Dynamics of a small globular protein in terms of low-frequency vibrational modes

(bovine pancreatic trypsin inhibitor/conformational energy/fluctuations/anharmonicity)

NOBUHIRO GŌ, TOSIYUKI NOGUTI, AND TESTUO NISHIKAWA

Department of Physics, Faculty of Science, Kyushu University 33, Fukuoka 812, Japan

ABSTRACT Normal modes of low-frequency vibrations are calculated for a small globular protein, bovine pancreatic trypsin inhibitor. In modes with frequencies below 120 cm^{-1} the protein molecule behaves like a continuous elastic body. Most modes with frequencies above 50 cm^{-1} are shown to behave harmonically within the range of thermal fluctuations at room temperature. Those with frequencies below 50 cm^{-1} show some anharmonicity. Magnitudes of displacements of atoms are mainly determined by the modes with frequencies below 30 cm^{-1}. These very-low-frequency modes contribute significantly to the entropy of the system. The dynamic structure of the globular protein is described as a superposition of harmonic high-frequency motions and coupled anharmonic low-frequency motions of collective variables corresponding to the normal modes of vibration.

The mechanism of function of protein molecules can be elucidated by knowledge of molecular conformations and their dynamics. The variety of internal motions in globular proteins is large. Certain structural transitions occur in the time range of more than 10^2 s, whereas localized vibrations are characterized by the time range around 0.1 ps (1). Existence of nonlocalized vibrational modes is expected in the time range between 0.1 and 10 ps (2, 3). We have developed a method of calculating these modes in globular proteins and have applied it to a small globular protein, bovine pancreatic trypsin inhibitor (BPTI), which consists of 58 amino acids. This article is devoted to characterization of these vibrational modes and to derivation therefrom of a picture of the dynamic structure of the native protein.

Vibrational motions in globular proteins with frequencies lower than 200 cm^{-1} (expressed in the wave number of corresponding light) are brought about mainly by concerted variations in soft variables—i.e., dihedral angles. Coupling with variations in hard variables (bond lengths and bond angles) is expected to be small in these modes. From this consideration we treat bond lengths and bond angles as fixed and treat only dihedral angles of the main chain (ϕ, ψ) and side chains (χ) as dynamic variables. In this treatment the conformational space consists of 241 variables (116 main-chain and 125 side-chain angles). Hydrogen atoms are not treated individually but are included with the atom to which they are bonded, except for those expected to participate in hydrogen bonding. Solvent effects are not considered; that is, the molecule is treated as being *in vacuo*.

The dynamics of a protein molecule is determined by both the kinetic and the potential energy functions expressed in terms of dynamic variables. In the analysis of conformations of biopolymers the potential energy is referred to as the conformational energy. In a recent paper (4) we examined the shape of the conformational energy surface of an isolated BPTI molecule

near a minimum point corresponding to the x-ray crystal structure (5). Very near to the minimum, the energy surface is a multidimensional parabola characterized by the second-derivative matrix F at the minimum. We observed that the energy surface was in fact a good parabola within the range of thermal fluctuations along 80% of orthogonal directions that diagonalize the second-derivative matrix F. Along the remaining 20% of orthogonal directions moderate anharmonicities of the energy surface were observed. This finding indicates that description of the dynamic structure of the protein molecule in terms of collective variables may be useful.

In the present article we assume that the conformational energy E can be approximated by the multidimensional parabola characterized by the second-derivative matrix F at the minimum—i.e.,

$$E = \frac{1}{2} \sum_{\mu\nu} f_{\mu\nu} \Delta\theta_\mu \Delta\theta_\nu, \qquad [1]$$

in which $\Delta\theta_\mu$ is the deviation of the μth dihedral angle from that at the minimum point and $f_{\mu\nu}$ is an element of the matrix F. The coefficient matrix $H = (h_{\mu\nu})$ in an expression for the kinetic energy T,

$$T = \frac{1}{2} \sum_{\mu\nu} h_{\mu\nu} \Delta\dot{\theta}_\mu \Delta\dot{\theta}_\nu, \qquad [2]$$

of the internal motions in the protein molecule is calculated by the method developed recently by us and described elsewhere. Normal modes of vibration and their frequencies are obtained by solving the generalized eigenvalue equation

$$F\Delta\theta_i = (2\pi\nu_i)^2 H\Delta\theta_i, \qquad [3]$$

in which ν_i is the frequency of the ith normal mode and $\Delta\theta_i$ is the vector whose μth component $\Delta\theta_{i\mu}$ is the amplitude of variations in dihedral angle θ_μ in the ith normal mode. Use is made of high-speed high-precision computer program NICER (6) for solving this equation.

In this article we characterize the calculated normal vibrational modes and then discuss possible effects of anharmonicity neglected in Eq. 1.

Frequency spectrum

A histogram of calculated frequencies is shown in Fig. 1. Because there are 241 variable dihedral angles in our treatment of the protein molecule, we obtain 241 different normal modes of vibration. Eighty-eight percent of these normal modes have frequencies below 200 cm^{-1}. The treatment of fixing bond lengths and bond angles is expected to be justified in these low-

---

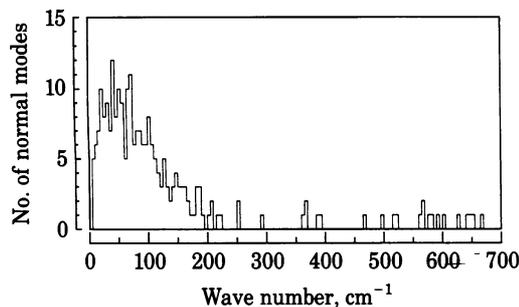Abbreviation: BPTI, bovine pancreatic trypsin inhibitor.

FIG. 1. Histogram of calculated frequencies of normal modes of vibration in a small globular protein, BPTI. Low-frequency motions are studied by treating only dihedral angles as variables. The frequencies are shown by corresponding light wave numbers. Number of normal modes in each interval of 5 cm^{-1} is shown.

frequency modes. In modes with frequencies higher than 200 cm^{-1}, coupling with variations in bond lengths and angles is expected to be significant, which makes the calculated frequencies in this high-frequency region approximate.

Protein molecules are expected to behave like an elastic body for very-low-frequency vibrational modes. By treating a protein molecule as a continuous elastic body, the number of such very-low-frequency vibrational modes has been estimated to be about 3 times the number of amino acid residues in a protein molecule. When the value of 10^{11} dyne cm^{-2} is assumed for Young's modulus E of the elastic material, these modes have been found to have frequencies in the range up to 120 cm^{-1} (7, 8). The number of calculated vibrational modes below 120 cm^{-1} in Fig. 1 is 174, exactly 3 times the number of amino acid residues, 58, in BPTI. This indicates that the assumed value of Young's modulus is in the right range.

The method of calculating the normal modes of vibration in this article is designed to be accurate in the lower-frequency region. The method of calculation of normal vibrations in proteins reported in a recent paper by Tasumi *et al.* (9) is complementary to ours. They calculated normal vibrations in a polypeptide hormone, glucagon, by treating Cartesian coordinates of atoms as independent variables and by using refined force fields for variations in bond lengths and bond angles, but by completely neglecting intramolecular interactions between atoms separated far from each other along the chain. The neglected interactions are the main source of the restoring forces from such deformations as are occurring in a continuous elastic body. Therefore, their method is not reliable for the lower-frequency modes in which the protein behaves like a continuous elastic body, but it is designed to be accurate for vibrational modes in which a group of atoms near each other along the chain is involved.

Normal modes of vibration

In the normal modes with frequencies higher than 200 cm^{-1}, motions of atoms in the protein molecule are localized in one or a few side chains. Most of the normal modes with frequencies in the range of 120–200 cm^{-1} have atomic motions localized in several residues that are close to each other. An example is given in Fig. 2B. In most of the normal modes with frequencies lower than 120 cm^{-1}, atomic motions are not localized—i.e., almost all atoms in the molecule move in concerted ways. An example is given in Fig. 2C for the second-lowest frequency mode. Displacement vectors of atoms close to each other are similar—i.e., their spatial changes are continuous. This property is general in most of the nonlocalized vibrational modes

with frequencies lower than 120 cm^{-1}. In these modes the protein molecule behaves like a continuous body. This fact justifies the previous treatment (2, 7, 8) of protein molecules as continuous elastic bodies for discussions of lower-frequency modes.

There are five modes in the range of wave number between 5 and 10 cm^{-1} or in the time range between 3 and 7 ps. The wave number of the lowest frequency mode is 5.7 cm^{-1}. This value is smaller than the value, 26 cm^{-1}, obtained previously by assuming that a protein molecule is a sphere of radius of 20 Å made of continuous elastic material with Young's modulus of 10^{11} dyne cm^{-2} (2). This difference may be partly due to the "slender" shape of BPTI and partly due to deviation of the behavior of BPTI from that expected for a continuous elastic body.

Anharmonicity

Because of the earlier observation of moderate anharmonicities of the energy surface (4), we examined the shape of the energy surface for conformational deformations along directions of the normal vibrational modes. These directions are not orthogonal to each other in general. Similar anharmonicities are again observed. Especially along directions of the lowest-frequency modes, the energy curves generally increase more rapidly than the parabolic behavior as the conformation deviates from the minimum point. This means that, when the calculated energy curves are represented by fourth-order polynomials, the fourth-order terms have significant positive contributions within the range of thermal fluctuations.

As an indicator of harmonicity of the energy surface along the directions of the normal vibrational modes, the mean-square deviation of the conformational point in the conformational space is calculated both from the calculated values of the second derivatives and from the calculated energy curves along the directions of the normal modes, and then their ratios are obtained. Results are given in Table 1. If a calculated energy curve is well approximated by a parabola, the ratio is close to unity. The results in Table 1 indicate that 70% of the modes are indeed good harmonic modes. In particular, most modes with frequencies higher than 50 cm^{-1} are harmonic. Anharmonic modes with frequencies higher than 180 cm^{-1}, including the six modes for which the energy curves cannot be fitted well by the fourth-order polynomials, correspond to isolated motions in one or two side chains on the surface of the molecule. Many modes with frequencies lower than 50 cm^{-1} show some anharmonicities with significant positive fourth-order terms in energy curves. The anharmonicities in these modes may explain the nonlinear or anharmonic aspects in dynamic protein structures that have been discussed in the literature (10–12).

It is likely that these anharmonic low-frequency modes are coupled to each other through higher-order cross terms. If this coupling is strong, each of the very-low-frequency modes loses the advantage of being isolated from other degrees of freedom, and their dynamics must be treated as one inseparable object. The wave number of 50 cm^{-1} corresponds to period of 0.7 ps for the vibration. The harmonic nature of most modes with frequencies higher than 50 cm^{-1} indicates that the dynamics of the structure in the time range faster than 0.7 ps is basically harmonic. In the time range between 0.7 and 7 ps, which corresponds to wave numbers between 50 and 5 cm^{-1}, the dynamics may be appreciably anharmonic. The dynamic structure in this time range is well described in terms of a relatively small number of the collective variables corresponding to the very-low-frequency modes. Thus, the concept of the collective variables corresponding to the very-low-frequency modes is useful even when nonlinear couplings among them are strong.
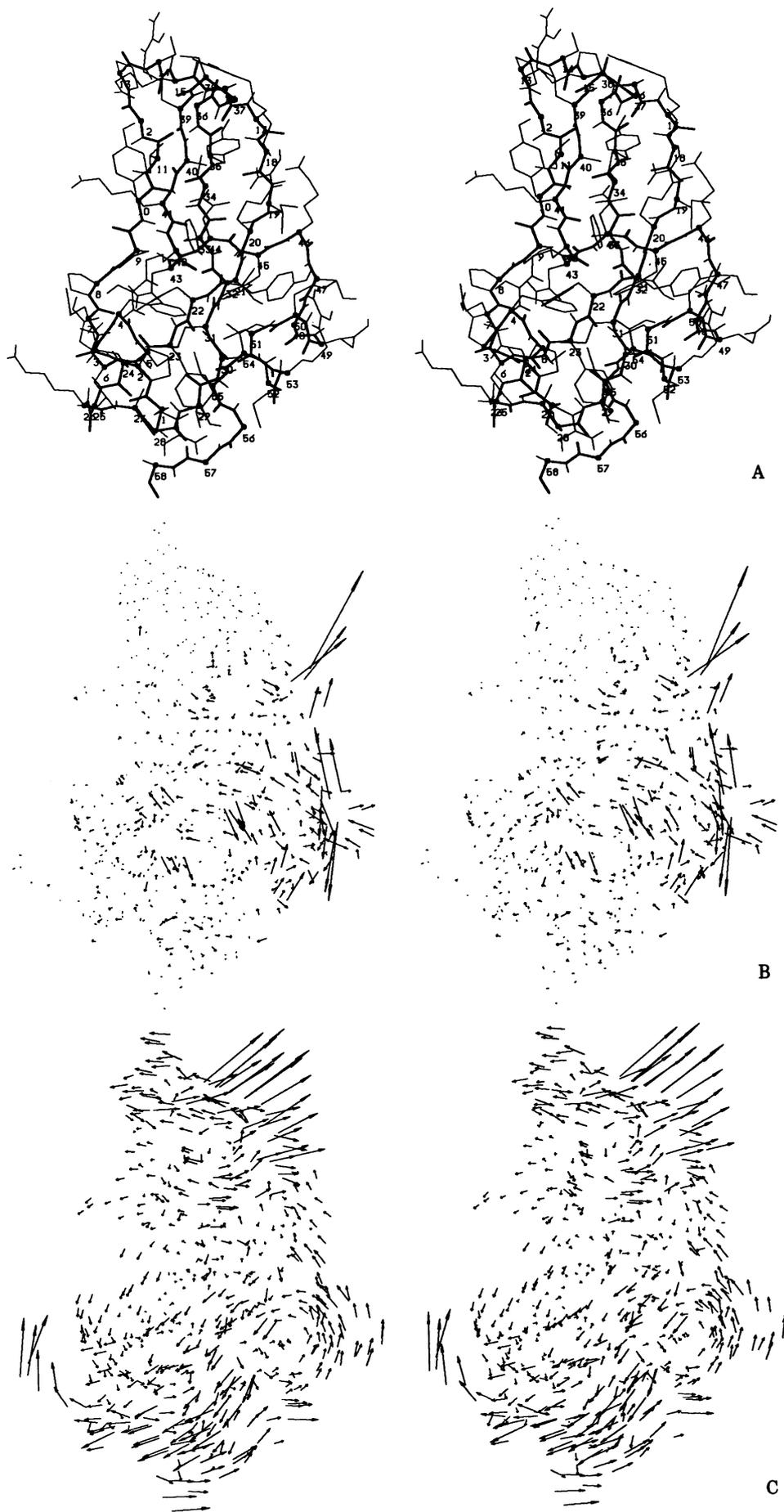
FIG. 2. Stereo drawings illustrating displacements of atoms in two normal modes of vibration in BPTI. (*A*) Main-chain (thick lines) and side-chain (thin lines) bonds in BPTI. The C^α atoms are numbered. (*B*) Displacement vectors of atoms in the 70th normal mode of vibration with frequency of 118.8 cm^{-1}. Root-mean-square displacements of atoms, calculated to occur when this mode is excited thermally at room temperature, are shown by vectors that are magnified 100 times for easy perception. Atom positions are aligned as in *A*. (*C*) Same as in *B* but for the 240th (the second from the lowest) normal mode, with frequency of 6.9 cm^{-1}. Magnification of vectors is 20 times.

Table 1. Degree of harmonicity of the normal modes in thermal fluctuations at room temperature

Frequency range, cm^{-1}	No. of normal modes							
	Value of $\langle(\Delta x)^2\rangle_c/\langle(\Delta x)^2\rangle_h$*							
	0.03–0.10	0.10–0.30	0.30–0.60	0.60–0.90	0.90–1.10	1.10–1.34	Others	Total
680–50			2	4	151	6	6†	169
50–45				2	6			8
45–40			1	5	6			12
40–35				4	3			7
35–30				8	1			9
30–25				8				8
25–20			2	7	1			10
20–15			3	4				7
15–10		1	2	3				6
10–5	1	2	2					5
Total	1	3	12	45	168	6	6	241

* Amplitude of fluctuations $(\Delta x)^2$ in each normal mode at room temperature is estimated by two methods: first, $\langle(\Delta x)^2\rangle_h$ from the second derivatives; second, $\langle(\Delta x)^2\rangle_c$ from a fourth-order polynomial that fits best with the cross section of the energy surface along the direction of a normal mode. Anharmonicity becomes strong as the ratio deviates far from unity. The 168 modes with the ratios between 0.90 and 1.10 are considered good harmonic modes. The amplitude of fluctuations $(\Delta x_i)^2$ in the ith normal mode is defined as $\Sigma_\mu (\Delta\theta_{i\mu})^2$.

† The energy curves for these six modes cannot be fitted well by fourth-order polynomials. Each of these modes corresponds to an isolated motion in one or two side chains on the surface of the molecule.

Mean-square displacements

Contribution from each normal vibrational mode to the mean-square displacements of atoms in the protein molecule can be calculated within the harmonic approximation by

$$\langle(\Delta \mathbf{r}_a)^2\rangle = \sum_{i,\mu\nu}\left(\frac{\partial \mathbf{r}_a}{\partial\theta_\mu}\cdot\frac{\partial \mathbf{r}_a}{\partial\theta_\mu}\right)\Delta\theta_{i\mu}\Delta\theta_{i\nu}\langle(\Delta x_i)^2\rangle_h, \qquad [4]$$

in which $\mathbf{r}_a$ is the position vector of the ath atom in the molecule, $\Delta \mathbf{r}_a$ is its derivation from the mean position, and $\langle(\Delta x_i)^2\rangle_h$ is the mean-square deviation of the conformational point along the direction of the ith normal mode calculated from the second-derivative matrix; explicit expressions for the derivatives on the right-hand side will be given elsewhere. Because of lack of knowledge, we neglect effects due to the higher-order cross terms between the very-low-frequency modes. The anharmonicities of the conformational energy surface along the directions of normal modes are such that the mean thermal fluctuations along such directions in the conformational space are generally smaller than the fluctuation expected from the second-derivative matrix (Table 1). In order to take these effects into account $\langle(\Delta x_i)^2\rangle_h$ are replaced in Eq. 4 by $\langle(\Delta x_i)^2\rangle_c$—i.e., by those estimated from the calculated energy curves.

Root-mean-square displacements averaged over main-chain atoms of each residue are plotted against residue number in Fig. 3. Because corrections for contributions from static lattice disorder are not made in the upper experimental curve, quantitative comparison cannot be made between the calculated and experimental curves. However, the qualitative agreement of the general trend is good. For example, regions of regular secondary structure—i.e., regions of two β-strands and an α-helix—have smaller root-mean-square displacements both experimentally and theoretically. A similar result was previously obtained by the method of molecular dynamics (10). The good qualitative agreement indicates soundness of the picture of the dynamic structure of the protein molecule based mainly on the second-derivative matrix at the minimum point. It also supports the usefulness of the description of the dynamic structure in terms of collective variables corresponding to the normal modes of vibration.

In a few regions in Fig. 3, especially in the region of residues 43–47, sharp peaks in the experimental curve are smoothed out in the calculated curve. These sharp peaks occur at residues that are well exposed on the surface of the molecule as judged from calculated values (not shown here) of the static accessible surface area (13, 14). Absence of these sharp peaks from the calculated curve suggests that these exposed residues fluctuate more than predicted from the shape of the energy surface around a single minimum point.

From the analysis of the mean square displacements calculated by Eq. 4, we have found that they are contributed mainly from very-low-frequency modes. The contributions to the mean square displacements of C$^\alpha$ atoms from low-frequency vibrational modes are shown in Fig. 4. We see very clearly that the very-low-frequency modes with wave number smaller than 30 cm^{-1} have dominant contributions.

As shown earlier, these modes generally involve nonlocalized motions of the molecule. Even though it is likely that these modes are coupled to each other nonlinearly, the concept of the collective variables corresponding to the very-low-frequency vibrational modes is expected to be useful. The dominant con-
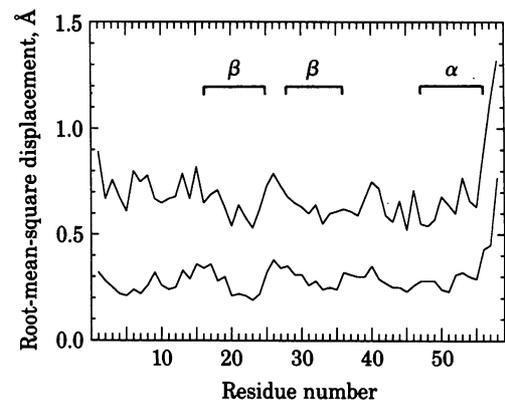
FIG. 3. Root-mean-square displacements of main-chain and C$^\beta$ atoms averaged over those of each residue are plotted against residue number. The lower curve is the result of the present calculation, which is based on the second-derivative matrix but is partially corrected for the anharmonicity observed mainly in the low-frequency modes. The upper curve is obtained from isotropic B factors determined x-ray-crystallographically by J. Deisenhofer for all individual nonhydrogen atoms (personal communication) without making corrections for static lattice disorder. Locations of secondary structures (β-sheet and α-helix) are indicated.
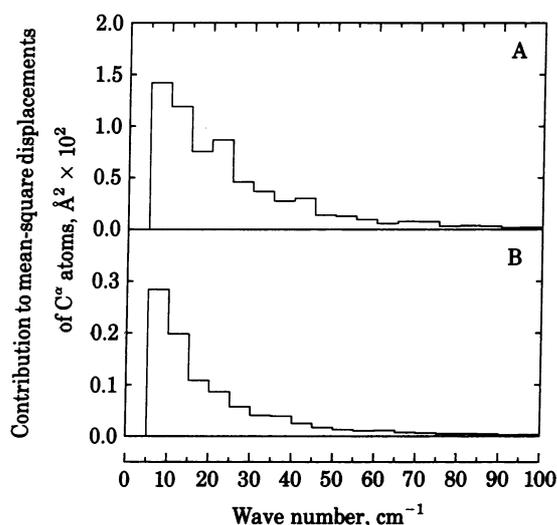
FIG. 4. (*A*) Contributions to the mean-square displacements of C^α atoms from all vibrational modes in each of the 5-cm^{-1} intervals in Fig. 1. (*B*) Mean contribution from each mode in each interval.

tribution mentioned above indicates that motions in these collective variables determine the basic dynamic structure of the protein.

Thermodynamics

The contribution to the thermodynamic quantities from motions arising from changes in dihedral angles can be calculated in the harmonic approximation as the sum of contributions from the 241 harmonic oscillators in Fig. 1. The calculated contributions from the harmonic oscillators to entropy and heat capacity of the molecule are 0.15 cal K^{-1} g^{-1} and 0.07 cal K^{-1} g^{-1}, respectively (1 cal = 4.18 J). These values are close to those obtained previously (8)—i.e., 0.11 cal K^{-1} g^{-1} for entropy and 0.06 cal K^{-1} g^{-1} for heat capacity, for the elastic protein model as contributions from such low-frequency vibrational motions in which a protein molecule behaves like a continuous elastic body. The value of entropy is about 35% larger in the present calculation based on the molecular structure. This is the result of the fact that the peak of the frequency spectrum density exists in a lower-frequency region than that estimated from the elastic protein model.

The value of entropy obtained in the present article accounts roughly for one-half of the entropy of solid dry "native" proteins measured by Hutchens *et al.* (15). The dihedral angles, which we are treating as variables in this article, constitute roughly 14% of all internal degrees of freedom in the protein molecule. The very-low-frequency modes with wave numbers smaller than 50 cm^{-1}, which constitute about 4.3% of all internal degrees of freedom, contribute about 22% of the entropy of the whole protein. These degrees of freedom, although a relatively small number, have a significant role in determining the thermodynamic functions of the protein molecule. The possible biochemical importance of variations in low-frequency vibrational entropy has been pointed out by Sturtevant (16).

Concluding remarks

Calculated normal modes of vibration in a small globular protein, BPTI, have been characterized. Reflecting the complexity of the protein structure and its dynamics, individual normal

modes have complex structures. However, once the normal modes are known, and if the conformational energy surface is harmonic within the range of thermal fluctuations at room temperature, the dynamics of a protein molecule can be described simply as a superposition of harmonic oscillations of collective variables corresponding to normal modes. If this were the case, the complexity could be confined to the structures of the normal modes and the dynamics should be basically simple. Analysis in this article indicates that most modes with frequencies above 50 cm^{-1} indeed behave harmonically at room temperature. This simplifies the dynamics of the protein structure significantly. Some anharmonicities are found in many modes with frequencies lower than 50 cm^{-1}. There are 72 modes between 5 cm^{-1} and 50 cm^{-1}. Magnitudes of displacements of atoms are mainly determined by modes with frequencies lower than 30 cm^{-1}. There are 36 modes in this range. It is also shown that 174 modes with frequencies lower than 120 cm^{-1} are generally nonlocal and continuous—i.e., the protein molecule behaves like a continuous elastic body in these low-frequency modes.

From the above results we obtain the following dynamic picture of the protein structure in the time range between 0.01 and 100 ps. The basic characteristics of the dynamic structure in the time range between 1 and 100 ps is determined by coupled nonlinear motions of collective variables corresponding to nonlocal vibrational modes with frequencies lower than 50 cm^{-1}. Motions in a large number of harmonic degrees of freedom with frequencies higher than 50 cm^{-1} are superimposed on the above nonlinear motions. Those degrees of freedom associated with bond length stretching and bond angle bending, which we did not treat as variables in this article, contribute mostly to harmonic degrees of freedom with frequencies higher than ≈120 cm^{-1}. They are also superimposed in the above nonlinear motions. It is interesting to see if this picture emerges from the trajectories generated by the method of molecular dynamics (17), whereby the effect of anharmonicity is automatically included. The nature of the nonlinear motions must be elucidated in the future.

1. Careri, G., Fasella, P. & Gratton, E. (1975) *CRC Crit. Rev. Biochem.* 3, 141–164.
2. Suezaki, Y. & Gō, N. (1975) *Int. J. Peptide Protein Res.* 7, 333–334.
3. Peticolas, W. L. (1978) *Methods Enzymol.* 61, 425–458.
4. Noguti, T. & Gō, N. (1982) *Nature (London)* 296, 776–778.
5. Disenhofer, J. & Steigemann, W. (1975) *Acta Crystallogr. Sect. B* 31, 238–250.
6. Beppu, T. & Ninomiya, I. (1981) *Comp. Phys. Com.* 23, 123–126.
7. Gō, N. (1978) *Biopolymers* 17, 1373–1379.
8. Gō, N. (1980) *Proc. Jpn. Acad.* 56, 414–419.
9. Tasumi, M., Takeuchi, H., Ataka, S., Dwivedi, A. M. & Krimm, S. (1982) *Biopolymers* 21, 711–714.
10. McCammon, J. A., Gelin, B. R. & Karplus, M. (1977) *Nature (London)* 267, 585–590.
11. Frauenfelder, H., Petsko, G. A. & Tsernoglou, D. (1979) *Nature (London)* 280, 558–563.
12. Mao, B., Pear, M. R., McCammon, J. A. & Northrup, S. H. (1982) *Biopolymers* 21, 1979–1989.
13. Lee, B. & Richards, F. M. (1971) *J. Mol. Biol.* 55, 379–400.
14. Shrake, A. & Rupley, J. A. (1973) *J. Mol. Biol.* 79, 351–371.
15. Hutchens, J. O., Cole, A. G. & Stout, J. W. (1969) *J. Biol. Chem.* 244, 26–32.
16. Sturtevant, J. M. (1977) *Proc. Natl Acad. Sci. USA* 74, 2236–2240.
17. McCammon, J. A. & Karplus, M. (1980) *Annu. Rev. Phys. Chem.* 31, 29–45.