PROTEINS: Structure, Function, and Genetics 27:425–437 (1997)

Model-Free Methods of Analyzing Domain Motions in Proteins From Simulation: A Comparison of Normal Mode Analysis and Molecular Dynamics Simulation of Lysozyme

Steven Hayward,[1] Akio Kitao,[2] and Herman J.C. Berendsen[1]*
[1]*BIOSON Research Institute, Laboratory of Biophysical Chemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands*
[2]*Department of Chemistry, Faculty of Science, Kyoto University, Kitashirakawa, Sakyo-ku, Kyoto 606, Japan*

ABSTRACT **Model-free methods are introduced to determine quantities pertaining to protein domain motions from normal mode analyses and molecular dynamics simulations. For the normal mode analysis, the methods are based on the assumption that in low frequency modes, domain motions can be well approximated by modes of motion external to the domains. To analyze the molecular dynamics trajectory, a principal component analysis tailored specifically to analyze interdomain motions is applied. A method based on the curl of the atomic displacements is described, which yields a sharp discrimination of domains, and which defines a unique interdomain screw-axis. Hinge axes are defined and classified as twist or closure axes depending on their direction. The methods have been tested on lysozyme. A remarkable correspondence was found between the first normal mode axis and the first principal mode axis, with both axes passing within 3 Å of the alpha-carbon atoms of residues 2, 39, and 56 of human lysozyme, and near the interdomain helix. The axes of the first modes are overwhelmingly closure axes. A lesser degree of correspondence is found for the second modes, but in both cases they are more twist axes than closure axes. Both analyses reveal that the interdomain connections allow only these two degrees of freedom, one more than provided by a pure mechanical hinge. Proteins 27:425-437, 1997.**

Key words: hinge bending; hinge axis; principal component analysis; essential dynamics

INTRODUCTION

Most proteins are constructed from secondary and super-secondary structures connected together by often less definable structures of varying flexibilities. These different levels of structure with their various degrees of rigidity give proteins a large range of time scales for their dynamics. It has been shown that for bovine pancreatic trypsin inhibitor, the Young's modulus for an alpha-helix is two orders of magnitude higher than that of the whole protein, giving a picture of a protein being constructed from rigid elements connected by softer elements.[1] Many large proteins are built from domains, and interdomain motions are likely to be the slowest of all motions in proteins. Functional sites are often located at interdomain clefts, implying that these interdomain motions are of functional significance.[2] It has been suggested that domain motions, which have been described as hinge, shear, or a combination of both hinge and shear motions, correspond to low energy conformational changes available to the protein at its functioning temperature.[3] This view has support from the crystallographic work on a mutant T4 lysozyme that crystallizes in five different conformations, each with a different hinge angle between the two domains.[4]

In the field of computer simulation of proteins, molecular dynamics (MD) simulation and normal mode analysis are the most common methods used to provide information on protein native state dynamics. Normal mode analysis provides a set of normal modes that describe collective motions at different frequencies.[5–7] Although it suffers from some drawbacks, being a harmonic analysis, it has provided a great deal of insight into the nature of collective motions in proteins[8,9] and has been found to give an accurate description of the motion of lysozyme in the crystal.[10,11] In MD simulations it is thought that multiple minima are visited. The motions occurring as a result of transitions between these minima may be considerably greater and different from those predicted by a normal mode analysis performed within a single minimum. Although the precise relationship between the two methods is still not fully understood, particularly regarding the low-

*Correspondence to: Professor H.J.C. Berendsen, Biophysical Chemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands.
Received 21 June 1996; accepted 30 September 1996.

frequency motions, a quantitative correspondence between a normal mode analysis and an MD simulation where multiple minima are visited has been demonstrated.[12,13] One further point concerns the robustness of the normal mode analysis in response to differences in the minimum energy conformation. It has been shown that the normal modes and frequencies vary little over minima generated from an MD simulation.[14]

The method of principal component analysis has been applied to MD trajectories to determine those collective motions that contribute most significantly to the total (possibly mass-weighted) mean-square fluctuation (msf) of a protein's motion.[15–19] The resulting modes, termed in the literature "principal modes," "quasi-harmonic modes," or "essential modes" also describe collective motions and would be expected to give a more accurate description of the actual motions occurring in proteins than the normal modes.

In this work we study only the modes with the highest msf's to discover more about the nature of the motions they describe. The study of these modes alone can be justified by the fact that the motions they describe dominate the overall motion to such an extent that their study should reveal, in a real sense, the nature of a protein's motion and in turn its relation to function.

The method developed here is one that assumes that some proteins can be regarded, in the first approximation, to consist of parts that can be treated as behaving dynamically as rigid bodies in the low-frequency modes. This approximation requires that the interaction between these parts be small when compared with interactions within a single part. This method, then, may be particularly suited to proteins built of domains. It should, in principle, be able to identify domains, determine the location and orientation of any existing hinge axes, and also find the residues involved in the hinge bending. The method of identification of domains belongs to the group of methods that are based on conformational change for their identification of domains rather than those methods based on purely structural information.

To test the method we compare the results from a normal mode analysis and an MD simulation of lysozyme, a protein that has received considerable attention in the past.[10,11,19,20–23] As mentioned above, the gulf between normal mode analysis, harmonic analysis performed in vacuum, and MD simulation in water is considerable. Therefore, if we were to find a good correspondence in location and orientation of the hinge axis from the two approaches, it would suggest that the hinge axis is determined by factors insensitive to approximations in either of the methods. This would give us the confidence to believe that it is close to the true hinge axis.

A related method has recently been developed for analysis of protein structures found in more than one X-ray conformer.[24]

METHODS

Theoretical Basis

Approximation of low-frequency normal modes in terms of external normal modes of quasi-independent parts

We express the formalism in terms of the normal mode analysis, although it can equally apply to a quasi-harmonic analysis. Let us assume that the Hessian can be written, by appropriate rearrangement of the atom order, as a matrix where the values of the elements in the off-diagonal blocks are small in comparison with the elements in the diagonal blocks. For example as:

$$H = \begin{bmatrix} \blacksquare & \blacksquare & & \\ \blacksquare & \blacksquare & & \\ & & \blacksquare & \blacksquare \\ & & \blacksquare & \blacksquare \end{bmatrix}$$

Each diagonal block will correspond to a part of the protein, and elements within each diagonal block will represent interactions between atoms within each part. Off-diagonal blocks correspond to interactions between parts. If these interactions are small enough, we can consider them as a perturbation on the system of non-interacting parts. In such a case the Hessian of the unperturbed system will be:

$$H_0 = \begin{bmatrix} \blacksquare & & & \\ & \blacksquare & & \\ & & \blacksquare & \\ & & & \blacksquare \end{bmatrix}$$

Taking each block individually, we know that there are six normal modes, each with an eigenvalue of

zero, describing the external motions of the corresponding part of the protein. If $\mathbf{H_o}$ consists of M blocks then there will be 6M modes describing the external motion of the corresponding parts of the protein. However, 6 of these will describe the external motion of the molecule as a whole. So there are 6(M-1) modes that describe the motions of the parts with respect to each other in the internal coordinate system. In the unperturbed case of $\mathbf{H_o}$ the modes are degenerate with eigenvalues 0, but the perturbation by the off-diagonal blocks lifts this degeneracy. The eigenvectors in the zeroth-order of the perturbation expansion are particular linear combinations of the 6(M-1) degenerate eigenvectors. Considering one of the normal modes of the perturbed system, we show in Appendix A that, in the zeroth-order of approximation, the displacement of the *ith* atom (in part k) for unit displacement of the normal mode variable is given as:

$$\mathbf{\Delta d}_i = \mathbf{\Delta \tau}^k + \mathbf{\Delta \psi}^k \times \mathbf{r}_i \qquad (1)$$

where $\Delta\tau^k$ and $\Delta\psi^k$ are constants for the *kth* part, and r_i is the position vector from the center of mass of the *kth* part to the *ith* atom. We see that Eq. (1) describes a rigid body motion of the *kth* part combining a translation and rotation about an axis through the center of mass.

Screw axis and interdomain motion

Eq. (1) implies that the infinitesimal displacement of an atom in a rigid body can be written in terms of the translation of the rigid body and its rotation about a single axis through the center of mass.

In 1763 Mozzi and Cauchy showed that any infinitesimal rigid body motion is a screw motion, i.e., a rotation about an axis and a translation along that axis.[25] Actually this is also true for finite rigid body displacements, in the sense that any rigid body displacement can be reproduced by a screw motion about an appropriately located screw axis. This was shown in the last century by Chasles[26] and has been used to determine the hinge axes of proteins from X-ray conformers where large angle rotations are observed.[24,27] *This means we can always find a rotation axis about which the motion appears as a pure screw motion.* This axis is unique. We can therefore rewrite Eqn. (1) as (dropping the subscript and superscript):

$$\mathbf{\Delta d} = \Delta\tau_{\parallel}\mathbf{n} + \Delta\psi\mathbf{n} \times \mathbf{r}^s \qquad (2)$$

where $\mathbf{n}$ is the unit vector in the direction of the screw axis, $\Delta\tau_{\parallel}$ is the translation along the axis, $\Delta\psi$ the angle of rotation, and $\mathbf{r}^s$ the position vector joining any point of origin on the screw axis to the atom concerned. The screw axis is the only rotation axis of the motion that is fixed in space. Points on this axis are those that are minimally displaced. If there is no translational motion along the axis, then all points on the axis are stationary.

Although we have been referring to "parts" of the protein, we have not related these parts to any specific structures of a protein. In this paper we are interested in the domain motion of lysozyme and so we will now develop the method specifically for a two-domain protein. As we see above, each domain will have its own screw axis. Knowing the screw axes of the domains, how can we determine any physically meaningful axis related to the interdomain motion? Selecting domain 2, say, and knowing its screw axis, we can perform an oppositely directed screw motion, $-\mathbf{\Delta d}_2$, on the *whole* protein such that domain 2 becomes stationary. We will then have the motion of domain 1 relative to domain 2. If $\mathbf{\Delta u}$ is the vector connecting the domain 1 screw axis origin to the domain 2 screw axis origin such that $\mathbf{r}_2^s = \mathbf{r}_1^s - \mathbf{\Delta u}$, then the displacement vectors of domain 1 will now be given by:

$$\mathbf{\Delta D}_1 = \mathbf{\Delta d}_1 - \mathbf{\Delta d}_2 = (\Delta\tau_{\parallel}^1\mathbf{n}_1 - \Delta\tau_{\parallel}^2\mathbf{n}_2 + \Delta\psi_2\mathbf{n}_2 \times \mathbf{\Delta u}) + (\Delta\psi_1\mathbf{n}_1\,[-]\,\Delta\psi_2\mathbf{n}_2) \times \mathbf{r}_1^S. \qquad (3)$$

This has the form of Eq. (1), which can also be transformed to the form of Eq. (2) to give the "interdomain screw axis." We see immediately from Eq. (3) that the rotation vector for this interdomain motion is $\Delta\psi_1\mathbf{n}_1 - \Delta\psi_2\mathbf{n}_2$. If, when domain 2 is fixed in space, domain 1 performs a screw motion about some physical (sliding) hinge axis located between the two domains, then this hinge axis will be the interdomain screw-axis as given in Eq. (3). Depending on the nature of the interdomain motion, it is possible for the interdomain screw axis to be located away from any physical connection between the two domains, even outside the body of the protein itself. This will occur if the interdomain connections allow translational motion perpendicular to any rotation axis. In such a case the interdomain screw axis will have no direct relationship to anything physical. The location of the interdomain screw axis will therefore tell us something about the nature of the domain motion. If it is located between the two domains near regions known to be involved in the interdomain motion, we could reasonably assume that the interdomain connections effectively create a physical hinge axis. If it were located elsewhere, away from those regions known to be involved in the interdomain motion, then we would know that the interdomain connections allow considerable translation in the direction perpendicular to the rotation. This is the essence of the method: its application reveals the nature of the domain motion.

Note that the location and orientation of the interdomain screw axis is independent of the choice of domain to be fixed in space.

Practical Determination of Dynamical Domains and Screw Axes

Dynamical domains

Although for some proteins the domain structure is obvious, for others, domains may exist in a dynamical sense, but may not be easily recognizable from the structure. The theoretical basis of our approach shows that in some low-frequency normal modes, parts of the protein may be behaving, to a certain approximation, as rigid bodies. These parts may or may not have any correspondence to structurally defined domains and to distinguish them from structurally defined domains we have termed them "dynamical domains." Below we describe a method to determine these dynamical domains.

From Eq. (2) we see that $\mathbf{\Delta d}$ is a vector function of $\mathbf{r}^s$ and so we can apply standard vector analysis to determine $\Delta\psi\mathbf{n}$ by calculating the curl of $\mathbf{\Delta d}(\mathbf{r}^s)$:

$$\frac{1}{2}\boldsymbol{\nabla} \times \mathbf{\Delta d}(\mathbf{r}^s) = \Delta\psi\mathbf{n}. \tag{4}$$

Here $\mathbf{\Delta d}(\mathbf{r}^s)$ is a vector field (sampled by atoms) and $\boldsymbol{\nabla} \times \mathbf{\Delta d}(\mathbf{r}^s)$ is the curl of that vector field. Provided we have a method to calculate the curl we can find the unit vector in the direction of the rotation axis. The practical calculation of the curl at each residue is described in Appendix B. The curl should give a similar value over each dynamical domain provided the eigenvector can be approximated by components of the form of Eq. (1).

If there are two domains rotating in largely opposite directions, as they must do to satisfy the condition of no overall rotation, then the $\mathbf{n}$'s in one domain will be generally opposite in direction to the $\mathbf{n}$'s in the other domain. Therefore by taking inner products between the $\mathbf{n}$'s at different locations over the whole protein to generate a "rotation-orientation matrix," we should be able to identify the dynamical domains. If there are two dynamical domains for which the rotation vectors are zero because the two domains are translating in an anticorrelated fashion as in the oscillation of a bond, then the taking of inner products of the mode vector components themselves at different locations over the whole protein will reveal the dynamical domains.

Determination of other parameters

We already have a method to determine $\Delta\psi\mathbf{n}$ at each residue, but we still need to determine the location of the screw axis for the whole domain. To determine the screw axis of a dynamical domain accurately, any internal motions within the domain should first be removed. This can be done by doing all atom mass-weighted least-squares best fits on each dynamical domain, as has been done previously to determine the external components of the motions of secondary structure elements in bovine pancreatic trypsin inhibitor.[1] To do this we shift each atomic coordinate in the domain in the direction of its normal mode component by its root-mean-square fluctuation (rmsf) in that normal mode, to give a new conformation. By best-fitting the original unshifted conformation to this new one, one can determine the external component to motion as given by the atomic displacement vectors between the atoms in their original unshifted position and in the best-fitted position. By doing this for both domains one can also calculate the relative contributions to the total mass-weighted msf of the interdomain motion and intradomain motion, for:

$$\sum_i |\mathbf{\Delta q}_i|^2 = \sum_i |\mathbf{\Delta q}_i^{ext}|^2 + \sum_i |\mathbf{\Delta q}_i^{int}|^2 \tag{5}$$

where $\mathbf{\Delta q}_i = \sqrt{m_i}\mathbf{\Delta d}_i$. This means we can simply calculate the proportion contributed by the interdomain motion of the domains as:

$$\frac{\sum_i |\mathbf{\Delta q}_i^{ext}|^2}{\sum_i |\mathbf{\Delta q}_i|^2}. \tag{6}$$

As the internal component in each domain has now been removed, the rotation vector determined by calculating the curl will give the same result at all residues within the domain. This value will be the rotation vector of the dynamical domain, $\Delta\psi\mathbf{n}$.

We now know $\mathbf{\Delta d}$ and $\mathbf{n}$ of Eq. (2), but we still need to know the location of the screw axis. We can easily determine $\Delta\tau_\parallel$ as:

$$\Delta\tau_\parallel = \mathbf{\Delta d}.\mathbf{n}. \tag{7}$$

As Eq. (2) is valid for $\mathbf{r}^s$ joining any point on the screw axis to the atom concerned, if we choose the origin such that $\mathbf{r}^s$ is perpendicular to the axis, then $\mathbf{r}^s$, $\mathbf{n}$, and $\mathbf{\Delta}$d $- \Delta\tau_\parallel\mathbf{n}$ are all perpendicular, and so from Eq. (2):

$$\mathbf{r}^s = (\mathbf{\Delta d} - \Delta\tau_\parallel\mathbf{n}) \times \frac{\mathbf{n}}{\Delta\psi}. \tag{8}$$

Knowing $-\mathbf{r}^s$ we know a point on the screw axis relative to the atom concerned and have therefore fixed the axis in space. Each dynamical domain has its own screw axis, and, using the methods that lead to Eq. (3), we can locate the interdomain screw axis and parameters relevant to the interdomain motion.

We summarize the steps involved in determining the interdomain screw axis of a protein from a normal mode eigenvector:

1. Determine the rotation vectors at each residue by calculating the curl as described in Appendix B.

2. Take the inner products between all the unit vectors, **n**, calculated at every residue of the protein to generate the rotation-orientation matrix. A plot of this matrix should reveal any existing dynamical domains.
3. Once the dynamical domains have been identified, residual internal motions within each domain are removed using all atom mass-weighted best fits. This calculation also provides the percentage contribution to the total msf of the interdomain motion.
4. The screw axis of one domain is located.
5. The motion in this domain is removed by performing an opposing motion on the whole protein. Note in the case of a protein comprising three or more domains, the motion of any domain can be removed and the motion of the other domains relative to this domain can be analyzed as described below.
6. The screw axis of the resulting motion in the other domain is determined. This is the interdomain screw axis.

This whole procedure involves little computation, and human judgment is required only at step 2.

Meaning of $\Delta\psi$ and $\Delta\tau_\parallel$

We now have a method to determine $\Delta\psi$ and $\Delta\tau_\parallel$, but as yet we have not related them to anything physically meaningful. We know, however, that:

$$\mathbf{\Delta d}(t) = (\Delta\tau_\parallel \mathbf{n} + \Delta\psi \mathbf{n} \times \mathbf{r}^s)\sigma(t) \qquad (9)$$

where $\sigma(t)$ is the normal mode variable. To evaluate the msf in translation and rotation, we rewrite this expression as:

$$\mathbf{\Delta d}(t) = \Delta\eta_\parallel(t)\mathbf{n} + \Delta\theta(t)\mathbf{n} \times \mathbf{r}^s. \qquad (10)$$

We see that $\Delta\theta(t)$ gives the magnitude of the angular rotation at time t and $\Delta\eta_\parallel(t)$ the magnitude of the translation along the axis. Given that for the normal mode analysis

$$\sqrt{\langle\sigma(t)^2\rangle} = \sqrt{\frac{k_B T}{\omega^2}} \qquad (11)$$

where ω is the angular frequency of the activated mode; we can then write:

$$\sqrt{\langle\Delta\theta(t)^2\rangle} = \sqrt{\frac{k_B T}{\omega^2}}\,\Delta\psi \qquad (12)$$

and

$$\sqrt{\langle\Delta\eta_\parallel(t)^2\rangle} = \sqrt{\frac{k_B T}{\omega^2}}\,\Delta\tau_\parallel. \qquad (13)$$

Therefore, from the knowledge of $\Delta\psi$ and $\Delta\tau_\parallel$, we can calculate the rmsf of quantities pertinent to the interdomain motion in the harmonic approximation.

Determination of Residues Involved in Interdomain Motion

Usually a number of residues are involved in the motion of one domain relative to another. For example, this may involve the bending of an interdomain helix. How can one determine these residues? Once the direction of the interdomain screw axis has been determined, the component of the rotation vector of each residue along the direction of the interdomain screw axis can easily be calculated. Residues involved in interdomain motion should be those located in the region where one sees a transition between positive and negative values of the components. These will be those located in the regions connecting the dynamical domains. If the interdomain screw axis is located near residues of a connecting region shown to be involved in the motion, then this implies the existence of a physical axis located between the two domains. Only in this case will we call the axis a *hinge axis.*

Principal Component Analysis for Interdomain Motions

In this study we have applied two principal component analyses or essential dynamics analyses to the MD trajectory (we do not mass-weight the coordinates). As we shall see below, a conventional application of the principal component analysis, was found to be inadequate.[9] Therefore, a new form of principal component analysis was developed, tailored to the problem of analyzing interdomain motions. This method requires that one has already identified the dynamical domains. It involves first best-fitting the conformations of domain 1 at each time frame to a single reference conformation of domain 1, repositioning the whole protein accordingly. This will yield a trajectory of the motion of domain 2 relative to domain 1. The trajectory of domain 2 will still contain internal motion and so to remove this, a single reference conformation of domain 2 is fitted to each frame of the trajectory of domain 2. We now have the trajectory of domain 2 moving as a rigid body relative to domain 1. A principal component analysis of this trajectory will give just six non-zero eigenvalues corresponding to the motion of domain 2 as a rigid body relative to domain 1. These modes can be analyzed directly to give the interdomain screw axis. Again, the result will not depend on which domain we choose to fix.

Twist Axis versus Closure Axis

In accordance with our intuition we define a hinge axis located along, or parallel to, the line joining the centers of mass of the two domains, a *twist axis.* We further define a hinge axis perpendicular to the line

joining the centers of mass as a *closure axis*. A twist axis is one that will slide groups between the two domains past each other, and, depending on the shape of the surface, may also cause the closure of some hydrophobic pocket. The motion implied by a closure axis is the most effective way of bringing the centers of mass of the two domains closer together, thus bringing some groups between the two domains into eventual contact. The unit vector of the rotation vector corresponding to any hinge axis can be decomposed into a component parallel to the line joining the centers of mass of the two domains, g_t, and another perpendicular to this line, g_c. Therefore

$$g_c^2 + g_t^2 = 1 \quad (14)$$

and so we can always say that rotation is $g_c^2 \times 100\%$ closure motion.

Normal Mode Analysis and MD Simulation

The authors of this study were pleased to receive the results of 800 ps MD simulation of hen lysozyme in water at 300 K from Alan Mark in Zurich.[28] The simulation used the GROMOS force-field parameters.[29] The normal mode analysis was performed on human lysozyme in cartesian coordinates using PRESTO[30] and the AMBER force field parameters,[31] which are similar to the GROMOS force-field parameters. No cutoff for the non-bonded interactions was used.

RESULTS AND DISCUSSION

In this work we are comparing results from human and hen lysozymes. Human and hen lysozymes have 130 and 129 residues, respectively. Human lysozyme has an extra residue inserted between residues 46 and 47 of hen, in the sense that residues 1–46 of both hen and human are structurally similar, as are residues 47–129 of hen and 48–130 of human. A best fit of the alpha-carbon (CA) atoms, fitting CA *i* in hen with CA *i* in human for $i = 1$–46, and CA *i* in hen with $i + 1$ in human for $i = 47$–129 gives an rms deviation of 0.65 Å, showing how very similar their structures are.

Normal Mode Analysis of Human Lysozyme

As we have mentioned in the Introduction, analysis of the low-frequency normal modes alone can be justified by the fact that they dominate the overall motion. The first two normal modes of lysozyme (out of a total of 6,057 modes) account for 17% of the total mass-weighted msf. Although this tells us that most fluctuation still comes from other modes, these modes taken individually have significantly smaller msfs than the lowest frequency modes, and in addition represent localized motions, rather than collective motions spanning the whole protein.[1]

Determination of the dynamical domains

Figure 1 shows the x, y, and z components of the rotation vector for each residue from the first normal mode. The average rms angular fluctuation is 1.17°. Figure 2 shows a plot of the rotation-orientation matrix calculated from the first and second normal modes. The clear blocking shows how well this method is able to locate the region of demarcation between the two domains from the normal mode analysis. From Figure 2 we can deduce that the demarcation between the two domains for human lysozyme is located at residue 39 and residues 93–95. On the basis of this result let us define domain 1 as comprising residues 1–38 together with residues 96–130, and domain 2 comprising residues 40–91. Figure 3 shows these domains in relation to the backbone structure.

The rotation-orientation matrix plots for higher modes do not reveal a simple two-domain demarcation. For these higher frequency modes, the protein fractures into a greater number of domains, which could also be analyzed using the methods developed here.

In revealing the demarcation between the two domains, the clarity of the rotation-orientation matrix plot of Figure 2 stands in contrast to the results from other methods, such as a difference-distance matrix method[22,32] or the atomic-covariance matrix method,[10] which are also methods based on conformational change to determine domains. The plots of these matrices do not reveal such clear blocking. The difference-distance matrix method is based on the idea that intradomain atomic distances should be largely unchanged in comparison with interdomain atomic distances. However, depending on the location of the hinge axis, even interdomain atomic distances may also remain unchanged, and the plot may be unclear as regards the demarcation of the domains. This will be particularly true if the axis is located between the two domains. This problem also applies to the atomic-covariance matrix plot.

First normal mode interdomain screw axis

The external component to the motion of each domain was determined as described in the Methods section. The horizontal lines in Figure 1 show the x, y, and z components of the rotation vector at each residue when the first normal mode has been decomposed so that it now only contains the relative motion of the two domains as rigid bodies. The relative motion of the two domains as rigid bodies accounts for 85% of the total mass-weighted msf of first normal mode motion. The interdomain screw axis, shown in yellow in Figure 3, passes within 3 Å of the CA atoms of residues 1, 2, 39, 40, 56, and 92. One immediately notices that the axis passes near the residues located at the demarcation region between the two domains as seen in Figure 2. The continuous line in Figure 4 reveals in more detail

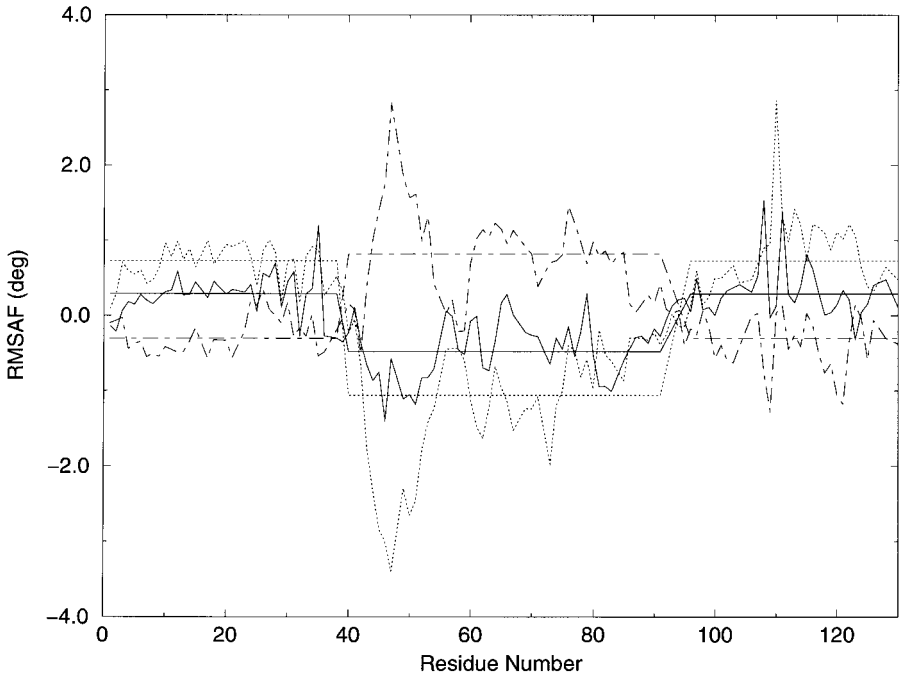

Fig. 1. The x (continuous line), y (long-dash, short-dash line), and z (dotted line) components of the rms angular fluctuation of each residue for the first normal mode of human lysozyme. The horizontal lines show the results after intradomain fluctuations have been removed.

that residues 39 and 40 as well as residues 85, 86, and 90–100 are those mainly involved in the interdomain motion. Residues 90–100 are in the interdomain alpha-helix and 39 and 40 in the interdomain loop. The first normal mode axis satisfies, therefore, the criteria for being a hinge axis. The axis makes an angle of 76° to the line joining the centers of mass and passes it at a distance of 3.6 Å. The motion is 94% closure motion, and the axis could therefore also be regarded as a closure axis.

McCammon et al.[20] guessed with their "rigid rotor model" that the hinge axis passed through the CA atoms of residues 38 and 97 of hen lysozyme and so their demarcation of the two domains almost, but not quite, corresponds with ours. Brooks and Karplus[21] compared the rigid rotor model with the first normal mode of lysozyme and estimated that there are "hinge points" at residues 40, 76, and 94. Gibrat and Go,[22] who fitted a model of hinge bending motion to the results from a normal mode analysis of lysozyme, found that the hinge axis passed through the CA atoms of residues 55 and 76.

At 300 K the rms angular fluctuation (rmsaf) is 2.25° and the rmsf of the translational motion along the axis has the very small value of 0.034 Å. As this is a harmonic analysis, in which the motion is confined within a single minimum, our values for the rmasf and the rmsf of the translation along the axis are likely to be underestimates of the true values. Our hope, however, is that the location of the hinge axis as determined by normal mode analysis is a good approximation to the actual one. Our comparison later with the results from an MD simulation will support or dash this hope.

Second normal mode interdomain screw axis

The relative motion of the two domains as rigid bodies in the second normal mode accounts for 76% of the total mass-weighted msf and it is therefore a mode that also predominantly describes the relative motion of the two domains. The interdomain screw axis passes within 3 Å of the CA atoms of residues 26, 27, 28, 90, 91, and 93. It is shown as the green line in Figure 3. The dotted line in Figure 4 shows which residues are involved in the interdomain motion for the second normal mode. They are largely the same ones as for the first normal mode. Again the axis satisfies our criteria for being a hinge axis. It passes the first normal mode hinge axis at a distance of 1.4 Å and makes an angle of 91.6° with it. If there is little translation along the axes, as is the case for both the first and second normal modes, then an angle of 90° is required by orthogonality. The fact that we find that this is the case gives us confidence that no errors have been made. The axis makes an angle of 23° to the line joining the centers of mass of the two domains and passes it at a distance of 3.1 Å. Decomposing into closure and twisting motions, we find that it is 85% twisting. In contrast to the first normal mode therefore, the second normal mode is more of a twist axis than a closure axis.

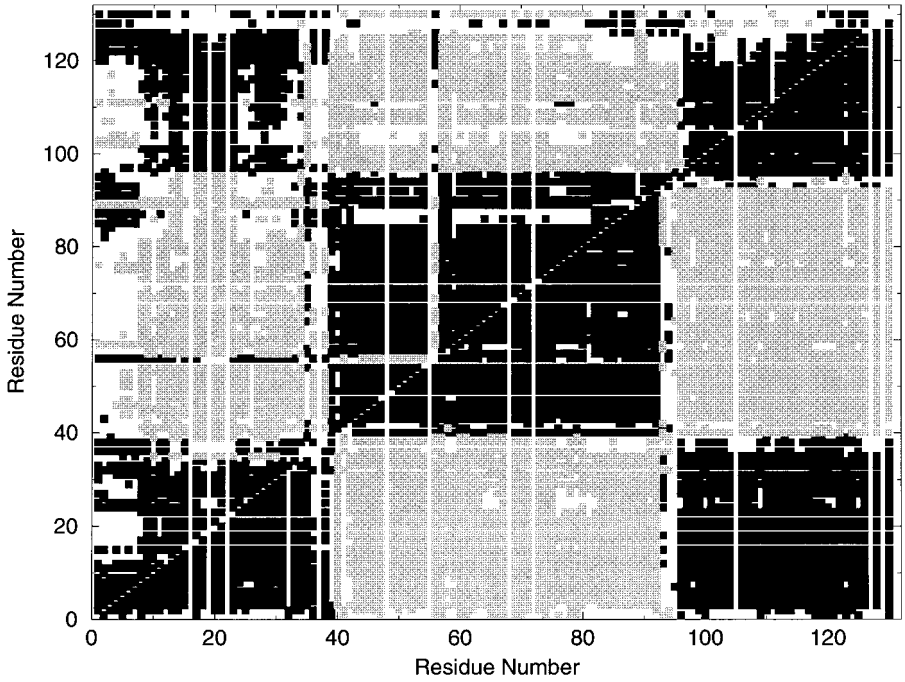

Fig. 2. Plot of the rotation-orientation matrix to determine the dynamical domains from the first (lower triangle) and second (upper triangle) normal modes. A black square means that the corresponding residues whose numbers are given by the x-y coordinates of that point have rotation vectors that are aligned such that the corresponding unit vectors yield an inner product greater than 0.5. For a gray square the value for the inner products is less than −0.5.

At 300 K the rmsaf is 1.1° and the rmsf of the translational motion along the axis is −0.011 Å.

Third and higher normal modes

As mentioned above, the rotation-orientation matrix plot for the third normal mode did not give the simple two-domain demarcation found for the first two normal modes. For this mode the relative motion of the two domains as rigid bodies accounts for only 32% of its total mass-weighted msf. The third normal mode, therefore, represents to a larger extent intradomain fluctuations for the domains as we have defined them, i.e., on the basis of the results from the first and second normal modes. This was more so, the higher the frequency of the mode and for that reason we did not analyze these modes for their interdomain screw axes. The fact that only the first two normal modes describe the relative motion of the two domains tells us something about the nature of the connecting region, that is, it allows only two degrees of freedom out of a possible six for a totally flexible connection, and one for a pure hinge. The obvious restraints on the motion of the interdomain helix are surely instrumental in this.

Principal Component Analysis of Hen Lysozyme MD Simulation

Initially a conventional principal component analysis was performed on the trajectory of hen lysozyme. The results were disappointing but also revealing. Choosing the same demarcation between the domains as was found by the normal mode analysis, the first principal mode represented interdomain motion of the two bodies as rigid bodies to only 17%. This result tells us that during the simulation most of the motion is intradomain fluctuation. However, if we accept that the main motion of lysozyme is its interdomain motion, as the normal mode analysis has indicated, then this result can only be due to the relative shortness of the simulation. The interdomain motion is expected to be a slow overdamped motion; during a short simulation much faster intradomain motions may have considerably larger msfs. Therefore the first principal mode, which has the largest msf, will not represent the slow interdomain motion but a faster intradomain motion. Only in a very long simulation will the interdomain motion become the one with the largest msf and thus the first principal mode. To overcome this limitation we have developed the principal component analysis for interdomain motions introduced in the Methods section. In removing the intradomain fluctuations, we can reveal the nature of the interdomain motions. We chose the same demarcation between the two domains as was determined from the normal mode analysis. Figure 5 shows the result of this analysis. It shows that, as expected, there are only 6 non-zero eigenvalues. What this plot really shows, though, is that the interdomain connection allows only two

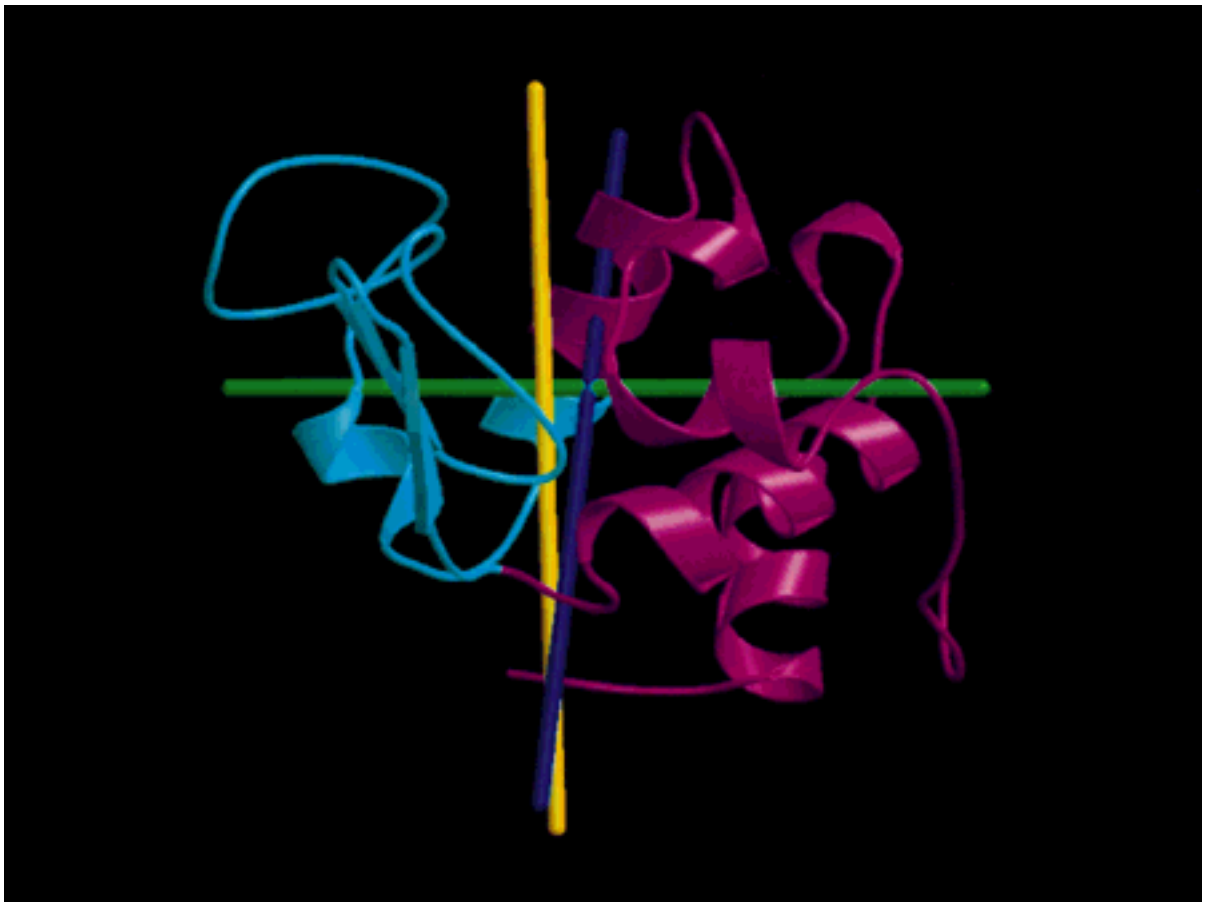

Fig. 3. The first domain of human lysozyme including residues 1–39 and 94–130 is shown in magenta, and the second domain, including residues 40–93, is shown in cyan. The yellow line represents the first normal mode axis, the green line the second normal mode axis, and the blue line the first principal mode axis. Figure created using "Molscript"[35] and "Raster3D."[36]

degrees of freedom out of a possible six. This was what we also found from the normal mode analysis.

First principal mode interdomain screw axis

The blue line in Figure 3 shows the axis determined from the first principal mode. It passes within 3 Å of the CA atoms of residues 2, 39, 56, 96, 98, and 99 of human lysozyme (by renumbering hen to human). It makes an angle of 12° with the first normal mode axis and passes it at a distance of 0.9 Å. The good agreement between the first normal mode axis and the first principal mode axis is a significant result. The location and orientation of the first principal mode axis is purely dependent on the relative motions of the two domains during the simulation. Therefore the correspondence is surely physical, indicating that the axis is determined by features largely independent of the considerable differences between the two methods. The rmsaf has a value of 3.8°, and the rmsf of the translational motion along the axis is −0.128 Å, considerably larger values than for the first normal mode. The axis makes an angle of 85° with the line joining the centers of mass of the two domains, and the resulting rotation is 99% closure motion.

Second principal mode interdomain screw axis

The axis of the second principal mode passes within 3 Å of residues 55, 82, and 83. It makes an angle of 39° with the second normal mode axis and passes it at a distance of 1.5 Å. There is therefore a considerable discrepancy between the results from the two methods for the second modes. The rmsaf is 2.7° and the rmsf of the translational motion along the axis is 0.2 Å. The motion is 55% twisting. Therefore, like the second normal mode axis, it is more of a twist axis than a closure axis.

CONCLUSIONS

The main aim of this paper is to introduce model-free, and therefore unbiased, methods to determine quantities pertaining to domain motions in proteins from normal mode analyses or principal component analyses. For the normal mode analysis the method is based on the assumption that intradomain interactions are large enough in comparison with interdo-

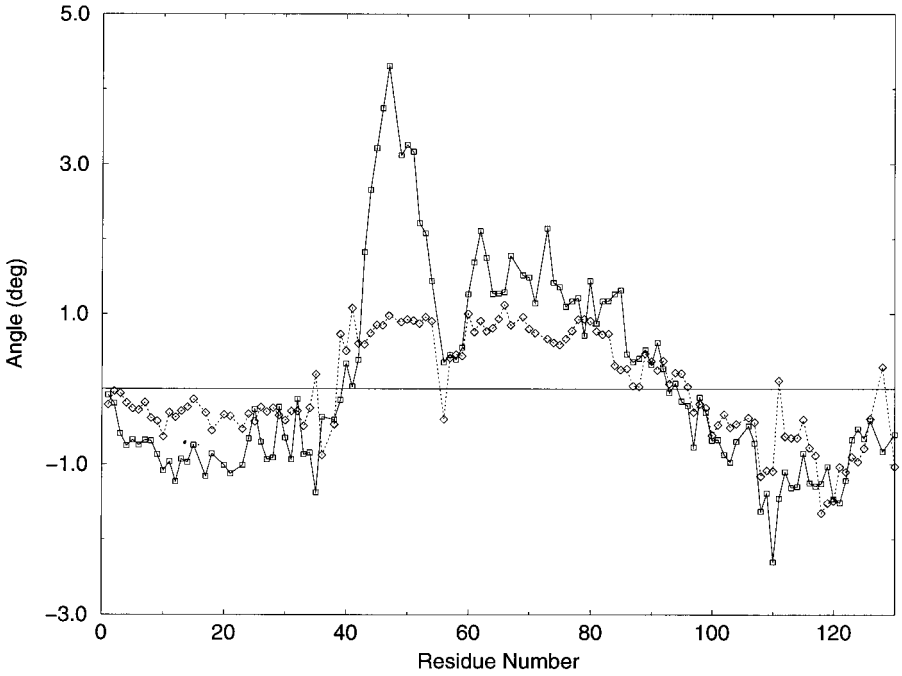

Fig. 4. Plot to show which residues are involved in the interdomain rotation. The continous line with circles shows the result of projecting the component of each residue's rotation vector onto the interdomain rotation vector for the first normal mode. The dotted line with diamonds shows the results for the second normal mode.

main interactions for the latter to be regarded as a perturbation. If the unperturbed system comprises M noninteracting domains, then one has 6(M-1) degenerate eigenvectors with eigenvalue 0 describing interdomain motions in the internal coordinate system. In degenerate perturbation theory the expansion up to the zeroth order delivers modes that are linear combinations of the degenerate eigenvectors of the unperturbed Hessian. We have shown in Appendix A that in this case the modes represent rigid body motions of the domains. We then assume that these modes are a reasonable approximation to the eigenvectors of the real, perturbed system. If this is the case then the eigenvectors from a normal mode analysis of the real system can be analyzed in terms of the properties of rigid body motion. By using the curl we can determine all the quantities relevant to the rigid body motion of a domain as well as the location of any hinge axis, etc. We have demonstrated the success of this method on a normal mode analysis of human lysozyme.

The first step in the process is to identify the demarcation between the two domains by plotting the rotation-orientation matrix. Once the dynamical domains have been identified, the interdomain screw axis can be determined. The method is such that the location of the interdomain screw axis reveals the type of motion occurring. If the interdomain screw axis is located near interdomain connecting regions determined to be involved in the interdomain motion, then the axis can be regarded as a physical axis, and only then is called a hinge axis. If the interdomain screw axis is located away from the interdomain connections, we would automatically know that there is translation, in a direction perpendicular to any rotation axis, involving the interdomain connection. In principle one could determine the degree of this translation. The fact that for the normal mode analysis of lysozyme, the interdomain screw axis is located in between the two domains, passing near interdomain connections shown to be involved in the interdomain motion, proves that lysozyme has a hinge axis. Hinge axes have further been classified into two extreme types, twist and closure axes. We have shown that according to the definition of closure axis and twist axis, the first normal mode motion is 94% closure motion. In addition, we have found that the second normal mode motion is 85% twisting. Higher normal modes do not characterize interdomain motion and therefore one result of the normal mode analysis is that the interdomain motion has two degrees of freedom.

In the case of a conventional principal component analysis it was found that the first principal mode does not characterize interdomain motion but rather intradomain motion. This was attributed to the relatively short time of the simulation, which does not allow the slow, overdamped, interdomain motion to converge. To overcome this, a principal component analysis of the interdomain motion was performed whereby the intradomain motion was removed. This

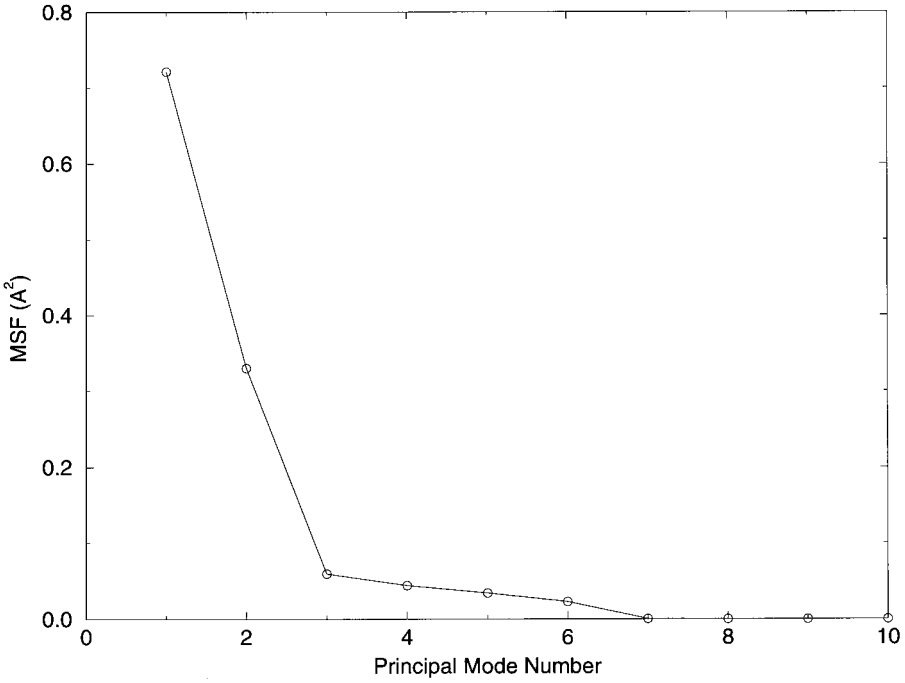

Fig. 5. The mean-square fluctuation of each principal mode from the principal component analysis on the interdomain motion, showing that interdomain motion has basically two degrees of freedom.

provided a first principal mode axis that corresponded very well with the first normal mode axis, with both passing through residues 2, 39, and 56, and both passing near the interdomain helix from 90 to 100.

In concordance with the normal mode analysis, the interdomain motion is dominated by two degrees of freedom. In both normal and principal modes, the first mode represents a closure motion and the second mode predominantly a twisting motion. In the case of the normal mode analysis, the rmsaf of twisting is half that of the closure motion; in the case of the principal component analysis it is about two-thirds. In both cases these results imply that the interdomain interactions in lysozyme, in allowing two degrees of freedom out of a possible six, are not as restrictive as a pure mechanical hinge, which only allows one.

The correspondence between normal mode analysis on lysozyme in vacuum and principal component analysis of an MD simulation of lysozyme in water suggests that the interdomain motion is largely determined by the nature of the connections between the two domains, in particular the interdomain alpha-helix. The fact that in both cases the axes pass near or through these connecting regions supports this, and suggests possible simple rules for the approximate determination of hinge axes. In particular, the interdomain alpha-helix appears to play a crucial role in the interdomain dynamics of lysozyme. Further analyses of this type and their comparison with experimental results should lead to a better understanding of interdomain motions and allow realistic efforts to engineer proteins with desired interdomain dynamic properties.

ACKNOWLEDGMENTS

The authors are grateful to Alan Mark at ETH, Zurich, Switzerland for the simulation results on hen lysozyme. S.H. is a recipient of a BIOTECH fellowship from the European Union.

APPENDIX A

The normal modes of external motion in the mass-weighted coordinate system are given as[33]:

$$\mathbf{\Phi} = \begin{vmatrix} \sqrt{m_1}C_T & 0 & 0 & 0 & \sqrt{m_1}C_y z_1 & -\sqrt{m_1}C_z y_1 \\ 0 & \sqrt{m_1}C_T & 0 & -\sqrt{m_1}C_x z_1 & 0 & \sqrt{m_1}C_z x_1 \\ 0 & 0 & \sqrt{m_1}C_T & \sqrt{m_1}C_x y_1 & -\sqrt{m_1}C_y x_1 & 0 \\ \sqrt{m_2}C_T & 0 & 0 & 0 & \sqrt{m_2}C_y z_2 & -\sqrt{m_2}C_z y_2 \\ 0 & \sqrt{m_2}C_T & 0 & -\sqrt{m_2}C_x z_2 & 0 & \sqrt{m_2}C_z x_2 \\ 0 & 0 & \sqrt{m_2}C_T & \sqrt{m_2}C_x y_2 & -\sqrt{m_2}C_y x_2 & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \end{vmatrix} \quad \text{(A1)}$$

where

$$C_T = \left(\sum_{i=1}^{N} m_i\right)^{-1/2}$$

$$C_x = \left(\sum_{i=1}^{N} m_i(z_i^2 + y_i^2)\right)^{-1/2}$$

$$C_y = \left(\sum_{i=1}^{N} m_i(z_i^2 + x_i^2)\right)^{-1/2}$$

$$C_z = \left(\sum_{i=1}^{N} m_i(y_i^2 + x_i^2)\right)^{-1/2}$$

The matrix $\mathbf{\Phi}$ satisfies the orthonormality condition $\mathbf{\Phi}^t\mathbf{\Phi} = \mathbf{I}$, where t denotes the transpose. The column vectors in $\mathbf{\Phi}$ correspond to the eigenvectors of translation in the x, y, and z directions and rotation about the x, y, and z axes, which are the principal axes of the part of the protein corresponding to this block. Let us term the eigenvectors of these external motions for the *kth* part of the protein corresponding to block k $\mathbf{T}_x^k$, $\mathbf{T}_y^k$, $\mathbf{T}_z^k$, $\mathbf{R}_x^k$, $\mathbf{R}_y^k$, and $\mathbf{R}_z^k$, respectively. Any linear combination of these vectors will give an eigenvalue of 0 when acted on by block k in $\mathbf{H_o}$. Let $\mathbf{v}_k$ be any such combination for block k in $\mathbf{H_o}$:

$$\mathbf{v}_k = t_x^k\mathbf{T}_x^k + t_y^k\mathbf{T}_y^k + t_z^k\mathbf{T}_z^k + r_x^k\mathbf{R}_x^k + r_y^k\mathbf{R}_y^k + r_z^k\mathbf{R}_z^k. \quad \text{(A2)}$$

If $\mathbf{V}$ is given by:

$$\mathbf{V} = \sum_k \lambda_k\mathbf{v}_k \quad \text{(A3)}$$

then this vector will be an eigenvector of $\mathbf{H_o}$ with an eigenvalue of 0. We are not interested in the overall external motion of the protein, and so there are six conditions that constrain the modes to describe the internal motion of the protein. These are $\mathbf{V.T_x} = 0$, $\mathbf{V.T_y} = 0$, $\mathbf{V.T_z} = 0$, $\mathbf{V.R_x} = 0$, $\mathbf{V.R_y} = 0$, and $\mathbf{V.R_z} = 0$, where $\mathbf{T_x}$, $\mathbf{T_y}$, $\mathbf{T_z}$, $\mathbf{R_x}$, $\mathbf{R_y}$, and $\mathbf{R_z}$ represent the external modes of motion for the whole molecule. We also require that $\mathbf{V.V} = 1$.

Using Eqs. (A1) and (A2) we can now write the component of $\mathbf{V}$ corresponding to the *ith* atom, which is in part k.

$$\mathbf{V}_i = \lambda_k\sqrt{m_i}[C_T^k(t_x^k\mathbf{i} + t_y^k\mathbf{j} + t_z^k\mathbf{k}) + (r_x^kC_x^k\mathbf{i} + r_y^kC_y^k\mathbf{j} + r_z^kC_z^k\mathbf{k}) \times (x_i\mathbf{i} + y_i\mathbf{j} + z_i\mathbf{k})] \quad \text{(A4)}$$

where x, y, z, and $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ correspond to the coordinate system at the center of mass and oriented along the principal axes of part k of the protein. Since $\mathbf{V}_i$ is a mass-weighted coordinate, the actual atomic displacement corresponding to this eigenvector is:

$$\mathbf{\Delta d_i} = \frac{\mathbf{V}_i}{\sqrt{m_i}} = \mathbf{\Delta}\tau^k + \mathbf{\Delta}\psi^k \times \mathbf{r}_i \quad \text{(A5)}$$

which shows clearly that the *kth* part of the eigenvector corresponds to a rigid body motion of that part, combining a translation, and a rotation about a single axis through the center of mass.

In degenerate perturbation theory, particular linear combinations of the degenerate eigenvectors of the unperturbed system will approximate the eigenvectors of the perturbed system where the degeneracy is likely to have been removed. For example, there will be one combination of the coefficients t_x^k, t_y^k, t_z^k, r_x^k, r_y^k, r_z^k, and λ_k that determines a $\mathbf{V}$ that gives the mode that best approximates the lowest frequency mode of the perturbed system. This combination will depend on the energetics of the perturbation, which in our case is an interaction between the parts. These particular linear combinations of the unperturbed system eigenvectors are provided by the method of degenerate perturbation theory.[34] These modes are the *zeroth*-order approximation to the eigenvectors of the perturbed system. Providing the perturbation is sufficiently small, the low-frequency eigenvectors of the real system Hessian $\mathbf{H}$ will be approximately of the form of $\mathbf{V}$, as given in Eq. (A5).

APPENDIX B
Practical Calculation of the Curl

In cartesian coordinates the curl is given as:

$$\nabla \times \mathbf{V} = \left(\frac{\partial V_z}{\partial y} - \frac{\partial V_y}{\partial z}\right)\mathbf{i} + \left(\frac{\partial V_x}{\partial z} - \frac{\partial V_z}{\partial x}\right)\mathbf{j} + \left(\frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y}\right)\mathbf{k} \quad \text{(B1)}$$

where $\mathbf{V}$ is a continuous vector field, which is our case is equal to $\mathbf{\Delta d}(\mathbf{r}^s)$. $\mathbf{V}$ is sampled by the displacement $\mathbf{\Delta d}_i$ at each atom i. To calculate the derivatives necessary to determine the curl, we select the N-CA-C-CB (in PDB data bank nomenclature) tetrahedra formed by the backbone atoms on each residue (apart from glycines of course). This group of atoms forms a rigid element in the low-frequency modes and is therefore an appropriate group with which to calculate the curl in such modes. Considering the x-component of $\mathbf{V}$, and assigning position vector $\mathbf{r}_1$ to the CA atom, and $\mathbf{r}_2$, $\mathbf{r}_3$, and $\mathbf{r}_4$ to the other atoms, we can write:

$$\Delta_2V_x = V_x(\mathbf{r}_2) - V_x(\mathbf{r}_1) = \Delta x_2\frac{\partial V_x}{\partial x} + \Delta y_2\frac{\partial V_x}{\partial y} + \Delta z_2\frac{\partial V_x}{\partial z} \quad \text{(B2)}$$

where $\Delta x_2 = (x_2 - x_1)$, and $\Delta_2 V_x$ is the difference in the displacements in the x-direction at $\mathbf{r}_1$ and $\mathbf{r}_2$. Similar expressions hold for the change in V_x between $\mathbf{r}_3$ and $\mathbf{r}_1$, and between $\mathbf{r}_4$ and $\mathbf{r}_1$. Writing these in matrix form, one has:

$$\begin{pmatrix} \Delta x_2 & \Delta y_2 & \Delta z_2 \\ \Delta x_3 & \Delta y_3 & \Delta z_3 \\ \Delta x_4 & \Delta y_4 & \Delta z_4 \end{pmatrix} \begin{pmatrix} \dfrac{\partial V_x}{\partial x} \\ \dfrac{\partial V_x}{\partial y} \\ \dfrac{\partial V_x}{\partial z} \end{pmatrix} = \begin{pmatrix} \Delta_2 V_x \\ \Delta_3 V_x \\ \Delta_4 V_x \end{pmatrix}. \tag{B3}$$

If we write this as:

$$\boldsymbol{\Delta} \cdot \boldsymbol{\nabla} \mathbf{V}_x = \boldsymbol{\Delta} \mathbf{V}_x \tag{B4}$$

then the desired derivatives of the x component are given by:

$$\boldsymbol{\nabla} \mathbf{V}_x = \boldsymbol{\Delta}^{-1} \cdot \boldsymbol{\Delta} \mathbf{V}_x. \tag{B5}$$

Similar expressions for the y and z components of $\mathbf{V}$ give all the derivatives needed to calculate the curl.

REFERENCES

1. Nishikawa, T., Go, N. Normal modes of vibration in bovine pancreatic inhibitor and its mechanical properties. Proteins 2:308–329, 1987.
2. Schulz, G.E. Domain motions in proteins. Curr. Opin. Struct. Biol. 1:883–888, 1991.
3. Gerstein, M., Lesk, A.M., Chothia, C. Structural mechanisms for domain movements in proteins. Biochemistry 33:6739–6749, 1994.
4. Faber, H.R., Matthews, B.W. A mutant T4 lysozyme displays five different crystal conformations. Nature 348:263–266, 1990.
5. Go, N., Noguti, T., Nishikawa, T. Dynamics of a small globular protein in terms of low-frequency vibrational modes. Proc. Natl. Acad. Sci. U.S.A. 80:3696–3700, 1983.
6. Brooks, B., Karplus, M. Harmonic dynamics of proteins: Normal modes and fluctuations in bovine pancreatic trypsin inhibitor. Proc. Natl. Acad. Sci. U.S.A. 80:6571–6575, 1983.
7. Levitt, M., Sander, C., Stern, P.S. The normal modes of a protein: Native bovine pancreatic trypsin inhibitor. Int. J. Quant. Chem. 10:181–199, 1983.
8. Case, D.A. Normal mode analysis of protein dynamics. Curr. Opin. Struct. Biol. 4:285–290, 1994.
9. Hayward, S., Go, N. Collective variable description of native protein dynamics. Annu. Rev. Phys. Chem. 46:223–250, 1995.
10. Kidera, A., Go, N. Normal mode refinement: Crystallographic refinement of protein dynamic structure. I. Theory and test by simulated diffraction data. J. Mol. Biol. 225:457–475, 1992.
11. Kidera, A., Inaka, K., Matsushima, M., Go, N. Normal mode refinement: Crystallographic refinement of protein dynamic structure. II. Application to human lysozyme. J. Mol. Biol. 225:477–486, 1992.
12. Hayward, S., Kitao, A., Go, N. Harmonic and anharmonic aspects in the dynamics of BPTI: A normal mode analysis and principal component analysis. Protein Sci. 3:936–943, 1994.
13. Hayward, S., Kitao, A., Go, N. Harmonicity and anharmonicity in protein dynamics: A normal mode analysis and principal component analysis. Proteins 23:177–186, 1995.
14. Janezic, D., Venable, R.M., Brooks, B.R. Harmonic analysis of large systems. III. Comparison with molecular dynamics. J. Comp. Chem. 16:1554–1566, 1995.
15. Levy, R.M., Srinivasan, A.R., Olson, W.K., McCammon, J.A. Quasi-harmonic method for studying very low frequency modes in proteins. Biopolymers 23:1099–1112, 1984.
16. Kitao, A., Go, N. Conformational dynamics of polypeptides and proteins in the dihredral angle space and in the cartesian coordinate space: Normal mode analysis of decaalanine. J. Comp. Chem. 12:359–368, 1991.
17. Garcia, A.E. Large-amplitude nonlinear motions in proteins. Phys. Rev. Lett. 68:2696–2699, 1992.
18. Hayward, S., Kitao, A., Hirata, F., Go, N. Effect of solvent on collective motions in globular proteins. J. Mol. Biol. 234:1207–1217, 1993.
19. Amadei, A., Linssen, A.B.M., Berendsen, H.J.C. Essential dynamics of proteins. Proteins 17:412–425, 1993.
20. McCammon, J.A., Gelin, B.R., Karplus, M., Wolynes, P.G. Hinge bending mode in lysozyme. Nature 262:325–326, 1976.
21. Brooks, B., Karplus, M. Normal modes for specific motions of macromolecules: Application to the hinge bending mode of lysozyme. Proc. Natl. Acad. Sci. U.S.A. 82:4995–4999, 1985.
22. 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 8:258–279, 1990.
23. Horiuchi, T., Go, N. Projection of Monte Carlo and molecular dynamics trajectories onto the normal mode axes: Human lysozyme. Proteins 10:106–116, 1991.
24. Hayward, S., Berendsen, H.J.C. Model free method to analyse domain motions in proteins from X-ray conformers. (submitted).
25. Kreyszig, E. "Differential Geometry." Toronto: University of Toronto Press, 1959.
26. Goldstein, H. "Classical Mechanics." 2nd edit. L, Reading, Mass., USA: Addison Wesley World Student Series, 1980.
27. Gerstein, M., Anderson, B.F., Norris, G.E., Baker, E.N., Lesk, A.M., Chothia, C. Domain closure in lactoferrin. J. Mol. Biol. 234:357–372, 1993.
28. Smith, L.J., Mark, A.E., Dobson, C.M., van Gunsteren, W.F. Comparison of MD simulations and NMR experiments from hen lysozyme. Analysis of local fluctuations, coorperative motions and global changes. Biochemistry 34:10918–10931, 1995.
29. van Gunsteren, W.F., Berendsen, H.J.C. Gromos-87 manual. BIOMOS BV, Nijenborgh 4, 9747 AG Groningen, The Netherlands, 1987.
30. Morikami, K., Nakai, T., Kidera, A., Saito, M. PRESTO(PRotein Engineering SimulaTOr): A vectorized molecular mechanics program for biopolymers. Comput. Chem. 16:243–248, 1992.
31. Weiner, S.J., Kollman, P.A., Nguyen, D.T., Case, D.A. An all atom force field for simulations of proteins and nucleic acids. J. Comp. Chem. 7:230–252, 1986.
32. Nishikawa, K., Ooi, T., Isogai, Y., Saito, N. Tertiary structure of proteins. I. Representation and computation of the conformations. J. Phys. Soc. (Jpn.) 32:1331–1337, 1972.
33. Bright Wilson, Jr., E., Decius, J.C., Cross, P.C. "Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra." New York: Dover Publications, 1980.
34. Mandl, F. "Quantum Mechanics." 2nd edit. London: Butterworths, 1957.
35. Kraulis, P.J. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallogr. 24:946–950, 1991.
36. Merrit, E.A., Murphy, M.E.P. Raster3D Version 2.0. A program for photorealistic molecular graphics. Acta Crystallogr. D50:869–873, 1994.