Journal of Computational Physics **151,** 169–189 (1999)
Article ID jcph.1999.6232, available online at http://www.idealibrary.com on IDEAL®

Low Frequency Motion in Proteins

Comparison of Normal Mode and Molecular Dynamics of Streptomyces Griseus Protease A

Pnina Dauber-Osguthorpe,* David J. Osguthorpe,† Peter S. Stern,‡ and John Moult§

*Department of Biology and Biochemistry and †School of Chemistry, University of Bath, Bath, Avon, BA2 7AY, England; ‡Chemical Physics Department, Weizmann Institute of Science, Rehovot, Israel; and §Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, 9600 Gudelsky Drive, Rockville, Maryland 20850

E-mail: *pnina@mgu.bath.ac.uk, †djosg@gmgu.bath.ac.uk, ‡Peter.Stern@weizmann.ac.il, and §jmoult@carb.nist.gov

Received July 20, 1998; revised February 8, 1999

The motion of a chymotrypsin-like serine protease, *SGPA*, has been studied by torsion space normal mode analysis and by Cartesian space molecular dynamics, and the results have been compared. The molecular dynamics trajectory was analyzed using digital signal processing techniques to provide a set of characteristic modes that can be compared directly with the normal modes. The results were also compared with the motion implied by the crystallographic temperature factors. We find that in spite of the radically different approximations used in the two methods, agreement between the resulting motions and with the experimental data is surprisingly high. We conclude that this agreement probably reflects an underlying robustness in the motion, dictated primarily by van der Waals packing. In contrast to other proteins, there are no large amplitude inter-domain motions. Rather, the low frequency, high amplitude motions are concentrated in three surface hairpin loops. The movement of one these loops, the specificity loop, appears to facilitate substrate binding.

Key Words: molecular dynamics; normal modes; protein structure; protein motion.

INTRODUCTION

Conformational flexibility is essential to many biological events. A variety of experimental methods give access to partial information on protein motion: X-ray crystallography provides accurate mean atomic positions, but only approximate values for the isotropic amplitudes of motion about those positions [1]. Limited information about protein flexibility

0021-9991/99 $30.00

and movement is also available from diffuse x-ray scattering [2]. Some insight into the time scale of protein motion has been obtained from Raman spectroscopy [3], fluorescence studies [4, 5], inelastic neutron scattering [6], and in particular, from NMR experiments [7]. However, these experimental techniques provide very few data on the spatial displacements involved. Computational methods offer a means of exploring the details of the possible motions. The two main approaches to investigating molecular motion are normal mode dynamics and molecular dynamics simulations.

Normal mode dynamics (NM) has long been used as a tool in interpreting vibrational spectra of small molecules [8, 9] and homobiopolymers [10]. In recent years it has been extended to the study of large systems such as proteins [11, 13]. The motion is modeled as a superposition of a set of independent harmonic oscillations about the equilibrium atomic positions. For protein work, the required forces are represented by the same type of approximate empirical potential as that used in molecular dynamics. The method provides an informative visual model which characterizes the available types of motion in terms of frequency and of atomic amplitudes and directions of motion. Of particular interest is the ability to study slow collective motions of large biological molecules. It also allows for the efficient calculation of time-averaged properties associated with positional fluctuations. However, it is not clear *a priori* that the underlying assumptions are justified for large molecules. First, the motion is assumed to involve small oscillations about one minimum energy conformation and to be harmonic (quadratic) in nature. Thus, its application in systems such as proteins, with many local minima and large fluctuations, may not be valid. Second, it has not been possible to apply this approach to systems with inherently disordered components, so that solvent effects cannot be taken into account.

The other major technique, molecular dynamics (MD) [14, 15], involves solving Newton's equations of motion to yield a trajectory of atomic positions. In principle, these represent a realistic description of molecular motion, including small and large structural fluctuations and conformational transitions. Moreover, this technique can take solvent effects into account explicitly. Limitations are imposed by the approximate nature of the force fields and by the relatively short time scale (of the order of a nanosecond) computationally accessible. Although the laws of motion are very simple, the resultant trajectories are very complicated and interpreting the complex motion is not trivial. In particular, it is difficult to investigate long range collective motions. This problem has been addressed recently in a number of studies, using different approaches. One method for identifying collective motions in proteins from MD simulations makes use of the equal-time covariances and cross-correlations of atomic fluctuations, and has revealed that regions of secondary structure move in a correlated manner [16]. Diagonalization of the equal-time correlation and projection of the MD equations of motion onto the resulting eigenvectors is another approach to analyzing the important motions in the simulation [17]. A hybrid method projects the atomic trajectories in MD simulations onto previously calculated axes of normal modes [18], or onto axes obtained by principal component analysis [19]. Recently, a method using either NM analysis or a form of principal component analysis of a MD simulation was used to characterize the hinge-bending motion between two dynamical domains in lysozyme [20].

We have used digital signal processing techniques to characterize the motion in MD simulations. Fourier transforming all the atomic trajectories yields the overall frequency distribution. We then choose the frequency ranges corresponding to motions of interest and eliminate the rest. In this way, as was demonstrated for small molecules [21, 22] and for a protein [23], it is possible to remove high frequency bond stretches and valence angle

bending and focus on the low frequency conformational motion. Moreover, we are able to extend this approach to extract the vectors defining the characteristic motion for each frequency of interest in a MD simulation [22, 24]. These vectors are analogous to those obtained from NM and provide a pictorial description of the motion as well as a means for comparing the results of the two methods. The technique has been used for the study of motion of small molecules, where distinct, well-resolved modes can be extracted from the MD trajectory [22, 24].

We now apply this method to a protein molecule, *Streptomyces Griseus Protease A* (*SGPA*), and compare the results to those obtained from NM analysis. We compare the amplitudes of the motion about the mean positions for the two methods with each other and with x-ray data and examine the correlation of the directions of atomic fluctuations from the two simulations. We address the issues of how well the three descriptions of motion agree and what the implications of the results are for the nature of the factors controlling motion in protein molecules. For the two types of simulation, the motions have been characterized in additional detail. This includes examining the nature of the motion as a function of the frequency, revealing the dependence of overall correlated motion on structural properties of the protein, and comparing specific modes of motion obtained by NM and MD.

SGPA is a small monomeric protein with a single polypeptide chain 181 amino acid residues in length. It is a member of the chymotrypsin class of serine proteases [25] with a fold made up of two similar anti-parallel β-barrel domains. The structure has been solved at a resolution of 1.5 Å, and carefully refined to an R factor of 12.5% [25]. MD analysis is based on a 216 ps MD simulation of the motion of the protein in its crystal environment, explicitly including the solvating water molecules and ions, and the neighboring protein molecules [26].

METHODS

Minimization and Normal Mode Dynamics

The normal mode analysis was carried out on a single protein molecule, *in vacuo*. The initial equilibrium conformation of the protein was obtained by minimizing the energy of the x-ray structure, first with respect to the 7413 Cartesian coordinates (including all hydrogen atoms), using 1567 steps of conjugate gradient minimization with smoothing and an 8 Å cutoff [27] and then with respect to the 574 dihedral angles using a variable metric minimization routine. The minimizations were carried out with the program ENCAD (by Michael Levitt). Details of the force field are given in Ref. [28]. The force field was modified to use electrostatic interactions based on the partial charges of Hagler, Huler, and Lifson [29], rather than a specific hydrogen bond energy term. The 8 Å cutoff was used to maintain compatibility with earlier NM work. The minimization in Cartesian coordinates was stopped when the energy change was less than 10^{-7} kcal/mol per step. The rms shift in coordinates at this point was less than 10^{-5} Å per step. The total rms change in coordinates from the x-ray coordinates was 0.48 Å. The rms value of the final derivatives after minimization with respect to dihedral angles was 4.0×10^{-5} kcal/mol-rad (the largest derivative was 1.5×10^{-4}). The second derivative matrix of the potential energy with respect to the dihedral angles was calculated numerically from the analytically determined values of the first derivatives and was then diagonalized to yield the normal modes, $\mathbf{L}_k$ (eigenvectors) and frequencies, ν_k (eigenvalues). (For details, see Refs. [11, 30].) The complete dynamic behavior of the

system is described, in torsion space, by

$$\theta_j(t) = \theta_j^0 + \sum_k^{N_\theta} L_{jk}\alpha_k \cos(2\pi\nu_k t + \varepsilon_k), \tag{1}$$

where $\theta_j(t)$ is the jth torsion angle at time t, θ_j^0 is the equilibrium value of torsion j, while α_k and ε_k are the amplitude and phase of mode k, respectively. The amplitudes α_k are proportional to the temperature, T,

$$\alpha_k = (2k_B T)^{1/2}/2\pi\nu_k, \tag{2}$$

where k_B is Boltzmann's constant. The corresponding dynamic behavior in Cartesian space is given by

$$q_i(t) = \sum_k l_{ik}\alpha_k \cos(2\pi\nu_k t + \varepsilon_k), \tag{3}$$

where the ith component of the kth normalized Cartesian mode amplitude, l_{ik}, is given by

$$l_{ik} = \sum_j (\partial x_i/\partial\theta_j) L_{jk} \tag{4}$$

and q_i is the mass-weighted Cartesian coordinate i, $q_i = m_i^{1/2}(x_i - x_i^0)$. In practice, the Cartesian displacements for mode k were calculated numerically from the change in Cartesian coordinates obtained by perturbing the torsion angles, j, by L_{jk}.

Molecular Dynamics

Newton's equations of motion were solved numerically [14] using a Verlet algorithm [31]. This simulation was performed in Cartesian space. The MD trajectory used for the analysis was a continuation of a simulation previously reported in [26]. Full details of the method can be found there. The motion of the contents of two asymmetric units of the *SGPA* crystal were simulated, including two protein molecules with all hydrogen atoms, 1429 water molecules, 26 dihydrogen phosphate ions, and 16 sodium ions. Protein heavy atom starting positions were taken from the 1.5 Å resolution x-ray structure [25]. The simulation was performed with a forerunner of the program DISCOVER (Molecular Simulations Inc., San Diego, CA) which uses a valence force field [32]. The dynamics was run with crystal symmetry boundary conditions, coupling to a temperature bath [33] and a 1 fs time step. Interactions closer than 15 Å between group centers were included, with a switching function smoothly decreasing the potential from its full value to zero between 11 and 15 Å. After initial Monte Carlo and molecular dynamics relaxation of the solvent positions, a total of 216 ps of dynamics was run. Coordinates and energies were saved energy 10 fs. This rate of sampling corresponds to a maximum frequency of 1670 cm^{-1} [21, 34]. Only a few bond stretching vibrations occur above this frequency. Since we are interested in the low frequency heavy atom motions, this rate of sampling is adequate. The last 16,384 history points (approximately 164 ps) of the simulation were used for the analysis. This corresponds to a frequency resolution of 0.2 cm^{-1}. The analysis was carried out using the program FOCUS [24].

Derivation of the frequency distribution from a MD trajectory. The frequency distribution function, $g(\nu_a)$, gives the density of motions at frequency ν_a. This function is obtained from the trajectories of the atomic coordinates [21]

$$g(\nu_a) = (1/k_B T)\nu_a^2 \sum_i H_i^2(\nu_a) \tag{5}$$

$$H_i(\nu_a) = 1/N \sum_{m=1}^{N} q_i(t_m)\, e^{-j2\pi\nu_a t_m/N}, \tag{6}$$

where $q_i(t_m)$ is the value of the mass weighted displacement coordinate i at time step m during the MD trajectory, and $H_i(\nu_a)$ is the value of the Fourier transform of q_i at frequency ν_a. The frequency distribution can be calculated for the whole system or for a subset of atoms of interest.

Extracting modes. The normal modes were defined in Eq. (1) in terms of characteristic frequencies and associated amplitudes of motion. An analogous description of the motion as a function of frequency can be obtained from the MD trajectory using Fourier transform techniques. The Fourier transform of a coordinate trajectory (Eq. (6)) gives the amplitude of oscillation of this coordinate for each frequency. A sample mode, a, is defined by the set of normalized amplitudes, l_{ia}, of all coordinates, i, at a specific frequency, ν_a. The absolute value of l_{ia} is given by

$$|l_{ia}| = |H_i(\nu_a)| \tag{7}$$

and the relative direction is determined by

$$\text{sign}[l_{ia}] = \text{sign}[\text{real}[H_i(\nu_a)]] \text{ for real}[H_i(\nu_a)] > \text{imag}[H_i(\nu_a)] \tag{8}$$

$$\text{sign}[l_{ia}] = \text{sign}[\text{imag}[H_i(\nu_a)]] \text{ for real}[H_i(\nu_a)] < \text{imag}[H_i(\nu_a)], \tag{9}$$

where $H_i(\nu_a)$ is defined in Eq. (6). For further details see Ref. [35].

It should be noted that the sample modes do not necessarily represent independent motions. In principle, each independent mode should correspond to a peak in the frequency distribution. Fluctuations in conformation during the molecular dynamics, and the discrete and finite nature of the trajectories, cause line broadening in the frequency distribution. In addition, for a large molecular such as a protein, there are a large number of closely spaced frequencies. These factors result in an unresolved, but still useful, frequency distribution.

Criteria for the Similarity of Modes

The similarity between a normal mode k and a MD sample mode a may be measured by the dot product, S_{ka}, between their respective Cartesian displacement vectors,

$$S_{ka} = \mathbf{l}_k \cdot \mathbf{l}_a, \tag{10}$$

where $\mathbf{l}_k$ is the kth NM, and $\mathbf{l}_a$ is the ath extracted sample mode [22]. Such dot products have been widely used for comparison of normal modes [36] and modes obtained by principal component analysis of MD simulations [37, 38]. The significance of the calculated dot products was assessed by comparing them with a set of random modes [37]. The random

modes were generated by choosing three random numbers between 0 and 1 to provide the x, y, and z components of each atom. The dot product of random vectors is expected to be of the order of $1/\sqrt{N}$, where N is the size of the vector.

The overall similarity of motion in selected regions of a molecule may be represented by the fraction F of the MD sample modes that have dot products greater than some threshold value with at least one normal mode. The threshold dot product value was taken to be 0.4. Since we are interested primarily in the correlation of the large amplitude motions, it is convenient to define a modified criterion, F_{amp}, which is weighted by the mode amplitudes. For this purpose, we use the total amplitude, $\sum_{ia'} l_{ia'}$, of the subset of atoms i of interest for a subset of MD sample modes a' in the frequency range under consideration. Each of the modes included in the a' set satisfies the following criteria: (a) the amplitude, $l_{ia'}$, of at least one of the atoms in the fragment is greater than 0.4 of the largest atomic amplitude found in mode a'; (b) there is a dot product $S_{k'a'} > 0.4$, with at least one of the normal modes k', where the amplitude, $l_{ik'}$, of at least one of the atoms in the fragment is greater than 0.4 of the largest atomic amplitude found in mode k'. The fraction of the total amplitude accounted for by the amplitudes of sample modes similar to normal modes is thus

$$F_{\text{amp}} = \sum_{ia'} l_{ia'} \Big/ \sum_{ia} l_{ia}, \tag{11}$$

where $\sum_{ia} l_{ia}$ is the total amplitude of all a MD sample modes under consideration.

Removal of rigid body motion. In order to facilitate comparison of the NM and MD motion, rigid body motion of the protein during the MD trajectory was removed. Each transient structure along the analyzed trajectory was transformed to produce the best rigid body fit to the structure of the first time step in the analysis.

Evaluation of the non-oscillatory component of the motion. In the molecular dynamics simulation the atomic fluctuations are due to rigid body motion and conformational transitions between conformational substates as well as oscillatory motions. Examination of the characteristics of the Fourier transform of a few of the atoms which undergo significant conformational transitions reveals a large component at the first frequency sample (0.2 cm^{-1}) which decreases rapidly with frequency. This is due to (a) "edge effects," i.e., the first and last values in the atomic trajectory are significantly different; (b) "discontinuities," i.e., relatively sudden changes in value during the trajectories. Rigid body rotation and translation mainly contribute to the first effect, while conformational transitions are usually (but not always) of the second type. The approximate "non-oscillatory" mode, $\mathbf{x}^{\text{no}}$ is defined by

$$\mathbf{x}^{\text{no}} = \mathbf{x}^{\text{fin}} - \mathbf{x}^{\text{init}}, \tag{12}$$

where $\mathbf{x}^{\text{init}}$ and $\mathbf{x}^{\text{fin}}$ are the coordinates at the beginning and end of the analyzed trajectory, respectively.

Calculation of Root Mean Square Atomic Fluctuations

The observed and calculated atomic fluctuations were derived as follows:

1. Experimental rms fluctuations about the mean atomic position of atom i are derived from the corresponding crystallographic isotropic temperature factor, B_i, by [39]

$$\left\langle u_i^2 \right\rangle_{\text{exp}}^{1/2} = \left(3B_i/8\pi^2\right)^{1/2}. \tag{13}$$

2. The rms amplitude of motion of atom i about its mean position due to all normal modes is [11]

$$\langle u_i^2 \rangle_{\mathrm{nm}}^{1/2} = \left(1/2 \sum_k l_{ik}^2 \alpha_k^2 \right)^{1/2}. \tag{14}$$

3. The rms amplitude of motion about the mean position of atom i from the molecular dynamics trajectory is determined by

$$\langle u_i^2 \rangle_{\mathrm{md}}^{1/2} = \left[\sum_{m=1}^{N} \Delta \mathbf{x}_i(t_m)^2 / N \right]^{1/2}, \tag{15}$$

where $\Delta \mathbf{x}$ are the atomic fluctuation vectors, $\Delta \mathbf{x}_i(t_m) = \mathbf{x}_i(t_m) - \bar{\mathbf{x}}_i$ and $\mathbf{x}_i(t_m)$ is the coordinate vector of atom i at time step m, $\bar{\mathbf{x}}_i$ is the mean position of atom i during the sample period, and N is the number of samples.

Equal-Time Correlations

The correlation in the motions of atoms i and j in the MD simulation is given by [16]

$$C(i, j) = \langle \Delta \mathbf{x}_i(t_m) \cdot \Delta \mathbf{x}_j(t_m) \rangle / \langle \Delta \mathbf{x}_i(t_m)^2 \rangle^{1/2} \langle \Delta \mathbf{x}_j(t_m)^2 \rangle^{1/2}. \tag{16}$$

The corresponding correlation in the NM calculation is given by

$$C(i, j) = \sum_k l'_{ik} \cdot l'_{jk} \Big/ \left(\sum_k l'^2_{ik} \sum_k l'^2_{jk} \right)^{1/2}, \tag{17}$$

where $l'_{ik} = l_{ik}/\nu_k$.

RESULTS AND DISCUSSION

Root Mean Square Atomic Fluctuations

The characteristics of the atomic motion as obtained from experiment, NM dynamics, and MD simulations are shown in Fig. 1a and Table I. The average rms atomic fluctuations for the C_α atoms obtained from x-ray and from the original molecular dynamics are similar. The average rms atomic fluctuations calculated by NM dynamics are significantly lower. It is commonly observed that average atomic fluctuations from molecular mechanics simulations such as NM dynamics [11, 40–43] and MD [23, 44, 45] differ in scale from the corresponding experimental values. (In most cases the calculated values are lower.) This is partly due to the fact that the three sources reflect different types of motion. Normal mode dynamics represents only the harmonic component of intra-molecular motion, and does not include the effects of neighboring protein molecules, solvent, or ions. Molecular dynamics simulations include some conformational transitions between different local minima, and in periodic boundary systems and translational and rotational rigid body motion. Since NM calculations of large systems are, for simplicity, carried out in torsion space, while MD simulations include all internal degrees of freedom, some additional flexibility is to be expected in the MD simulations. Experimental values are based on an isotropic harmonic model of

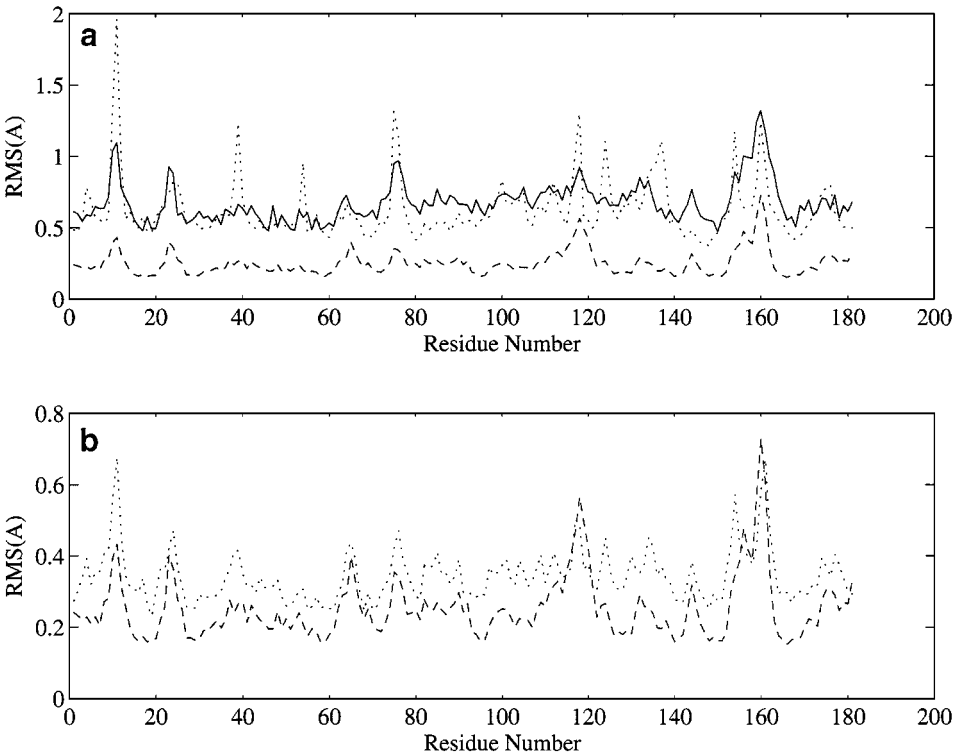


FIG. 1. Root mean square fluctuations in C_α as a function of residue number. Results from experiment and NM and MD simulations are plotted in solid, dashed, and dotted lines, respectively. The atomic fluctuations in the original MD simulation are shown in (a), while the fluctuations after removing rigid body motion and filtering motions with frequency less then 1.6 cm^{-1} are shown in (b). Although the average amplitudes of the motions differ, there is high overall correlations as a function of position in the protein (see Table I).

the individual atomic motions, but indirectly include some anharmonic motion effects, as well as contribution from alternative conformations not sampled on the MD time scale and imperfections in the crystal lattice.

Surprisingly, in spite of the large difference in amplitude of the NM data, the overall behavior as a function of position along the chain is very similar. Inspection of the curves

TABLE I
Comparison of C_α RMS Fluctuations

	Average RMS amplitude (Å)	Correlation coefficients			
		NM	MD[a]	MD[b]	MD[c]
Exp.	0.67	0.82	0.60	0.64	0.79
NM	0.25		0.53	0.62	0.76
MD[a]	0.63			0.95	0.76
MD[b]	0.50				0.86
MD[c]	0.34				

[a] Original MD.
[b] MD without rigid body motion.
[c] Filtered MD (>1.6 cm^{-1}).

shows that the local maxima in the x-ray structure (at residues 11, 24, 64, 76, 118, 137, and 160) are all local maxima in the MD and NM curves as well. This similarity in pattern is reflected in the correlation coefficients of the calculated fluctuations and the experimental ones—0.82 and 0.60 for NM and MD, respectively (Table I). The expected similarity between the simulated values and the experimental ones is limited by the experimental accuracy: Liao and Herzberg [1] calculated the correlation coefficients between the C_α atom crystallographic temperature factors for a series of protein structures solved independently in two different laboratories. They found a wide range of variability in the correlation coefficents, from a low of 0.07 to a high of 0.98. Higher values (0.5 or better), tend to be found for higher resolution structures.

As shown in Table I, removing rigid body motion from the MD trajectory reduces the overall C_α rms fluctuations from 0.63 to 0.50 Å, without qualitatively changing the distribution as a function of residue number. The correlation with the NM motion increases to 0.62 and there is also a small increase in correlation with the experimental motion. Increased correlation with the NM motion is to be expected, since rigid body motion is not represented there. However, increased agreement with the experimental values is not expected, and suggests that the MD rigid body motion is unreliable (cf. below).

An examination of the trajectories of residues with fluctuations significantly higher than the corresponding experimental values showed that these are typically involved in non-oscillatory conformational transitions, as reflected in the large amplitude of the "non-oscillatory mode" (Eq. (12)). The low frequency sample modes (up to $\sim$1.6 cm^{-1}) were found to be highly correlated with the approximate "non-oscillatory mode." That is, sample modes up to this frequency have large amplitudes for atoms that undergo conformational transitions and relatively negligible amplitudes for all other atoms. The effect of such transitions was reduced by filtering the trajectory to eliminate motion with frequency less than $\sim$1.6 cm^{-1}, leaving mainly intramolecular oscillatory motion. The resulting average rms fluctuations in the filtered trajectory is reduced to 0.34 Å. Although this value is much lower than the experimental one, the correlation with experiment increases to 0.76, and a new peak in the MD fluctuation around residue 146 appears, in agreement with experiment and the NM results. In addition, the correlation coefficient with the NM calculations increased to 0.79. This similarity in intramolecular oscillatory motions obtained by the two methods is also evident in Fig. 1b.

Correlated Motion of Residues

A general indication of the degree of collective motion in the protein can be obtained from the equal-time cross-correlations of atomic fluctuations. We have calculated the cross-correlations of the fluctuations of all C_α's, for the lowest 32 normal modes and for the MD simulation. To gain further insight into the relation between the cross-correlations and the inter-residue distance, we compared the patterns in the distance map and cross-correlation maps. The distance map (Fig. 2a) shows a pattern characteristic of a β sheet structure with lines perpendicular to the diagonal corresponding to close residues along adjacent strands. A short helix at the C-terminus of the protein gives rise to the band parallel to the diagonal at the top right hand corner. The cross-correlations as a function of pairs of residues are shown in Figs. 2b–d. For the normal modes (Fig. 2b), the positive correlations closely mimic the regions of close inter-residue distance in Fig. 2a. There are also some long range positive and negative correlations. In the MD simulation, a similar short range positive correlation pattern

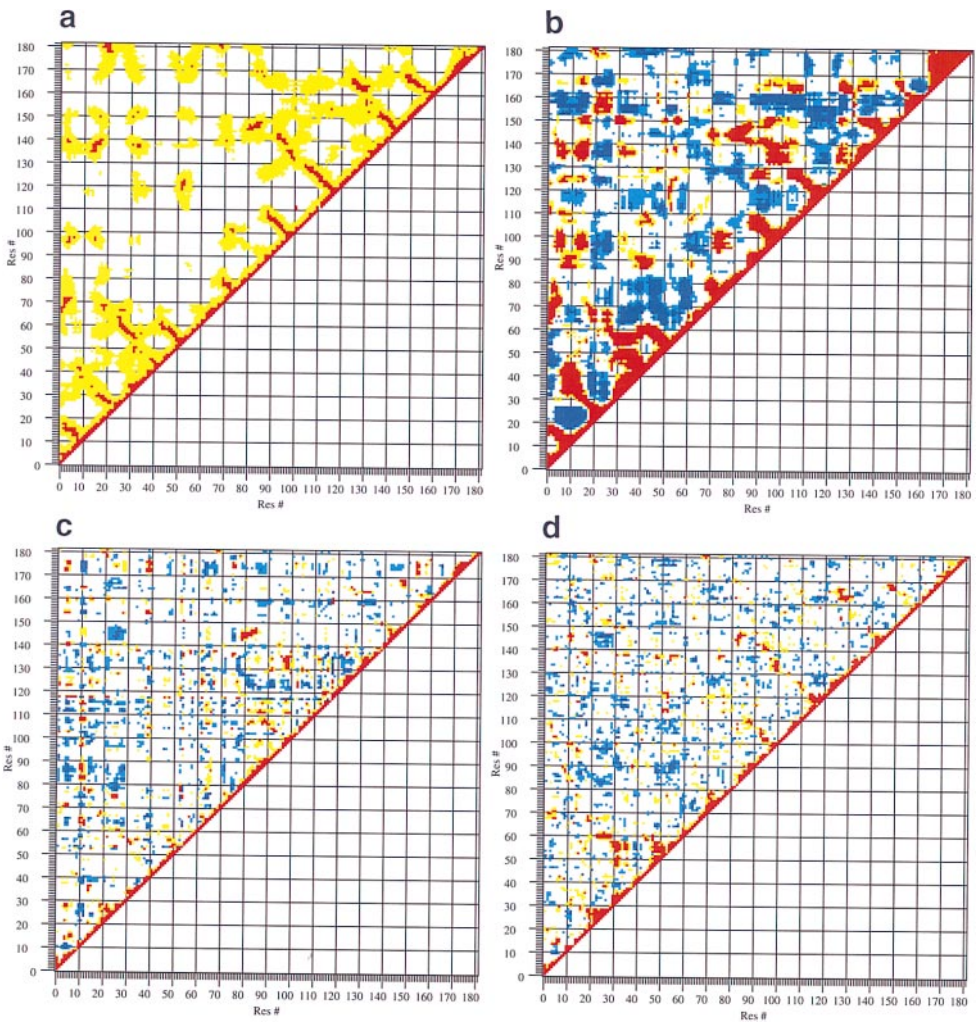


FIG. 2. (a) Inter-residue $C_\alpha \cdots C_\alpha$ distance map for *SGPA*. Distances less than 9 Å are shown in red, and those between 9 and 27 Å are yellow. Red regions extending perpendicular to the diagonal show the close contacts between neighboring anti-parallel β strands. Inter-residue equal-time cross-correlation maps, obtained from (b) NM analysis, (c) MD trajectory, after removal of rigid body motion, conformational transitions, and high frequency motion (>50 cm^{-1}). (d) Modes extracted from the MD trajectory, no rigid body motion, frequencies $3.3 < \nu < 50$ cm^{-1}. Correlations greater than 0.3 are shown in red, correlations between 0.2 and 0.3 in yellow, correlations between -0.2 and -0.3 in light blue, and those less than -0.3 in dark blue. Both types of motion show extensive positive correlations, but negative correlations are more pronounced in the normal modes.

is obtained, after eliminating rigid body motion and conformational transitions (Fig. 2c). The negative correlations are much less prominent, and only two of the longer range ones are evident. It is not possible to judge from the available data whether the strong long range negative correlations in the NM results are real, and conformational drift in the MD prevents them from being apparent there, or whether they are artifacts of the NM analysis. It should be noted, however, that these kinds of long range correlations have been observed in NM analysis of other proteins [11].

Frequency Distribution

Since the amplitude of a mode is inversely proportional to the frequency [11, 46], the most significant motion is produced by the low frequency modes. These are the modes which result in the largest amplitude motion, with the least energy change. The conformational motion which is of importance for biological events is expected to involve small energetic changes. Thus, we have concentrated on the low frequency motion of the protein as determined from normal mode dynamics and extraction of characteristic modes from molecular dynamics. The frequencies of the normal modes up to 50 cm^{-1} (127 modes) are plotted in Fig. 3a. The frequency distribution, $g(\nu)$, in both the original MD trajectory and after removing the rigid body motion is shown in Fig. 3b (0–50 cm^{-1}). The frequency distributions with and without rigid body motion are very similar except for a reduction in the value of $g(\nu)$ for very low frequencies.

Unlike the discrete frequencies in the normal mode dynamics, the frequency distribution in the MD simulation is continuous. This is a consequence of the fact that conformational variations during the simulation broaden the frequencies of the individual modes, causing them to overlap. In order to better compare the discrete NM spectrum with the continuous MD distributions, the NM distribution has been expressed as a mode density by convoluting it with a Gaussian function generating a full width half maximum resolution of 1.5 cm^{-1} [6]. The density of modes in the MD simulation increases steadily up to $\sim$40 cm^{-1}. A corresponding increase is observed for the normal mode dynamics.

We have extracted sample modes from the molecular dynamics simulation at intervals of 0.2 cm^{-1}, the highest resolution possible for the available simulation length, up to 15 cm^{-1},

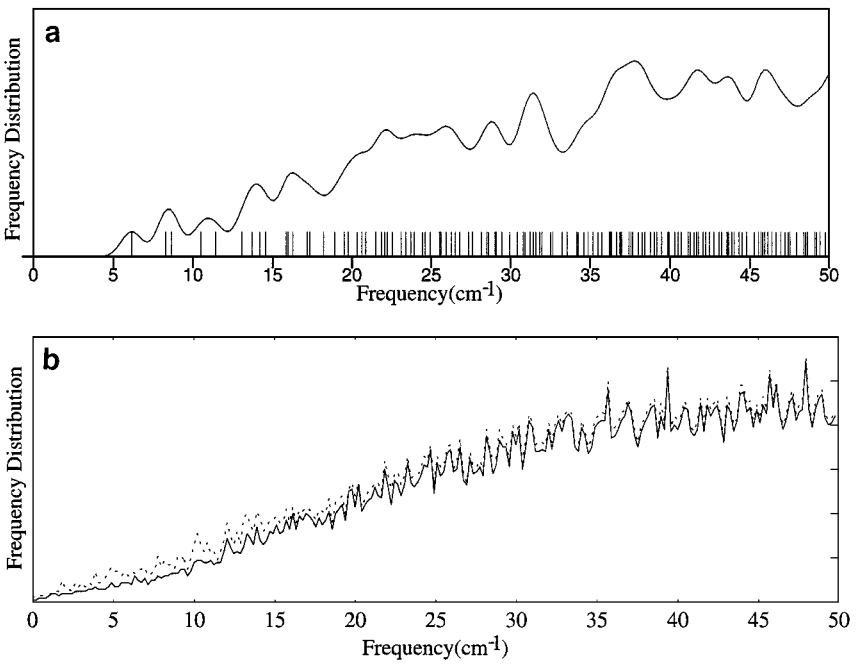


FIG. 3. Frequency distribution as obtained from (a) NM dynamics, (b) original MD trajectory (dotted line), and after removal of rigid body trajectory (solid line). The discrete frequencies of the normal modes are broadened with a resolution of 1.5 cm^{-1} to better compare them with the continuous MD distribution.

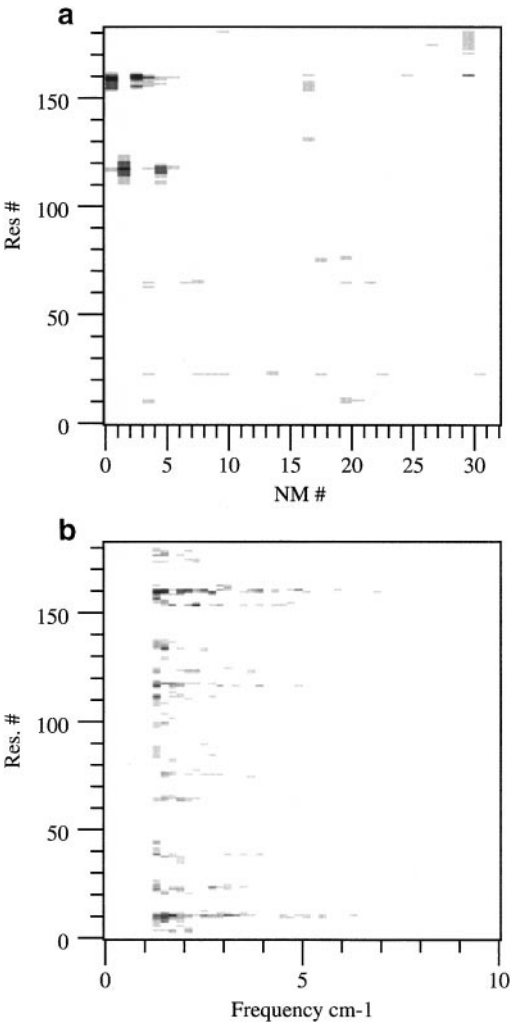


FIG. 4. Atomic amplitudes of motion as a function of residue number and frequency for (a) NM analysis, (b) sample modes from the MD trajectory after removal of rigid body motion and frequencies below 1.6 cm^{-1}. The shading indicates the relative amplitude of the motion of each residue in each mode. Residues with amplitudes between 70 and 100% of the largest value are black, amplitudes between 30 and 70% are in dark gray, amplitudes between 20 and 30% are in light gray (smaller amplitudes are not shown). MD modes are obtained by sampling from the continuous distribution. As expected, the largest amplitude motions are observed at the lowest frequencies. For the normal modes, motion is concentrated in a few local regions of the protein.

yielding a total of 75 modes. Thus, all motions with periods between ~164 and 2.2 ps are considered. At the low frequency end of the spectrum, the lowest normal mode is 6 cm^{-1}, while the MD trajectory represents motions down to 0.2 cm^{-1}. The trajectory is, therefore, longer than needed to cover all relevant motions. Rigid geometry calculations in general are known to over-estimate the stiffness of the energy surface [11, 47], so that normal mode calculations in dihedral space yield characteristic frequencies which are higher than those obtained in Cartesian space [48]. Therefore, in comparing the extracted MD modes to the normal modes we include a higher range of frequencies, the 32 modes up to 25 cm^{-1}, for the latter.

TABLE II
Regions in the Protein Selected for Analysis

Fragment	Residues		Description
	PDB[a]	Sequential	
a	29–44	5–16	Hairpin, very mobile
b	46–53	18–29	Hairpin, putative independent folding unit
c	89–105	47–58	Hairpin, low mobility
d	117–124	70–81	Insertion in domain linker, very mobile
e	166–179	111–124	Hairpin, very mobile
f	198–213	140–151	Hairpin, low mobility
g	214–228	152–166	Hairpin, specificity fragment, most mobile
h	231–241	169–180	C-terminal turn and helix

[a] Residue numbers as given in PDB entry 2SGA.

The very low frequency conformational transition mode (obtained from Eq. (12)) was found to correlate with the sample modes below 1.6 cm^{-1} with a correlation coefficient of 0.6–0.7, whereas the correlation with sample modes above 1.6 cm^{-1} is less than 0.4. It is, therefore, possible to approximately separate the oscillatory motion from the positional or conformational rearrangements: Sample modes below $\sim$1.6 cm^{-1} may be regarded as mainly due to non-oscillatory motion and those above as mainly due to oscillatory motion.

Analysis of Specific Structural Units

We have compared the low frequency modes from the two methods for the whole protein and for eight shorter structural elements (see Table II). These include the six β hairpin loops (three in each of the β barrel domains). Three of the β hairpins are the most mobile regions of structure seen in the crystal: residues 5–16, 111–124, and 152–166. The latter hairpin forms part of the primary specificity pocket of the enzyme. Its large amplitude motion is attributable to its surface location, and the presence of four glycine residues. However, the tip of the hairpin forms a disulfide bond with a remote piece of chain. Apparently, during the motion this link is able to lengthen in a manner that does not greatly restrict the amplitude.

By contrast, the two hairpins formed by residues 47–58 and 140–151 are two of the least mobile regions of the structure (cf. Fig. 1). (Residue 54 displays only small fluctuations in the experiment and NM dynamics, but larger fluctuations in the MD simulation. This is attributed to a conformational transition in the MD since the motion is small in the filtered trajectory.) The two other regions analyzed correspond to putative semi-independently folding fragments [49]. These are a hairpin-like insertion on the inter-domain linker region, relative to other chymotrypsin class proteins [25] (residues 70–81), and the C-terminal helix together with the turn preceding it (residues 169–180) [49]. Three of these regions (b, d, and the end of h) are involved in intermolecular interactions in the crystal which might affect the motion observed experimentally and by MD.

Analysis of Modes

Atomic amplitudes of motion. Figure 4a shows the relative amplitudes of the C_α atoms as a function of residue number for the 32 lowest normal modes. Examination of the

amplitudes of motion in the various normal modes reveals that the largest contribution to the overall motion comes from the low frequency modes, as expected. Similar behavior has been generally observed in previous NM calculations [11, 18, 40, 46]. As noted earlier, the low frequency modes are by nature low energy modes; i.e., it is less costly energetically to deform the molecule along the low frequency modes than to achieve fluctuations of the same magnitude by deforming it along high frequency modes. Since the overall energy is partitioned equally among the modes, distortions due to low frequency modes will account for most of the overall fluctuations. The largest amplitudes are obtained for fragments g and e in modes 1–5. Significant motion is also obtained for fragments a, b, d, h in some of the modes. The largest amplitudes within these fragments are obtained for the central part of the peptide, i.e., the turn regions. Examination of the distribution of amplitudes within each mode also reveals that the very lowest modes (i.e., 1–3, 5) are virtually localized on one fragment, whereas the motion of most of the higher modes is spread over larger parts of the protein.

The amplitudes of oscillatory motion in the MD simulation are given in Fig. 4b. At the low frequency end, the motion is very pronounced in the fragment regions, particularly fragments a, e, and g. At higher frequencies the motion is more equally distributed along the protein, in qualitative agreement with the results from the normal mode dynamics.

Correlated motion in the extracted MD modes. We have used the sample modes extracted from the MD trajectory to recalculate the cross-correlations discussed above, so as to examine whether the extraction process retained the characteristic features of the motion. Figure 2d shows the inter-residue cross-correlations obtained using all sample modes with frequencies between 3.3 and 50 cm^{-1}. The pattern created by the positive correlations is very similar to the pattern obtained from the original MD (Fig. 2c) and to that representing close inter-residue distances (Fig. 2a). However, the high correlations in Fig. 2c appear in regions corresponding to somewhat larger inter-residue distances than the high correlations in Fig. 2d. In addition, the negative correlations are smaller than those obtained from the original trajectory. Correlation maps which include sample modes with frequency lower than $\sim$3 cm^{-1} do not reveal these characteristic patterns as clearly, due to the effects of non-oscillatory motions.

Correlation between NM and MD modes. We compared the NM and MD motion in protein regions of high amplitude (fragments a, b, d, e, g, h) as well as for two fragments with relatively low amplitudes and for the whole protein. The similarity between the modes, as determined by the dot product between the corresponding modes, is shown in Fig. 5a for fragment g. In addition Fig. 5b shows the correlation between the NM and the MD modes for the whole protein. For comparison, we also calculated the correlation between the MD modes for fragment g with a set of 32 random modes. Overall measures for the similarity of motion of the various regions as described in Methods are given in Table III. The fraction of sample modes with a high correlation to at least one normal mode is given by F, while F_{amp} is weighted by the amplitudes of the modes (Eq. (11)).

For all fragments there is a significant correlation between the extracted MD modes and some of the normal modes. This is demonstrated both in terms of the number of "hits," F, and in terms of the overall amplitude, F_{amp}, accounted for by sample modes with similarity to normal modes. The correlations are particularly striking when compared to the extremely low values between the MD modes and modes generated randomly. Thus, for these short

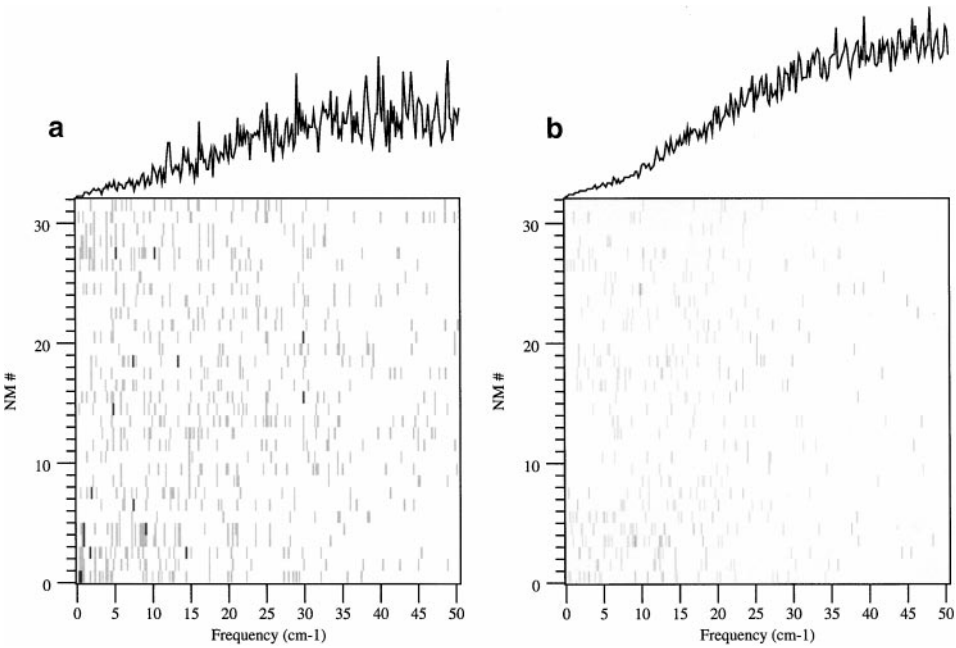


FIG. 5. Correlation between normal modes (vertical axis) and modes extracted from the MD trajectory, after removing rigid body motion (horizontal axis). The corresponding MD frequency distributions are shown above each plot. Correlated motions between individual NM and MD modes are shown as short bars, shaded as follows: Correlations between 0.7 and 1.0 are in black, between 0.3 and 0.7 in dark gray, and between 0.2 and 0.3 in light gray (smaller correlations are not shown). The correlations for fragment *g* are shown in (a) and for the whole protein in (b). For the fragment, there is a strong correlation between a few NM and MD modes. These correlations are not apparent at the whole protein level.

regions of the protein, there are typical modes of motion which are manifested in the normal modes and in the molecular dynamics simulation.

Figure 6a shows a comparison of the extracted MD modes with three of the normal modes of fragment *g*. Each diagram displays the atomic displacements corresponding to

TABLE III
Correlation between NM and MD Modes

Fragment	F^a	$F_{amp}{}^b$
a	0.71	0.70
b	0.84	0.75
c	0.85	0.61
d	0.85	0.82
e	0.68	0.66
f	0.84	0.65
g	0.68	0.62
h	0.87	0.79
g^c	<0.1	

[a] Fraction of MD sample modes with dot products greater than 0.4 with one or more normal modes.

[b] Weighted by mode amplitude (see Eq. (11)).

[c] Comparison with random modes.

one normal mode. Similar results were obtained for the other fragments. As seen in Fig. 5b, the correlation between the MD modes and the normal modes is much smaller when the C_α atoms of the whole protein are compared. For the whole protein, there were no correlations higher than 0.4. In Fig. 6b we show normal mode 4 along with the extracted MD mode with the highest correlation to it (0.4). It is clear that although in some regions the two modes are very similar (e.g., fragments *b*, *c*, and *g* have correlations of 0.5, 0.6, and 0.6, respectively), other regions differ in absolute amplitude or direction. For another pair of normal modes and MD modes with an overall correlation of 0.4, fragments *b*, *c*, and *f* have correlations of 0.5, 0.7, and 0.6, respectively. These results tend to indicate that although there are typical regional motions, which are similar in MD and NM, these do not combine in a fixed manner in terms of relative magnitude and phase. The results are in accord with the conclusions obtained by examining the equal-time correlated motion of residues, discussed above. That is, both NM and MD calculations display large positive correlations for residues in spatial proximity. However, while the NM dynamics result in large negative correlations for some distant residues, hardly any significant long range correlations are observed in MD.

CONCLUSIONS

Normal Mode, Molecular Dynamics, and Experimental Motions Are Similar

Molecular dynamics and normal mode analysis provide an approximate description of the motion of protein molecules. It is difficult to evaluate the validity of features of the motion obtained using either method alone. The use of a mode analysis of a molecular dynamics trajectory has made it possible to compare the two methods more directly than was previously possible. The surprising result is that although there are some detailed differences in behavior, the two methods produce a qualitatively similar overall picture of the motion of a protein molecule, and both are in agreement with the available experimental data. A range of diverse properties leads to this conclusion:

Amplitude of atomic fluctuations. There is a high correlation between the amplitude of the atomic fluctuations obtained by MD and NM analyses, and between both of these with the amplitudes deduced from the crystallographic temperature factors.

Correlated motions of residues. Equal-time correlations from normal modes, directly from MD, and reconstructed from extracted MD modes, all show positive correlations for motion of residues within the β hairpins in the structure. The results for longer range correlation do differ, however.

Frequency distributions. Although the frequency distribution is discrete in one case (NM) and continuous in the other, the density of modes as a function of frequency is similar.

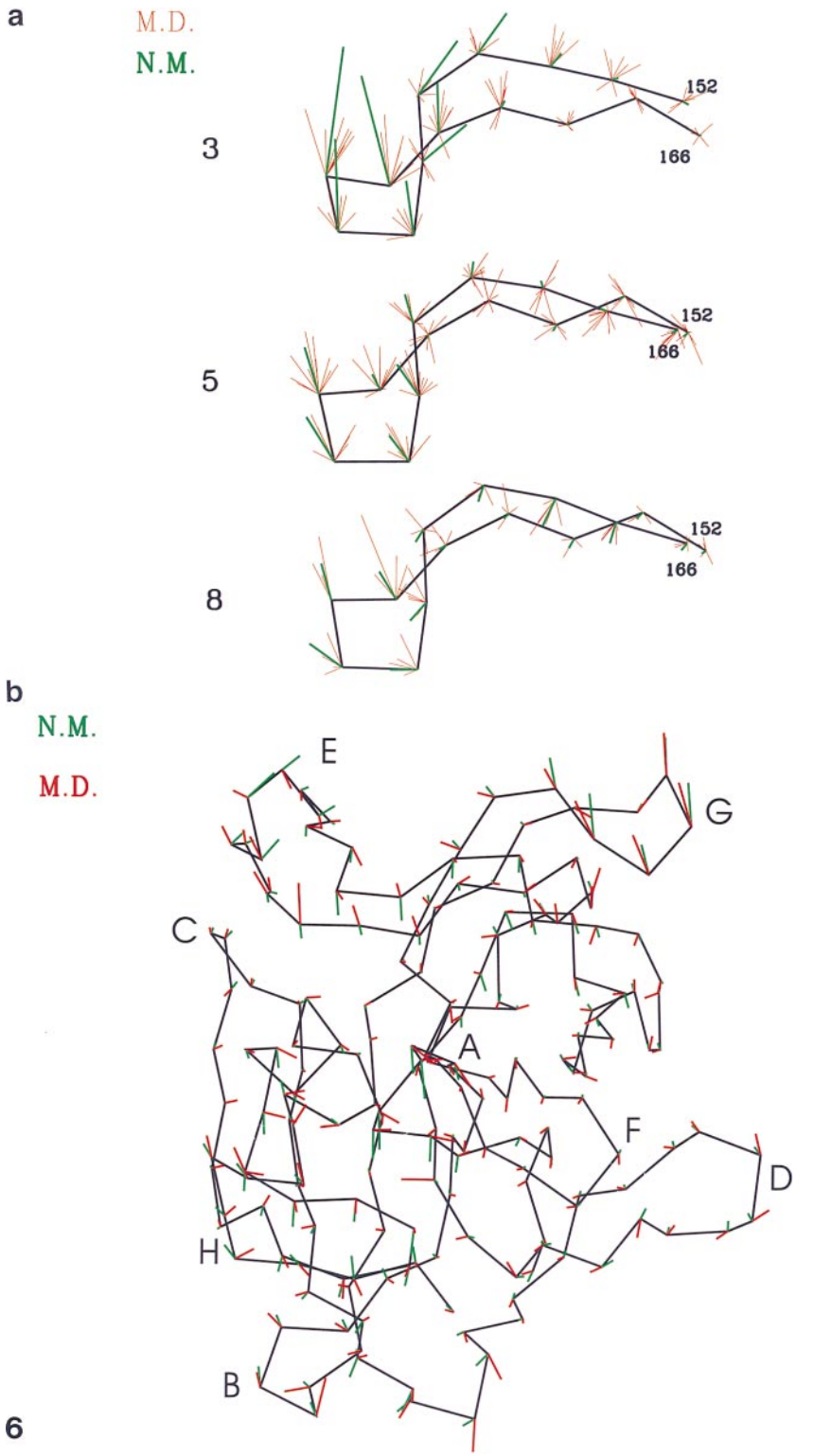
Correlation between normal modes and sample modes from the MD. For short fragments of the chain, there is a high correlation between the motion represented by individual MD modes and particular NM ones. This is not true for the whole protein.

Why are the Motions So Similar?

Both methods are affected by a number of approximations. In MD, imperfections in the potential lead to a slow drift away from the x-ray structure. In the NM calculations, the force field and cutoff used are different from those used in the MD calculations, and a torsion space representation is used, rather than the MD Cartesian one. Brooks *et al.* [36] and de Groot *et al.* [38] have presented criteria for the analysis of the similarity of low frequency normal modes. The latter authors used a penalty function emphasizing differences to compare sets of eigenvalues and found substantial overlap of the essential subspace derived from simulations, using different parameters, such as with and without solvent or using different cutoff values. Both Kitao and G$\bar{o}$ [48] and Janežič and Brooks [50] have shown that the torsion space representation produces normal modes similar to those obtained in Cartesian space, although the frequencies are shifted to higher values. In addition to these factors, the assumptions of harmonic motion, extrapolation of motion out from a minimum energy conformation, and neglect of the solvent and crystal environment might be expected to have a dramatic effect on the significance of the results. Yet, in spite of these factors, qualitatively similar behavior is found. This is in agreement with the findings of other studies comparing NM analysis and MD simulations on three different proteins [51, 52, 20]. Thus, we conclude that the general agreement of the results from the two methods, and with experiment, indicates that there is a basic robustness about the protein motion which is easily captured. It would appear that particular directions are relatively free of van der Waals clashes, and not restrained by strong interactions, while others offer strong resistance to motion. For example, the most mobile hairpin, fragment *g*, undergoes a flapping motion, away from the protein surface and back again. Such a motion is restrained only by the presence of solvent and the strength of the van der Waals interaction with the rest of the protein. By contrast, the least mobile hairpin, fragment *c*, is restrained from such a motion by helices lying approximately above and below the plane of the hairpin.

Are These Motions Co-incidental Properties of the Structure, or Do Some of Them Have a Functional Role?

A surprising feature of the modes is that none of the low frequency ones involve a relative motion of the two well defined domains. In other cases, for example, hen egg white lysozyme, such motions are very clear [11]. These inter-domain movements may facilitate the entrance and exit of substrates in enzymes. In the case of *SGPA*, the motion of the largest amplitude mode is concentrated in the "specificity loop" [25]. This β hairpin forms part of the lining of the pocket into which the side chain of the residue on the N-terminal side of the scissile bond of a substrate fits. Variations in the lining of this pocket are primarily responsible for the differences in specificity among the different members of the chymotrypsin class serine proteases. Figure 7 shows the main chain of *SGPA* in this region, with the specificity loop colored red. A peptide inhibitor is also shown, with co-ordinates taken from the PDB file 3SGA. The C-terminal phenlyalanine side chain occupies the pocket. In the apo enzyme, this pocket is presumably occupied by solvent, which must be displaced upon substrate binding. The large amplitude motion of the mode appears to facilitate squeezing out the solvent as the substrate side chain enters.


a
M.D.
N.M.
3
152
166
5
152
166
8
152
166
b
N.M.
M.D.
E
G
C
A
F
D
H
B
6

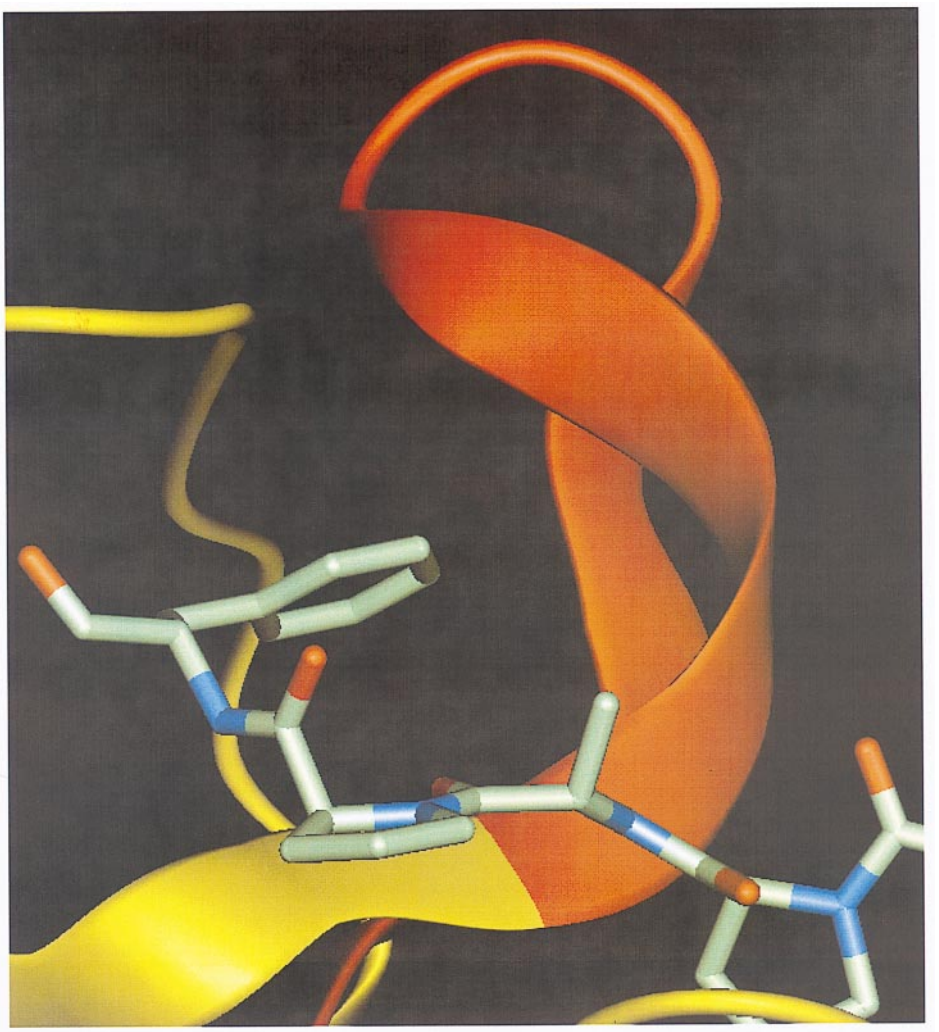


FIG. 7. The relationship between the most mobile loop (residues 152–166) and the primary specificity pocket of *SGPA*. The main chain of *SGPA* in this region is shown, with the specificity loop colored red. A peptide inhibitor is also shown, with co-ordinates taken from the PDB file 3SGA. The phenylalanine side chain occupies the primary specificity pocket.

FIG. 6. Normal mode vectors (green) at each C_α and corresponding sample modes extracted from the MD trajectory with no rigid body motion (red). (a) Normal modes 3, 5, and 8 for fragment *g* and all sample MD modes having correlations larger than 0.5 with these. (b) Normal mode 4 and the sample MD mode having the highest correlation with it, for the whole protein. As expected from the comparisons of NM and MD modes shown in Fig. 5, there is a high correlation at the fragment level, but not for the whole protein.

REFERENCES

1. D.-I. Liao and O. Herzberg, Refined structures of the active Ser83 → Cys and impaired Ser46 → Asp histidine-containing phosphocarrier proteins, *Structure* **2**, 1203 (1994).
2. D. L. D. Caspar, J. Clarage, D. M. Salunke, and M. Clarage, Liquid-like movements in crystalline insulin, *Nature (London)* **332**, 659 (1988).
3. N.-T. Yu, Raman spectroscopy: A conformational probe in biochemistry, *CRC Crit. Rev. Biochem.* **4**, 229 (1977).
4. J. R. Lakowicz, B. Maliwal, H. Cherek, and A. Balter, Rotational freedom of tryptophan residues in proteins and peptides, *Biochemistry* **22**, 1741 (1983).
5. M. C. Chang, A. J. Cross, and G. R. Fleming, Internal dynamics and overall motion of lysozyme studied by fluorescence depolarization of the eosin lysozyme complex, *J. Biomolec. Struct. Dyn.* **1**, 299 (1983).
6. S. Cusack, J. Smith, J. Finney, B. Tidor, and M. Karplus, Inelastic neutron scattering analysis of picosecond internal protein dynamics. Comparison of harmonic theory with experiment, *J. Mol. Biol.* **202**, 903 (1988).
7. R. J. P. Williams, NMR studies of mobility within protein structure, *Eur. J. Biochem.* **183**, 479 (1989).
8. G. Herzberg, *Molecular Spectra and Molecular Structure II. Infrared and Raman Spectra of Polyatomic Molecules* (Van Nostrand, New York, 1945).
9. E. B. Wilson, J. C. Decius, and P. C. Cross, *Molecular Vibrations* (McGraw–Hill, New York, 1955).
10. K. Itoh and T. Shimanouchi, Vibrational frequencies and modes of alpha-helix, *Biopolymers* **9**, 383 (1970).
11. M. Levitt, C. Sander, and P. S. Stern, Protein normal-made dynamics: Trypsin inhibitor, crambin, ribonuclease and lysozyme, *J. Mol. Biol.* **181**, 423 (1985).
12. B. Brooks and M. Karplus, Normal modes for specific motions of macromolecules: Application to the hinge-bending mode of lysozyme, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4995 (1985).
13. T. Nishikawa and N. Gō, Normal modes of vibration in bovine pancreatic trypsin inhibitor and its mechanical property, *Proteins* **2**, 308 (1987).
14. M. P. Allen and D. J. Tildesly, *Computer Simulations of Liquids* (Clarendon, Oxford, 1987).
15. J. A. McCammon and S. C. Harvey, *Dynamics of Proteins and Nucleic Acids* (Cambridge Univ. Press, Cambridge, UK, 1987).
16. T. Ichiye and M. Karplus, Collective motions in proteins: A covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations, *Proteins* **11**, 205 (1991).
17. A. Amadei, A. B. M. Linssen, and H. J. C. Berendsen, Essential dynamics of proteins, *Proteins* **17**, 412 (1993).
18. T. Horiuchi and N. Gō, Projection of Monte Carlo and molecular dynamics trajectories onto the normal mode axes: Human lysozyme, *Proteins* **10**, 106 (1991).
19. A. Kitao, F. Hirata, and N. Gō, The effects of solvent on the conformation and the collective motions of protein: Normal mode analysis and molecular dynamics simulations of melittin in water and in vacuum, *Chem. Phys.* **158**, 447 (1991).
20. S. Hayward, A. Kitao, and H. J. C. Berendsen, Model-free methods of analyzing domain motions in proteins from simulation: A comparison of normal mode analysis and molecular dynamics simulation of lysozyme, *Proteins* **27**, 425 (1997).
21. P. Dauber-Osguthorpe and D. J. Osguthorpe, Analysis of intra-molecular motions by filtering molecular dynamics trajectories, *J. Am. Chem. Soc.* **112**, 7921 (1990).
22. P. Dauber-Osguthorpe and D. J. Osguthorpe, Extraction of the energetics of selected types of motion from molecular dynamics trajectories by filtering, *Biochemistry* **29**, 8223 (1990).
23. R. B. Sessions, P. Dauber-Osguthorpe, and D. J. Osguthorpe, Filtering molecular dynamics trajectories to reveal low-frequency collective motions: Phospholipase A2, *J. Mol. Biol.* **210**, 617 (1989).
24. D. J. Osguthorpe and P. Dauber-Osguthorpe, FOCUS: A program for analyzing molecular dynamics simulations, featuring digital signal-processing techniques, *J. Mol. Graph.* **10**, 178 (1992).
25. A. Sielecki, W. A. Hendrickson, C. G. Broughton, L. T. J. Delbaere, G. D. Brayer, and M. N. G. James, Protein structure refinement: Streptomyces griseus serine protease A at 1.8 Å resolution, *J. Mol. Biol.* **134**, 781 (1979).

26. F. Avbelj, J. Moult, D. H. Kitson, M. N. G. James, and A. T. Hagler, Molecular dynamics study of the structure and dynamics of a protein molecule in a crystalline ionic environment, Streptomyces griseus protease A, *Biochemistry* **29**, 8658 (1990).
27. M. Levitt, Protein folding by restrained energy minimization and molecular dynamics, *J. Mol. Biol.* **170**, 723 (1983).
28. M. Levitt, Molecular dynamics of native protein. I. Computer simulation of trajectories, *J. Mol. Biol.* **168**, 595 (1983).
29. A. T. Hagler, E. Huler, and S. Lifson, Energy functions for peptides and proteins. I. Derivation of a consistent force field including the hydrogen bond from amide crystals, *J. Am. Chem. Soc.* **96**, 5319 (1974).
30. M. Levitt, C. Sander, and P. S. Stern, The normal modes of a protein: Native bovine pancreatic trypsin inhibitor, *Int. J. Quant. Chem.: Quant. Biol. Symp.* **10**, 181 (1983).
31. L. Verlet, Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules, *Phys. Rev.* **159**, 98 (1967).
32. P. Dauber-Osguthorpe, V. A. Roberts, D. J. Osguthorpe, J. Wolff, M. Genest, and A. T. Hagler, Structure and energetics of ligand binding to proteins: Escherichia coli dihydrofolate reductase-trimethoprim, drug–receptor system, *Proteins* **4**, 31 (1988).
33. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, Molecular dynamics with coupling to an external bath, *J. Chem. Phys.* **81**, 3684 (1984).
34. E. O. Brigham, *The Fast Fourier Transform and Its Applications* (Prentice-Hall, Englewood Cliffs, NJ, 1988).
35. P. Dauber-Osguthorpe and D. J. Osguthorpe, Partitioning the motion in molecular dynamics simulations into characteristic modes of motion, *J. Comput. Chem.* **14**, 1259 (1993).
36. B. R. Brooks, D. Janežič, and M. Karplus, Harmonic analysis of large systems. I. Methodology, *J. Comput. Chem.* **16**, 1522 (1995).
37. M. A. Balsera, W. Wriggers, Y. Oona, and K. Schulten, Principal component and long time protein dynamics, *J. Phys. Chem.* **100**, 2567 (1996).
38. B. L. de Groot, D. M. F. van Aalten, A. Amadei, and H. J. C. Berendsen, The consistency of large concerted motions in proteins in molecular dynamics simulations, *Biophys. J.* **71**, 1707 (1996).
39. R. W. James, *The Optical Principles of the Diffraction of X-Rays* (Ox Bow, Woodbridge, 1962).
40. J.-F. Gibrat and N. Gō, Normal mode analysis of human lysozyme: Study of the relative motion of the two domains and characterization of the harmonic motion, *Proteins* **8**, 258 (1990).
41. Y. Seno and N. Gō, Deoxymyoglobin studied by the conformational normal mode analysis. I. Dynamics of globin and the heme–globin interaction, *J. Mol. Biol.* **216**, 95 (1990).
42. P. Derreumaux and G. Vergoten, Effect of Urey–Bradley–Shimanouchi force field on the harmonic dynamics of proteins, *Proteins* **11**, 120 (1991).
43. M. M. Tirion and D. ben-Avraham, Normal mode analysis of G-actin, *J. Mol. Biol.* **230**, 186 (1993).
44. J. Kuriyan, G. A. Petsko, R. M. Levy, and M. Karplus, Effect of anisotropy and anharmonicity on protein crystallographic refinement. An evaluation by molecular dynamics, *J. Mol. Biol.* **190**, 227 (1986).
45. P. H. Hunenburger, A. E. Mark, and W. F. van Gunsteren, Fluctuation and cross-correlation analysis of protein motions observed in nanosecond molecular dynamics simulations, *J. Mol. Biol.* **252**, 492 (1995).
46. N. Gō, A theorem on amplitudes of thermal atomic fluctuations in large molecules assuming specific conformations calculated by normal mode analysis, *Biophys. Chem.* **35**, 105 (1990).
47. A. Kitao, S. Hayward, and N. Gō, Comparison of normal-mode analyses on a small globular protein in dihedral angle space and cartesian coordinate space, *Biophys. Chem.* **52**, 107 (1994).
48. A. Kitao and N. Gō, Conformational dynamics of polypeptides and proteins in the dihedral angle space and in the cartesian coordinate space, *J. Comput. Chem.* **12**, 359 (1991).
49. J. Mount and R. Unger, An analysis of protein folding pathways, *Biochemistry* **30**, 3816 (1991).
50. D. Janežič and B. R. Brooks, Harmonic analysis of large systems. II. Comparison of different protein models, *J. Comput. Chem.* **16**, 1543 (1995).
51. M. M. Teeter and D. A. Case, Harmonic and quasiharmonic descriptions of crambin, *J. Phys. Chem.* **94**, 8091 (1990).
52. S. Hayward, A. Kitao, and N. Gō, Harmonicity and anharmonicity in protein dynamics: A normal mode analysis and principal component analysis, *Proteins* **23**, 177 (1995).