Å²) has been obtained in several (>50 residues) proteins, assuming an interaction cutoff distance of $r_c = 7$ Å. A force constant of this strength may indeed be viewed as a generic property of proteins modeled as elastic networks. Note that the absolute value of γ does not affect the distribution (or relative size) of residue fluctuations, but only rescales (uniformly) their size.

The cutoff distance of $r_c = 7.0$ Å includes all pairs located within a first coordination shell in the neighborhood of a central residue (Bahar and Jernigan, 1997). Clearly, larger r_c values give rise to an increase in the number of interacting pairs, and consequently would cause a decrease in the amplitudes of fluctuations. A fit to experimental data would then require adopting a weaker force constant for pairwise interactions. Fig. 3 illustrates the dependence of γ on r_c . The dashed and solid curves refer to the results from GNM and ANM theories, respectively, applied to RBP.

Also for relatively short cutoff (r_c) values, the ANM analysis yields more than six zero eigenvalues, and there are extremely large amplitude fluctuations along particular directions for particular residues. To remove such physically unrealistic behavior, we thought we should adopt larger cutoff distances. This reduces the extreme differences among the numbers of neighbors of different residues in different directions and actually gives six zero eigenvalues, as expected. In addition, adoption of the larger r_c values has an additional advantage of yielding spring constants (which were automatically found from the normalization of the results against the experimental B-factors) comparable to those found by the GNM, as illustrated in Fig. 3. Comparison of the results from the two theories indicates that the typical force constant of 1.0 ± 0.5 kcal/(mol \cdot Å²) is

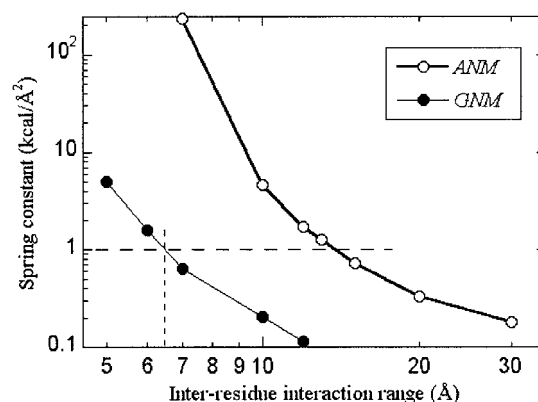


FIGURE 3 Comparison of the force constant for inter-residue potentials in the GNM (filled circles) and ANM (open circles) representations of RBP, as a function of the cutoff distance r_c . Adoption of a larger r_c and thereby a higher number of interacting pairs is required for the ANM in order to reduce the interaction strength (force constant) and match the experimentally observed fluctuation behavior.

reproduced in the ANM by adopting a longer interaction range, mainly 12–15 Å.

Vibrational frequencies

Further support for the adoption of r_c values of 12–15 Å in the ANM is provided by the histograms of vibrational frequencies. Fig. 4 illustrates the distributions of vibrational frequencies for $r_c = 10, 13$, and 20 Å. In the ANM (or GNM), vibrational frequencies are readily found from $\omega_i = (\gamma\lambda_i)^{1/2}$, where λ_i is the i th eigenvalue of $\gamma^{-1}\mathcal{H}$ (or Γ), for $1 \leq i \leq 3N - 6$ (or $N - 1$). We note that the relatively low frequency motions are overemphasized in part *a* of Fig. 4, whereas part *c* is biased toward higher frequency modes. A more realistic Gaussian-like distribution is observed for $r_c = 13$ Å (part *b*), conforming in shape with the densities of vibrational states previously calculated for globular proteins (Elber and Karplus, 1986; Haliloglu et al., 1997). We note that the histogram in part *b* also bears a close resemblance to the one obtained by NMA for TIM barrels (Haliloglu et al., 1997; Kobayashi et al., 1997).

X-ray crystallographic temperature factors

A first test of the validity of the ANM is to compare the predicted ms fluctuations of residues with those observed in experiments. Fig. 5 illustrates the results for RBP. Experimental B-factors, shown by the dashed curve, refer to the crystallographic measurements of Zanotti et al. (1998). Theoretical (solid curve) results are obtained with the ANM by using $r_c = 13$ Å. These are found directly from the diagonal elements of \mathcal{H}^{-1} , after summing up the three diagonal terms of the $N \times N$ super-elements corresponding to each residue. The agreement between experiments and theory is excel-

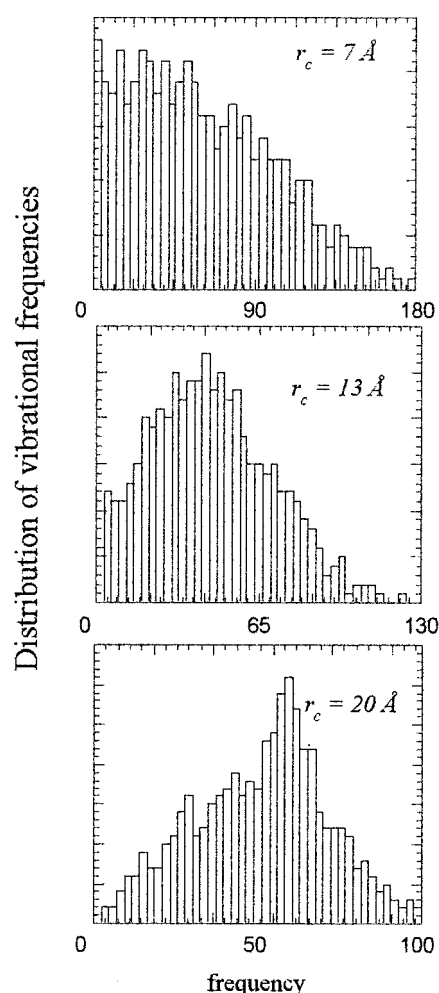


FIGURE 4 Histograms of frequencies of normal modes in the ANM of retinol binding protein (RBP) for three different interaction cutoff distances: (top) $r_c = 70 \text{ \AA}$, (middle) 13 \AA , and (bottom) 20 \AA .

lent, lending support to the use of ANM for further investigation of the dynamics of RBP. Calculations performed with r_c varying in the range $12 \leq r_c \leq 15 \text{ \AA}$ were found to yield almost indistinguishable results, whereas $r_c = 7.0$ yields excessively large (one or more) components ΔX_i , ΔY_i , and ΔZ_i of $\Delta \mathbf{R}_i$ for a number of surface-exposed residues (that have few neighbors along particular directions). For the other extreme case of $r_c \geq 20 \text{ \AA}$, the distributions of fluctuations among residues gradually become more diffuse, in general.

Fig. 6 displays the RBP structure, color-coded according to the sizes of fluctuations of α -carbons in the folded state. The most severely constrained regions (minima in Fig. 5) are colored red, and the most flexible ones (peaks in Fig. 5) are blue. The retinol molecule is shown in white. In the same figure are shown the two principal axes X and Y of the molecule, found by singular value decomposition of the $3 \times N$ matrix of the α -carbon positions of RBP, constructed using the crystal structure coordinates. The Z axis, not

shown, lies along the radial direction of the β -barrel. The X axis, which would be expected to be along the axial direction in the case of a perfectly symmetrical structure, is slightly tilted due to the presence of asymmetric structural units flanking the β -barrel. This molecule-embedded principal axis system is used below as the basis for describing the preferred directions of equilibrium fluctuations.

Anisotropy of motion

In the particular case of isotropic motion the fluctuations in all directions are equal and given by

$$\langle (\Delta X_i)^2 \rangle = \langle (\Delta Y_i)^2 \rangle = \langle (\Delta Z_i)^2 \rangle = \langle (\Delta \mathbf{R}_i)^2 \rangle / 3 \quad (27)$$

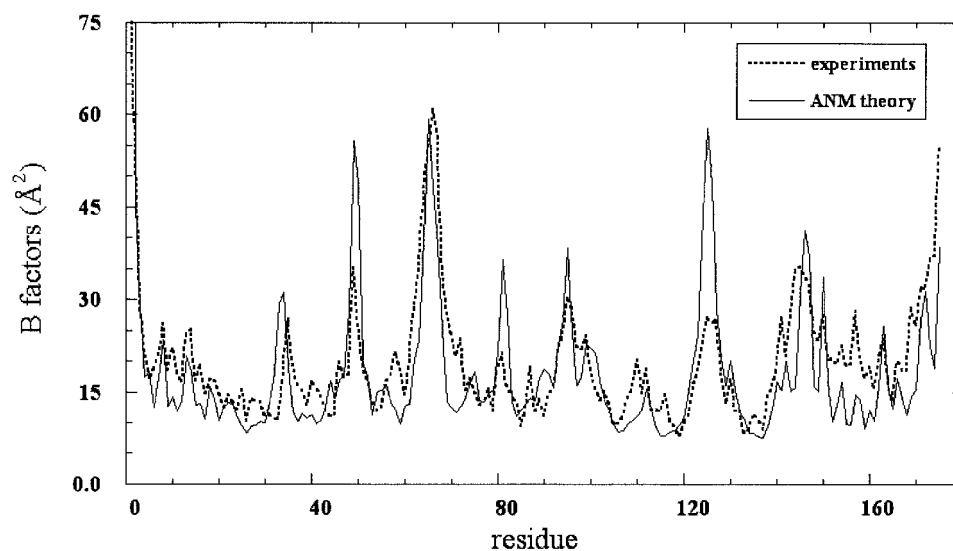
for each residue i . The above equality holds for the GNM. The ANM, by contrast, yields distinctly different fluctuations along each direction. The departure of the above ms components from their expected isotropic values $\langle (\Delta \mathbf{R}_i)^2 \rangle / 3$ provides a measure of the anisotropy of fluctuations. Fig. 7 illustrates these differences for RBP. In the lower part of the figure, rms total fluctuations of individual residues are displayed (left ordinate), and the upper part displays the departures of each ms component $\langle (\Delta X_i)^2 \rangle$, $\langle (\Delta Y_i)^2 \rangle$, and $\langle (\Delta Z_i)^2 \rangle$, from the isotropic value $[\langle (\Delta X_i)^2 \rangle + \langle (\Delta Y_i)^2 \rangle + \langle (\Delta Z_i)^2 \rangle] / 3$.

Using the molecule-embedded principal axes could give an indication of preferential fluctuations along the axial and lateral directions of the β -barrel. However, no such preference is discerned here. The local packing density apparently dominates the observed behavior, in that the strongest departures from isotropy occur in the most loosely packed regions of the structure, which enjoy the highest mobilities. The anisotropy manifested at these regions is actually a consequence of the fact that any tendency to undergo anisotropic fluctuations is not suppressed at these regions. Results obtained after normalization with respect to $\langle (\Delta \mathbf{R}_i)^2 \rangle$ indeed showed that the normalized departures are, instead, almost equally large at either interior or solvent-exposed regions.

Mechanism of dominant global motion

Fig. 8 displays the distributions of fluctuations driven by the global mode of RBP. The curves refer to the X -, Y - and Z -components of the residue displacement vectors driven by the second slowest mode. The first slowest mode was found to involve a large amplitude displacement of the N-terminus, accompanied by an overall rotation of the β -barrel, which emerged as the slowest mode apparently due to the low frequency/large amplitude movement of the tail. The second mode, referred to here as the global mode, however, drives cooperative structural fluctuations relevant to function, as is explained below.

FIGURE 5 Comparison of experimental (Zanotti et al., 1998) and theoretical (ANM) temperature factors obtained for RBP α -carbons, as a function of residue index.



The largest size displacements in the global mode (Fig. 8) are found to be performed by the loop 94–99, at the opening of the β -barrel, displayed in green in Fig. 9. The amplitude of these fluctuations reaches up to 8.7 Å, as calculated from the weighted ($1/\lambda_i$) contribution of the fluctuation vector driven by the global mode. The motion of this loop is accompanied by that of the loop 63–68, shown in yellow in Fig. 9. See also Fig. 10 *a* for the fluctuating conformations associated with this mode. It is interesting to note that the other parts of the molecule undergo fluctuations shorter than 1.0 Å in general, while these two loops move by 7–8 Å.

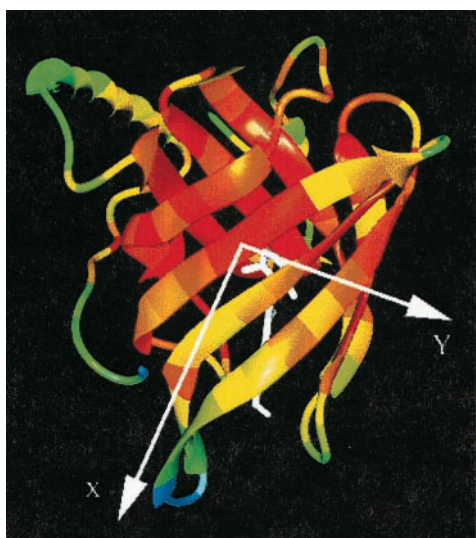
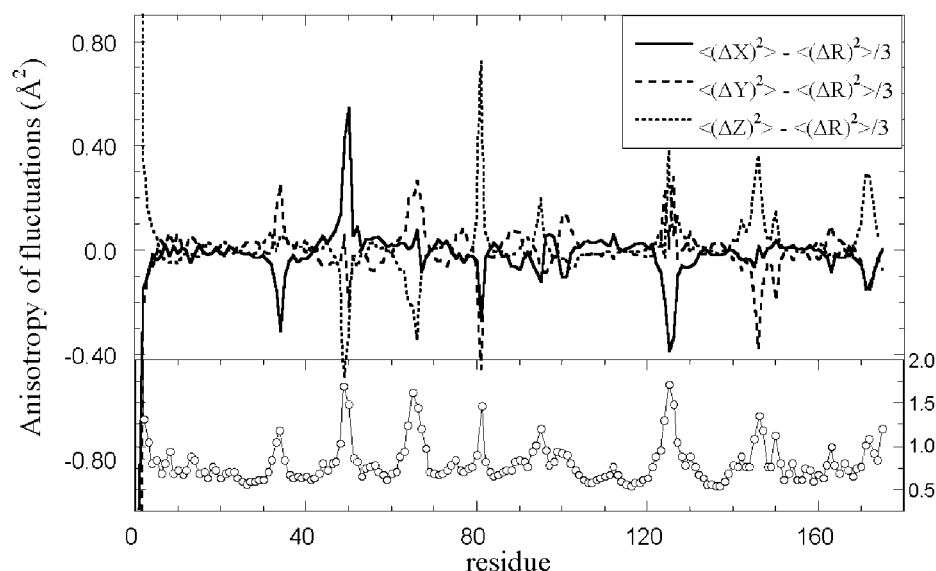


FIGURE 6 Ribbon diagram of RBP color-coded for the degree of flexibility of the individual residues, as indicated by their B-factors (see Fig. 5). The most severely constrained regions are shown in red, and the most flexible parts are in blue. The molecule-embedded principal axes *X* and *Y* are displayed. The *Z* axis (not shown) lies along the radial direction and completes a right-handed coordinate system.

The coupled motions of these two loops suggest that these structural elements cooperate for opening/closing the entrance of the cavity for ligand binding. The ligand (retinol) is shown in red in Fig. 9. Interestingly, these loops contain two aromatic side chains, Phe-97 and Trp-67, which appear to play the role of a gate at the opening of the β -barrel. Residues Gly-92 and Gly-100, whose small size and less restricted dihedral angles are suitable for forming kinks and making conformational transitions, probably ensure the enhanced flexibility of residues 94–99. Two other regions emerging here as the most flexible in the collective modes are residues 48–52 and 120–127, shown in blue and cyan, respectively, in Fig. 9. The amplitudes of fluctuations in these regions were calculated to be around 2.0 Å.

We note that a conserved glycine (Gly-67) of cellular RBP has been pointed out to play a functional role—effectively allowing for hinge bending motions to occur near the entrance of the β -barrel transporter protein—to control retinol intake and release. This was inferred both from biochemical studies (reduced retinol binding of the G67S mutant) and essential dynamics analysis of cellular RBP MD trajectories (van Aalten et al., 1997b). In the presently analyzed plasma RBP, there is another residue (Trp-67) at the same position, and the hinge-bending center is apparently shifted a few residues along the sequence, yet the two RBPs and the bovine serum binding RBP (Chau et al., 1999) are all found to exhibit the same type of collective motion in the slowest, dominant (or essential) modes, mainly hinge-bending motions opening and closing the retinol entry site. The implication of glycine residues being near hinge-bending sites appears to be a common behavior. Remarkably, the loops around residues 35, 65, and 95 were pointed out to move together in the slowest two modes of serum RBPs, controlling the diameter of the tunnel toward the center of the β -barrel (Chau et al., 1999), in agreement

FIGURE 7 In the top part is shown the anisotropy of fluctuations of each residue as measured by the departures of the three components of the ms fluctuations $\langle(\Delta X_i)^2\rangle$, $\langle(\Delta Y_i)^2\rangle$, and $\langle(\Delta Z_i)^2\rangle$, from the isotropic value $\langle(\Delta R_i)^2\rangle/3$. The lower part of the figure displays the total rms fluctuations of each residue $\langle(\Delta R_i)^2\rangle^{1/2}$. Notably the largest departures from isotropy coincide precisely with the regions enjoying the highest total flexibility.



with the present results for the plasma RBP. This essential motion was deduced from the structural variations found in a cluster of crystal structures for the same protein and from the essential dynamics of the C^α atoms positions visited during full atomic MD trajectories. Notably, ANM indicates the same mechanism of motion in the slowest modes, despite the neglect of atomic details and specific energetics.

CONCLUSIONS

A simple analytical approach (ANM) for estimating the mechanism of collective motions in proteins has been presented here. Both this ANM and the previous GNM are based on an elastic network model. The GNM has proven itself to accurately describe the vibrational dynamics of proteins and their complexes in numerous studies. Whereas the GNM is limited to the evaluation of the mean-square displacements and cross-correlations between fluctuations, the motion being projected to a mode space of N dimensions, the ANM approach permits us to evaluate directional

preferences and thus provides 3-D descriptions of the $3N - 6$ internal modes.

In principle, the ANM may be viewed as a simplified form of NMA, in which inter-residue interactions are assumed to be nonspecific. A harmonic potential with a generic force constant γ is assigned to all pairs located within r_c , similar to the distance-dependent generic force constant adopted by Hinsen (Hinsen et al., 1999; Hinsen and Kneller, 1999; Hinsen, 1998). The nonspecificity of inter-residue interactions might seem unrealistic at first sight. However, the main point is that residue specificity, which can play a major role in selecting a fold for a given sequence, is of relatively secondary importance in maintaining the stability of the folded state, once the folded structure has already been formed. At this stage, the molecule has rather solid-like properties, and an elastic network model that incorporates the geometry and distributions of inter-residue contacts as in the model displayed in Fig. 1 can yield an entirely satisfactory description of the molecular machinery imprinted by the particular architecture, even though the pro-

FIGURE 8 Shapes of the slowest collective modes, using $r_c = 13$ Å. The curves display the X-, Y-, and Z-components of the displacement vectors driven by the global mode. Four regions are distinguished by their high mobilities: the loops at the opening end of the β -barrel, shown in green (residues 94–99) and yellow (residues 63–68) in Fig. 9, and residues 48–52 and 120–127, shown in blue and cyan in the same figure.

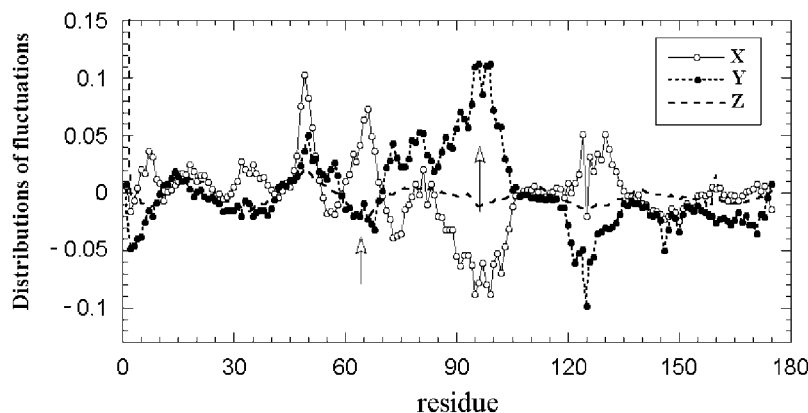




FIGURE 9 Ribbon diagrams of RBP displaying the regions active in the collective slowest modes 1 and 2. (a) Side view and (b) top view of the β -barrel structure displaying the ligand (red), and the regions of the structure exhibiting the largest displacements in the slowest modes: residues 63–68 (yellow) and 94–99 (green), 48–52 (blue) and 120–127 (cyan). Side chains of residues 63–68 and 94–99 are displayed. Note the presence of two highly flexible aromatic residues, Phe-96 (green) and Trp-67 (yellow) at the entrance of the cavity accommodating the substrate.

tein's surface can exhibit more liquid-like characteristics (Zhou et al., 1999).

It is worth recalling that a knowledge-based study has also shown that a single generic potential typifying inter-residue interactions in folded structures plays the major role in determining the total free energy (or stability) of folded structures, with the contributions from specific interactions being far less important (Bahar and Jernigan, 1997).

Direct comparison of the fluctuation dynamics predicted by such single-parameter harmonic potentials and results from detailed NMA have shown that the dynamics predicted by the simplified models closely conform to those found from more detailed treatments. For example, Tirion showed that the residue fluctuations predicted for G-actin with a single-parameter model are almost indistinguishable from

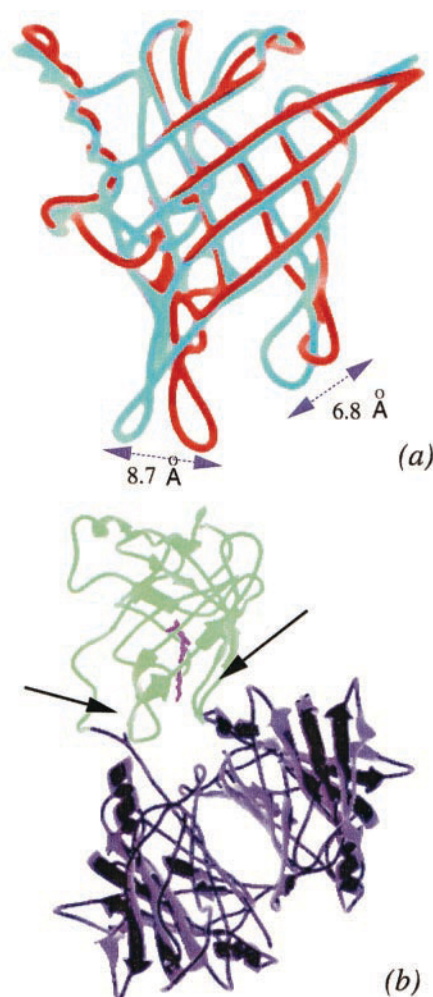


FIGURE 10 (a) RBP conformations visited by the action of the slowest mode (2), illustrating the large amplitude fluctuations of the loops at the entrance of the β -barrel. The fluctuation sizes refer to those of residues 65 (left) and 95 (right) on the loops. (b) Complexation of RBP (green) with its carrier protein transthyretin (purple), as observed in the complex crystallized by Naylor and Newcomer (1999). Note that the most flexible parts in (a) are involved in the recognition of the transthyretin, along with the C-terminus. Arrows indicate protein-protein close contacts. Vitamin A is shown in magenta.

those obtained with a classical NMA using an elaborate force field (Tirion, 1996). Likewise, Hinsen pointed out that a harmonic model taking into account the chemical bond structure and a generic mid-range deformation energy yields a description of protein dynamics around an equilibrium state that is as good as one obtained from detailed molecular mechanics force fields (Hinsen and Kneller, 1999). The density of states obtained from the harmonic potentials in NMA was shown to be as close to the experimental data as was a density of states calculation from a MD simulation with a standard atomic force field (Hinsen and Kneller, 1999). Finally, our recent comparison of full-atomic MD results with those from the coarse-grained GNM and ANM

demonstrated that the collective modes are rather insensitive to the details of the model (Doruker et al., 2000). The robustness of the global modes has indeed been emphasized in previous work (Kitao and Go, 1999).

However, MD simulations include another effect, the anharmonicity of inter-residue interactions, which is present neither in the present linear approach nor in conventional NMA. The absence of anharmonic effects, often claimed to be a serious defect of linear models, was recently shown (Hinsen et al., 1999; Hinsen and Kneller, 1999) to have a negligible influence on the frequency spectrum of small motions around a stable conformation. A more direct examination of the contribution of anharmonic or multimodal fluctuations to protein dynamics has been performed by Go and collaborators (Hayward et al., 1994, 1995). They projected their 200 ps MD trajectories for BPTI onto a space of principal components, and compared the motion associated with each mode to those indicated from NMA. The first principal mode observed in MD was found to deviate from the NMA first mode, the corresponding distribution of fluctuations being bimodal, as opposed to the Gaussian distributions observed in both MD and NMA for the higher modes. Bimodal or multimodal distributions were observed for the slowest mode(s) in other studies, as well (Amadei et al., 1993; Garcia and Harman, 1996), including our recent comparison of MD and GNM/ANM modes (Doruker et al., 2000). However, individual MD trajectories often exhibit different time evolutions, and multiple runs are usually needed to capture a reproducible pattern (Caves et al., 1998), especially when examining the large-amplitude, collective motions. The mode shapes associated with the slowest modes, which exhibit bimodal distributions, may therefore be biased by sampling inefficiencies, which would be overcome by the multiple runs; thus requiring averaging. It is likely that such averaging techniques would result in an overall broadening and smoothing of the energy minimum near the native conformation, and an apparent harmonic behavior; a relatively weak force constant could yield a reasonable first approximation for the average behavior. Not surprisingly, the net effect of neglecting anharmonic modes in NMA appears to be an overall underestimation of the fluctuation amplitudes (with little distortion of the residue distribution of fluctuations). This effect could indeed be removed to a large extent by properly rescaling amplitudes of fluctuations (or the effective force constant). See Fig. 3 in Hayward et al. (1994).

It is worth noting that in a crystal environment large-scale motions are severely restricted. Evidence of this effect comes from a comparison of the solution and crystal structures of the calcium-binding protein calmodulin, for example. NMR experiments show that calmodulin experiences extremely large hinge-bending at the central helix connecting its two lobes (Barbato et al., 1992). This helix is, however, found to be fixed in the x-ray structure. More evidence is offered by the crystal structure of myoglobin,

which does not reveal the cooperative flexibility required for oxygen binding (Mozarelli et al., 1991). Simulations of proteins in noncrystalline environments show large-amplitude nonlinear motions (Garcia, 1992), such as subunit transitions between closed and open forms (Garcia and Harman, 1996). Go and co-workers proposed that the largest scale concerted motions represented by the anharmonic modes are frozen when a protein is in its crystal environment (Hayward et al., 1994). The good agreement between crystallographic temperature factors and the ms fluctuations predicted by ANM calculations likely is partly due to the constraints imposed on large-scale structural changes in the tightly packed environment of crystals.

It remains to be seen whether the regions indicated by the GNM or ANM calculations to act as hinges in the collective modes can be used as pivots about which a global rotational motion could be developed. The directional vectors revealed by the ANM for the slowest mode can be used to this aim. Application to RBP clearly indicates the high mobility of the two loops near the entrance of the cavity, in agreement with the essential dynamics derived from MD simulations (van Aalten et al., 1997b; Chau et al., 1999). Interestingly, the same regions are involved in another function, that of recognition of the carrier protein transthyretin (also referred to as prealbumin) in plasma (Naylor and Newcomer, 1999), suggesting that the high mobility of these regions in the global mode is required not only for the gating action of the vitamin A binding cavity, but also for recognizing its carrier protein. The involvement of the most flexible regions emerging in the slowest, most cooperative collective modes for the recognition of substrates is a phenomenon already revealed in several GNM studies of the equilibrium dynamics of other proteins and their complexes. The flexibility in the collective modes indeed emerges as a prerequisite for the effective functioning of the recognition sites.

Partial support by Bogazici University Research Funds Project 99HA503 is gratefully acknowledged by I.B.

REFERENCES

- Amadei, A., A. B. M. Linssen, and H. J. C. Berendsen. 1993. Essential dynamics of proteins. *Proteins*. 17:412–425.
- Bahar, I., A. R. Atilgan, M. C. Demirel, and B. Erman. 1998a. Vibrational dynamics of folded proteins. Significance of slow and fast modes in relation to function and stability. *Phys. Rev. Lett.* 80:2733–2736.
- Bahar, I., A. R. Atilgan, and B. Erman. 1997. Direct evaluation of thermal fluctuations in proteins using a single parameter harmonic potential. *Fold. Des.* 2:173–181.
- Bahar, I., B. Erman, R. L. Jernigan, A. R. Atilgan, and D. G. Covell. 1999. Examination of collective motions in HIV-1 reverse transcriptase. Examination of flexibility and enzyme function. *J. Mol. Biol.* 285: 1023–1037.
- Bahar, I., and R. L. Jernigan. 1997. Inter-residue potentials in globular proteins and the dominance of highly specific hydrophilic interactions at close separation. *J. Mol. Biol.* 266:195–214.

- Bahar, I., and R. L. Jernigan. 1998. Vibrational dynamics of transfer RNAs: comparison of the free and synthetase bound forms. *J. Mol. Biol.* 281:871–885.
- Bahar, I., and R. L. Jernigan. 1999. Cooperative fluctuations and subunit communication in tryptophan synthase. *Biochemistry*. 38:3478–3490.
- Bahar, I., A. Wallquist, D. G. Covell, and R. L. Jernigan. 1998b. Correlation between native state hydrogen exchange and cooperative residue fluctuations from a simple model. *Biochemistry*. 37:1067–1075.
- Barbato, G., M. Ikura, L. E. Kay, R. W. Pastor, and A. Bax. 1992. Backbone dynamics of calmodulin studied by ¹⁵N relaxation using inverse detected two-dimensional NMR spectroscopy: the central helix is flexible. *Biochemistry*. 31:5269–5278.
- Brooks, B., and M. Karplus. 1983. Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor. *Proc. Natl. Acad. Sci. USA*. 80:6571–6575.
- Brooks, C. L., M. Karplus, and B. M. Pettitt. 1988. A theoretical perspective of dynamics, structure, and thermodynamics. *Adv. Chem. Phys.* 71.
- Caves, L. S. D., J. D. Evanseck, and M. Karplus. 1998. Locally accessible conformations of proteins: multiple molecular dynamics simulations of proteins. *Protein Sci.* 7:649–666.
- Chau, P.-L., D. M. F. van Aalten, R. P. Bywater, and J. B. C. Findlay. 1999. Functional concerted motions in the bovine serum retinol-binding protein. *J. Comp.-Aided Mol. Des.* 13:11–20.
- Damaschun, G., H. Damaschun, K. Gast, and D. Zirwer. 1999. Proteins can adopt totally different folded conformations. *J. Mol. Biol.* 291:715–725.
- Demirel, M. C., A. R. Atilgan, R. L. Jernigan, B. Erman, and I. Bahar. 1998. Identification of kinetically hot residues in proteins. *Protein Sci.* 7:2522–2532.
- Demirel, M. C., I. Bahar, and A. R. Atilgan. 1999. Predicting the ensemble of unfolding pathways for proteins: an updated incremental Lagrangian model. *Biophys. J.* 76:176a (Abstr.).
- Doi, M., and S. F. Edwards. 1986. *The Theory of Polymer Dynamics*. Clarendon Press, Oxford Science Publications, New York.
- Doruker, P., A. R. Atilgan, and I. Bahar. 2000. Dynamics of proteins predicted by molecular dynamics simulations and analytical approaches: application to α -amylase inhibitor. *Proteins*. 40:512–524.
- Elber, R., and M. Karplus. 1986. Low frequency modes in proteins: use of effective-medium approximation to interpret the fractal dimension observed in electron-spin relaxation measurements. *Phys. Rev. Lett.* 56:394–397.
- Flory, P. J. 1976. Statistical thermodynamics of random networks. *Proc. R. Soc. Lond. A*. 351:351–380.
- Frauenfelder, H., and B. McMahon. 1998. Dynamics and function of proteins: the search for general concepts. *Proc. Natl. Acad. Sci. USA*. 95:4795–4797.
- Frauenfelder, H., S. G. Sligar, and P. G. Wolynes. 1991. The energy landscapes and motions of proteins. *Science*. 254:1598–1603.
- Garcia, A. E. 1992. Large-amplitude nonlinear motions in proteins. *Phys. Rev. Lett.* 68:2696–2699.
- Garcia, A. E., and J. G. Harman. 1996. Simulations of CRP:(cAMP)₂ in noncrystalline environments show a subunit transition from the open to the closed conformation. *Protein Sci.* 5:62–71.
- Go, N., T. Noguti, and T. Nishikawa. 1983. Dynamics of a small globular protein in terms of low-frequency vibrational modes. *Proc. Natl. Acad. Sci. USA*. 80:3696–3700.
- Haliloglu, T., and I. Bahar. 1999. Structure-based analysis of protein dynamics. Comparison of theoretical results for hen lysozyme with x-ray diffraction and NMR relaxation data. *Proteins*. 37:654–667.
- Haliloglu, T., I. Bahar, and B. Erman. 1997. Gaussian dynamics of folded proteins. *Phys. Rev. Lett.* 79:3090–3093.
- Hayward, S., A. Kitao, and N. Go. 1994. Harmonicity and anharmonic aspects in the dynamics of BPTI: a normal mode analysis and principal component analysis. *Protein Sci.* 3:936–943.
- Hayward, S., A. Kitao, and N. Go. 1995. Harmonicity and anharmonicity in protein dynamics: a normal-mode analysis and principal component analysis. *Proteins*. 23:177–186.
- Hinsen, K. 1998. Analysis of domain motions by approximate normal mode calculations. *Proteins*. 33:417–429.
- Hinsen, K., and G. R. Kneller. 1999. A simplified force field for describing the vibrational protein dynamics over the whole frequency range. *J. Chem. Phys.* 111:10766–10769.
- Hinsen, K., A. Thomas, and M. J. Field. 1999. Analysis of domain motions in large proteins. *Proteins*. 34:369–382.
- Ichiye, T., and M. Karplus. 1987. Anisotropy and anharmonicity of atomic fluctuations in proteins: analysis of a molecular dynamics simulations. *Proteins*. 2:236–239.
- Jernigan, R. L., M. C. Demirel, and I. Bahar. 1999. Relating structure to function through the dominant modes of motion of DNA topoisomerase II. *Int. J. Quantum Chem.* 75:301–312.
- Keskin, O., R. L. Jernigan, and I. Bahar. 2000. Proteins with similar architecture exhibit common large scale dynamics. *Biophys. J.* 78:2093–2106.
- Kitao, A., and N. Go. 1999. Investigating protein dynamics in collective coordinate space. *Curr. Opin. Struct. Biol.* 9:164–169.
- Kobayashi, N., T. Yamato, and N. Go. 1997. Mechanical property of a TIM-barrel protein. *Proteins*. 28:109–116.
- Kuriyan, J., G. A. Petsko, R. M. Levy, and M. Karplus. 1986. Effect of anisotropy and anharmonicity on protein crystallographic refinement. An evaluation by molecular dynamics. *J. Mol. Biol.* 190:227–254.
- McCammon, A., and S. C. Harvey. 1987. *Dynamics of Proteins and Nucleic Acids*. Cambridge University Press, Cambridge.
- Mozarelli, A., C. Rivetti, G. L. Rossi, E. R. Henry, and W. A. Eaton. 1991. Crystals of the haemoglobin with the T quaternary structure bind oxygen noncooperatively with the Bohr effect. *Nature*. 351:416–419.
- Murzin, A. G., S. E. Brenner, T. Hubbard, and C. Chothia. 1995. SCOP: a structural classification of protein databases for the investigation of sequences and structures. *J. Mol. Biol.* 247:536–540.
- Naylor, H. M., and M. E. Newcomer. 1999. The structure of human retinol-binding protein (RBP) with its carrier protein transthyretin reveals an interaction with the carboxy terminus of RBP. *Biochemistry*. 38:2647–2653.
- Rouse, P. E. 1953. A theory of the linear viscoelastic properties of dilute solutions of coiling polymers. *J. Chem. Phys.* 21:1272–1280.
- Tirion, M. M. 1996. Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Phys. Rev. Lett.* 77:1905–1908.
- Van Aalten, D. M. F., D. A. Conn, B. L. de Groot, H. J. C. Berendsen, J. B. C. Findlay, and A. Amadei. 1997a. Protein dynamics derived from clusters of crystal structures. *Biophys. J.* 73:2891–2896.
- Van Aalten, D. M. F., P. C. Jones, M. de Sousa, and J. B. C. Findlay. 1997b. Engineering protein mechanics: inhibition of concerted motions of the cellular retinol binding protein by site-directed mutagenesis. *Protein Eng.* 10:31–37.
- Zaccai, G. 2000. How soft is a protein? A protein dynamics force constant measured by neutron scattering. *Science*. 288:1604–1607.
- Zanotti, G., M. Panzavorto, A. Marcato, G. Malpeli, C. Folli, and R. Berni. 1998. Structure of pig plasma retinol-binding protein at 1.65 Å resolution. *Acta Crystallogr. D*. 54:1049–1052.
- Zhou, Y., D. Vitkup, and M. Karplus. 1999. Native proteins are surface-molten solids: application of the Lindemann criterion for the solid versus liquid state. *J. Mol. Biol.* 285:1371–1375.