Low-frequency vibrations of helical structures in protein molecules

Kuo-Chen CHOU
*Shanghai Institute of Biochemistry, Chinese Academy of Sciences, 320 Yo-Yang Road, Shanghai 200031,
China*

A physically intuitive and mathematically easily handled formula is presented for calculating the low-frequency vibrations of helical structures in protein molecules. α-Chymotrypsin is taken as an example, and the calculated result shows precise agreement with observations of the low-frequency Raman spectra. As reflected in the formula, this kind of low frequency is very sensitive to the conformation of a biomacromolecule, and can therefore serve as a vehicle to investigate the mechanism of action of a biomacromolecule from the viewpoint of dynamics. On this basis a feasible experiment is suggested by which one can examine the relationship between a presumed mode of low-frequency vibration in a biomacromolecule and its activity.

Among the component elements of protein molecules, the α-helix occupies a prominent position. This is because: (1) the regularity of α-helices makes them the only feature amenable to theoretical analysis; (2) the helical structure can have more influence on the stability and organization of a protein than any other individual structure; (3) as much as 80% of a structure can be helical, and only seven proteins are known that have no helix whatsoever (Richardson, 1981); (4) the flexibility or elasticity, one of the most significant characters of biomacromolecules (Koshland, 1958; Chou & Chen, 1977), is to a considerable degree related to their α-helical component (e.g. the α-keratin structure). In fact, the α-helix is regarded as the classic element of protein structure (Richardson, 1981).

Recently, there have been a number of experimental reports and theoretical analyses concerning the low-frequency motions in biomacromolecules. For instance, Brown *et al.* (1972) observed that low-frequency Raman bands (lower than 50cm^{-1}) exist in certain proteins, and these vibrations appeared to be sensitive to the conformation of a protein. Chou & Chen (1977) pointed out that the association of some biomacromolecules would concomitantly excite low-frequency phonons, otherwise some perplexing thermodynamic phenomena could not be explained self-consistently. Chou *et al.* (1981) further discussed the relationship between the co-operativity of an oligomeric protein and the effects of creating and annihilating low-frequency phonons, and described a physical picture concerning the microscopic mechanism of co-operative effects in biomacromolecules. Karplus & McCammon

(1979) found low-frequency fluctuations from a molecular-dynamics simulation of the bovine pancreatic trypsin inhibitor at the atomic level on a picosecond time scale. Sobell *et al.* (1979, 1982) deduced that low-frequency (acoustic) phonons can play an important role in DNA 'breathing' and drug intercalation. Zhou (1981) further investigated the low-frequency vibrational energy of a ring-like DNA and described some related biophysical phenomena. Careri *et al.* (1975) and Englander (1980), from another viewpoint, discussed the relationship between the hydrogen-exchange properties and this kind of internal motions in proteins and nucleic acids.

Now that it is accepted that low-frequency vibrations do exist in biomacromolecules and their functions have been widely discussed, a question is naturally raised: how do we calculate or predict this kind of low-frequency motions from a known molecular structure? In principle, the normal mode calculation method developed by Wilson (1939), Itoh & Shimanouchi (1970) and Fanconi *et al.* (1971) could serve such a purpose, but in practice this is computationally impossible, owing to lack of molecular symmetry and limitations on computer size and speed. To circumvent these difficulties, Fanconi & Peticolas (1971), following the procedure of normal mode calculations, suggested a highly simplified model, in which the α-helix consisted of only two masses per repeat unit, to compute the low-frequency motions. Because of the extreme complexity of biomacromolecules, it is advantageous and instructive to try to find some other route by which to approach this subject. Further-

more, it would be more useful for most molecular biologists if a physically more intuitive model and a mathematically more-easily handled formula could be supplied for calculating the low-frequency motions in biomacromolecules. The present study was initiated in an attempt to explore the ways of attaining such a goal. Realizing the prominent features cited above for the α-helical element of proteins, as a first step, let us consider the low-frequency motions of α-helices in protein molecules.

Continuity model

Recognizing the extreme complexity and inherent flexibility of biomacromolecules, rather than the discrete model suitable for the normal mode calculation method, I prefer to adopt a continuity model for consideration of the internal low-frequency motions in biomacromolecules. When discussing the high-frequency vibrations of a molecule, which refer to very small relative displacements and very strong molecular forces between neighbouring individual atoms, one has to resort to the discrete model. But for the low-frequency motions in a biomacromolecule, which involve much bigger effective masses and much weaker force constants (Chou & Chen, 1977), and whose modes can be compared with an 'accordion-like' motion (Brown et al., 1972), it is not only more convenient and intuitive but also physically rational to adopt the continuity model. In doing so, we will of course lose the information regarding the high-frequency motions in a biomacromolecule. Nevertheless, this is worthwhile, since the low-frequency motions in a biomacromolecule seem to be much more significant than the high-frequency motions (Careri et al., 1975; Chou & Chen, 1977; Englander, 1980; Chou et al., 1981; Sobell et al., 1982).

It is reasonable and instructive to liken an α-helix to a spring whose mass, however, is not negligible, namely a spring with distributed mass. On such an assumption, let us discuss the fundamental frequency (lowest frequency) of a spring system as illustrated in Fig. 1. Suppose O is the reference point of the vibration system. The two sides of this point will always move in opposite directions. Such a point actually divides the spring L into two parts, L_1 and L_2, whose force constants are assumed to be k_1 and k_2 respectively. According to the relation given in the Appendix, we have:

$$1/k = 1/k_1 + 1/k_2 \qquad (1)$$

where k is the force constant of the spring L. If σ_1 and σ_2 are the maximum stretch amounts of the springs L_1 and L_2 respectively, along the x-axis (Fig. 1), then according to the property of an even spring we obviously have:

$$\sigma_1/L_1 = \sigma_2/L_2 \qquad (2)$$

with:

$$L_1 + L_2 = L, \qquad \sigma_1 + \sigma_2 = \sigma \qquad (3)$$

where σ is the maximum amount of stretch of the spring L.

On the other hand, according to the force equilibrium of the system, we have:

$$K_1\sigma_1 = K_2\sigma_2 \qquad (4)$$

where K_1 and K_2 are the force constants of the two mass-negligible springs linked to the two ends of the spring L, as shown in Fig. 1.

From eqns. (1)–(4) it follows that:

$$\left. \begin{array}{ll} L_1 = \alpha_1 L, & L_2 = \alpha_2 L \\ \sigma_1 = \alpha_1\sigma, & \sigma_2 = \alpha_2\sigma \\ k_1 = k/\alpha_1, & k_2 = k/\alpha_2 \end{array} \right\} \qquad (5)$$
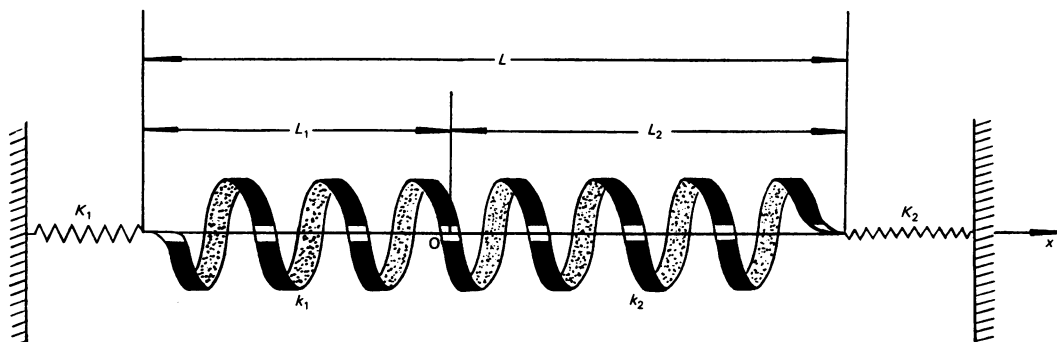

Fig. 1. *Vibration system in which a mass-distributed spring is linked to two other mass-negligible springs whose force constants are K_1 and K_2 respectively*

The point O is a reference point of the vibration system. The two parts of the system from this point move always in opposite directions.

where:

$$\alpha_1 = K_2/(K_1 + K_2), \qquad \alpha_2 = K_1/(K_1 + K_2) \quad (6)$$

The displacements of any points on L_1 and L_2 at any time can be described by

$$\left.\begin{array}{ll} u_1(x,t) = \dfrac{x}{L_1} \cdot \sigma_1 \cdot \sin \omega \cdot t & (-L_1 \le x \le 0) \\[2ex] u_2(x,t) = \dfrac{x}{L_2} \cdot \sigma_2 \cdot \sin \omega \cdot t & (0 \le x \le L_2) \end{array}\right\} \quad (7)$$

respectively, where ω is the round frequency. Suppose ρ is the mass per unit length of the mass-distributed spring; then the maximum kinetic energy of an element $\rho\Delta x$ of the spring L_1, a distance x_i from the point O, is:

$$\text{max. } (\Delta T_1{}^i) =$$

$$\frac{\rho\Delta x}{2} \cdot \text{max.} \left(\frac{du_1}{dt}\right)^2_{x_i} = \frac{\rho\Delta x}{2} \cdot \left(\frac{x_i}{L_1} \cdot \sigma_1\omega\right)^2 \quad (8)$$

The total maximum kinetic energy of the spring L_1 is thus:

$$\text{max. } T_1 =$$

$$\lim_{\Delta x \to 0}\left[\sum_i \text{max. } (\Delta T_1{}^i)\right] = \frac{\rho}{2} \cdot \frac{\sigma_1{}^2\omega^2}{L_1{}^2} \cdot \int_0^{L_1} x^2 \cdot dx \quad (9)$$

$$= \rho L_1 \omega^2 \sigma_1{}^2/6$$

Similarily, the total maximum kinetic energy of the spring L_2 is:

$$\text{max. } T_2 = \rho L_2 \omega^2 \sigma_2{}^2/6 \quad (10)$$

On the other hand, the total maximum potential energy of the whole spring system is:

$$\text{max. } U = (k_1 + K_1)\sigma_1{}^2/2 + (k_2 + K_2)\sigma_2{}^2/2 \quad (11)$$

According to energy conservation, i.e.:

$$\text{max. } T_1 + \text{max. } T_2 = \text{max. } U$$

we have:

$$\rho(L_1\sigma_1{}^2 + L_2\sigma_2{}^2)\omega^2/3 = (k_1 + K_1)\sigma_1{}^2$$
$$+ (k_2 + K_2)\sigma_2{}^2 \quad (12)$$

Substituting eqn. (5) into the above, we obtain:

$$\tilde{\nu} = \frac{\nu}{c} = \frac{\omega}{2\pi c} = \frac{1}{2\pi c} \cdot \sqrt{\frac{k + K^*}{(\alpha_1{}^3 + \alpha_2{}^3)\rho L/3}} \quad (13)$$

where ν is the fundamental frequency, $\tilde{\nu}$ the wave number, c the speed of light in vacuum and:

$$K^* = K_1 K_2/(K_1 + K_2) \quad (14)$$

Note that, if one end of the spring L is linked by a mass-negligible spring in the way illustrated in Fig. 2, then, instead of a stretching force constant, in eqn. (14) K_2 should be substituted by a corresponding bending force constant.

Stretching force constant of an α-helix

In this section, let us discuss the approach of calculating the stretching force constant of an α-helix. Obviously, for an α-helix in an 'accordion-like' vibration, its force constant is essentially related to the constituent hydrogen bonds. As is well known, the normal α-helix has 3.6 residues per turn, with a hydrogen bond between the CO group of residue n and the NH group of residue $n+4$ (Pauling et al., 1951), as illustrated in Fig. 3. Now, the problem is how to find the relation between the force constant of the whole α-helix and those of the constituent hydrogen bonds formed in such a mode.

For such a purpose, it would be helpful to compare an α-helix with a cylinder, then imagine that its flank is cut off along a straight line parallel to the helix axis, and is then flattened as shown in Fig. 4. As is well known, the hydrogen bonds in an α-helix are not precisely parallel to the helix axis, but there is a deviation angle, say θ, between them. Suppose the stretching and bending force constants of a hydrogen bond are k_s and k_B respectively. Then the force constant for such a hydrogen bond stretching along the helix axis should be:

$$k^* = \sqrt{(k_s \cdot \cos\theta)^2 + (k_B \cdot \sin\theta)^2} \quad (15)$$

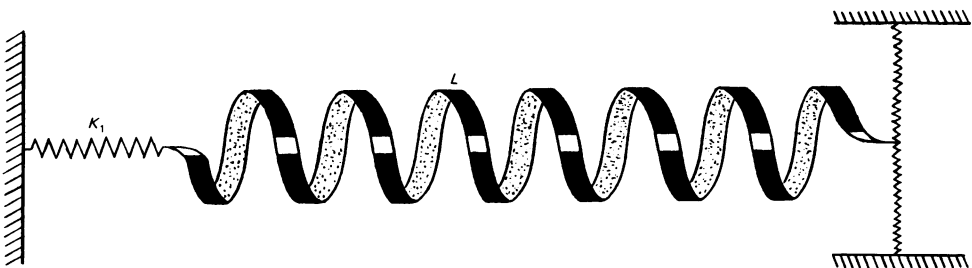According to the relation given in the Appendix, the

Fig. 2. Vibration system in which one end of the mass-distributed spring L is linked to a mass-negligible spring in the same way as in Fig. 1, but the other end is linked to a mass-neglibible spring in a mutual perpendicular mode
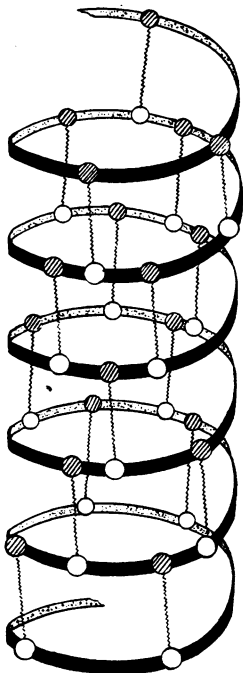
Fig. 3. *Illustration of an α-helix*
O, Peptide O atom; ⊘, peptide N atom; ∿∿∿,
hydrogen bond.

resultant force constant k for the spring system shown in Fig. 4 can thus be expressed as:

$$k = \frac{1}{\dfrac{1}{4k^*} + \dfrac{1}{4k^*} + \dfrac{1}{3k^*} + \dfrac{1}{4k^*} + \dfrac{1}{3k^*}} \qquad (16)$$

Following such a method, for any α-helix, we can always derive the approximate expression of its force constant in terms of the force constants of the individual hydrogen bonds, e.g. for an α-helix with eleven amino acid residues the corresponding force constant k is:

$$k = \frac{1}{\dfrac{1}{4k^*} + \dfrac{1}{3k^*}} = \frac{12}{7} \cdot k^* \qquad (17)$$

Note that the number of the hydrogen bonds in an α-helix is normally $n-4$, where n is the number of the amino acid residues in the α-helix.

Example and discussion

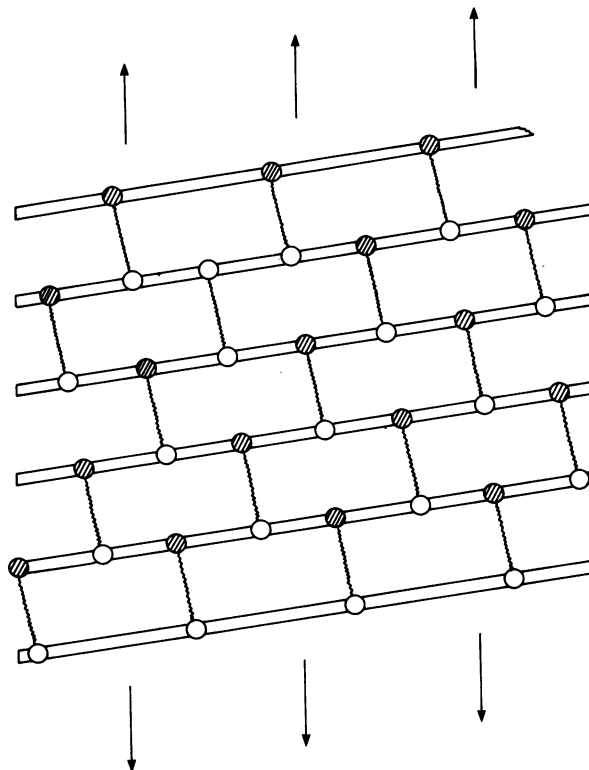According to a report presented by Brown *et al.* (1972), low-frequency Raman spectra were obtained

Fig. 4. *Illustration obtained by cutting and flattening the flank of the α-helix in Fig. 3*

from samples of α-chymotrypsin that had been prepared in several ways. A peak at about 29 cm⁻¹ was found for all samples except the one that had been denatured with sodium dodecyl sulphate. They further pointed out that this kind of vibration 'appears to be sensitive to the conformation of the protein' and 'is of considerable interest even though the exact assignment of the motion is not possible at present'.

Now, let us try to reveal the physical mechanism of the low-frequency vibration appearing in α-chymotrypsin. α-Chymotrypsin consists of 245 residues. A view of its complete polypeptide chain is outlined in Fig. 5, where residues 57, 102 and 195 are the components of the active site. The β-barrel 1, formed by six adjacent anti-parallel chains along the

sequence from residue 29 to 112, appears at the upper left, and the β-barrel 2, also formed by six anti-parallel chains between residues 133 to 230, is at the lower right. The C-terminal α-helix (residues 235–245) forms a tail on the second barrel and makes hydrophobic contacts with both barrels, and the short contiguous helix (residues 230–235) is almost inside the molecule. Their axes make an angle of about 40°. The only hydrogen bond between the C-terminal helix and the remainder of the molecule is that made by the terminal carboxy group with Lys-107.

For the following reasons, the whole α-chymotrypsin molecule can be treated as a vibration system as illustrated in Fig. 2: (1) the C-terminal α-helix is on the enzyme surface; (2) although one end of the
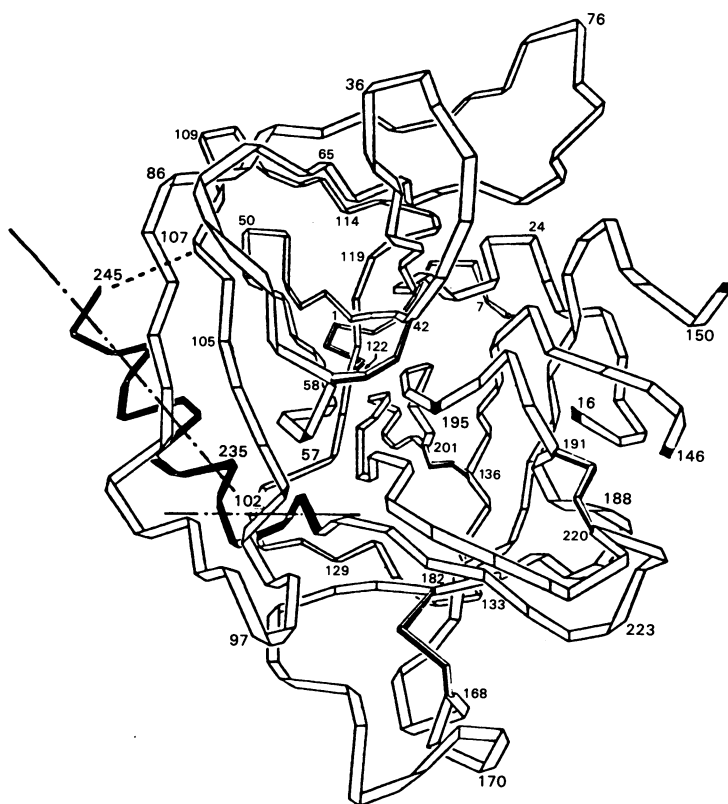

Fig. 5. *View of the α-chymotrypsin molecule*

The polypeptide chain is represented by a ribbon folded at each α-carbon atom, and the part thereof with helical structure is specially marked in black. The C-terminal α-helix (residues 235–245) is on the enzyme surface, and the short contiguous helix (residues 230–235) is almost buried inside the molecule. Their axes (indicated by —·—·) make an angle of about 40°. β-Barrel 1 (residues 29–122) appears at the upper left of the Figure (in front of the C-terminal helix), and β-barrel 2 (residues 133–230) is at the lower right. It looks as if that the C-terminal helix forms a tail on the second cylinder (barrel) and makes hydrophobic contacts with both cylinders (β-barrel 1 and β-barrel 2). The only hydrogen bond between C-terminal helix and the remainder of the molecule is that made by the terminal carboxyl group with Lys-107 (illustrated by ------), holding the C-terminal helix more firmly to the enzyme surface. Residues 57, 102, and 195 are the components of the active site. Disulphide linkage are indicated by hatched bars.

C-terminal helix is linked to a short contiguous helix, the latter is buried inside the molecule, and therefore this end is actually linked by a covalent bond to an object whose mass is much bigger than that of the C-terminal helix; (3) the other end of the helix is fixed by a salt bridge to Lys-107, the only hydrogen bond between the C-terminal helix and the remainder of the molecule, and such a hydrogen bond is almost

(Itoh & Shimanouchi, 1970), and therefore eqns. (6) and (14) can be reduced to:

$$\alpha_1 = 0, \qquad \alpha_2 = 1 \tag{20}$$

and

$$K^* = k_B = 0.015 \times 10^5 \, \text{dyn/cm} \tag{21}$$

respectively.

Substituting eqns. (17)–(21) into eqn. (13), we obtain:

$$\tilde{\nu} = \frac{1}{2\pi \times 3 \times 10^{10}} \cdot \sqrt{\frac{\left(\dfrac{12}{7} \times 0.117 + 0.015\right) \times 10^5 \times 6.02 \times 10^{23}}{\dfrac{1242}{3}}} \tag{22}$$

$$= 29.7 \, \text{cm}^{-1}$$

perpendicular to the C-terminal helix axis; (4) both the covalent bond and the hydrogen bond at the two ends of the C-terminal helix can be regarded as mass-negligible springs.

The amino acid sequence of the C-terminal helix is (Birktoft & Blow, 1972):

Val-Asn-Trp-Val-Gln-Gln-Thr-Leu-Ala-Ala-Asn

whose mass can be easily calculated by adding the masses of all the residues together; the sum comes out at 1242 a.m.u. (atomic mass units). To an extent of fair approximation, we can suppose that ρ, the mass per unit length of the helix, is even. We thus have:

$$\rho \cdot L = 1242 \, \text{a.m.u.} = \frac{1242}{N} \, \text{g} \tag{18}$$

where L is the length of the helix and N is the Avogardro constant.

The stretching force constant of an α-helix can be calculated by means of the method presented in the preceding section. For a helix with eleven residues, its stretching force constant has already given as in eqn. (17). The stretching force constant, k_S, of the hydrogen bond is 0.130×10^5 dyn/cm, and its bending force constant k_B is 0.015×10^5 dyn/cm (Itoh & Shimanouchi, 1970). [Note that the bending point in our case is at the end of the hydrogen bond, rather than the middle of the bond as considered by Itoh & Shimanouchi (1970). So the force constant of the former should be half of the latter.] The deviation angle θ of the hydrogen bonds from the helix axis is about $26°$. Then, according to eqn. (15), we have

$$k^* = \sqrt{(0.130 \cos 26°)^2 + (0.015 \sin 26°)^2} \times 10^5$$

$$= 0.117 \times 10^5 \, \text{dyn/cm} \tag{19}$$

As is well known, the force constant of a covalent bond is much bigger than that of a hydrogen bond

which is exactly the same as the result observed by Brown et al. (1972).

It should be pointed out that in the above calculation the 'frictional' interactions of the side chains on the C-terminal helix with the rest of the protein were ignored. However, according to the classical analysis for a simple vibration system, such an approximate treatment is warranted if the 'frictional coefficient' due to this kind of interaction is much smaller than the product of the mass of the oscillator and its natural frequency. And here this condition would most probably hold, because the C-terminal helix is situated on the enzyme surface. Besides, as is well known, an oscillator with frequency ν will excite the phonons with energy of $h\nu$. Since phonons are bosons (Chou & Chen, 1977), under thermal equilibrium, the mean number of phonons thus excited is, according to Bose–Einstein statistics, given by:

$$\langle n \rangle = \frac{1}{e^{h\nu/k_B T} - 1} \tag{23}$$

where k_B is the Boltzmann constant. On the other hand, under general temperature, for phonons with low frequencies ($\sim 30 \, \text{cm}^{-1}$), we obviously have $h\nu \ll k_B T$. Consequently, the average vibration energy of the oscillator is:

$$\langle E \rangle = \langle n \rangle h\nu = \left. \frac{h\nu}{e^{h\nu/k_B T} - 1} \right|_{h\nu \ll k_B T} \simeq k_B T \tag{24}$$

This indicates that the C-terminal helix undergoing such a low-frequency motion possesses the same magnitude of energy as that excited by thermal movement, and hence the loss of its energy due to the 'friction' of its side chains with the rest of the protein can be easily compensated through Brownian collisions with the solvent molecules around the enzyme molecule.

Besides, from eqn. (11), in terms of eqns. (5), (6) and (14), as well as the energy conservation principle, we have the amplitude given by:

$$\sigma = \sqrt{\frac{2<E>}{k+K^*}} \Bigg|_{h\nu \ll k_B T} \simeq \sqrt{\frac{2k_B T}{k+K^*}} \quad (25)$$

When $T = 300\,\text{K}$, substituting eqns. (17), (19) and (21) into eqn. (25), we obtain $\sigma \simeq 0.02\,\text{nm}$ (0.2 Å), which is in the same order of magnitude as the protein structural fluctuations obtained through the molecular-dynamics simulation by Karplus & McCammon (1979).

Native conformation and activating low frequency

When the α-chymotrypsin was denatured with sodium dodecyl sulphate, the low-frequency peak at $29\,\text{cm}^{-1}$, which is closely related to the micro-environment of the C-terminal helix as well as its conformation as illustrated in the aforementioned calculation, will of course vanish, as a result of a dramatic change on the secondary and tertiary structure of the protein. Instead, 'rather intense Raman scattering throughout the region of 20–$150\,\text{cm}^{-1}$ is observed on the denatured material, but it is broad and structureless' (Brown et al., 1972). This apparently reflects the increase of the background noise and the decrease in order of the protein conformation.

From this example, we can see how the activity of the enzyme as well as its conformation are remarkably related to some characteristic low frequency. Of course, with such a change in protein conformation, all the other portions of the corresponding molecular spectra will change as well. Nevertheless, considering the exceptional biological functions of low-frequency motions in biomacromolecules (Careri et al., 1975; Chou & Chen, 1977; Englander, 1980; Chou et al., 1981; Sobel et al., 1982), it is instructive to single out this kind of characteristic low frequency, and term it the 'activating low frequency' (Chou et al., 1981). The introduction of such a concept will no doubt be useful for further investigation into the action mechanisms of biomacromolecules from the viewpoint of dynamics.

According to the above calculation and discussion, it is intriguing to devise an experiment as follows. Cut off the C-terminal helix, the 'tail' of α-chymotrypsin, by some chemical method, and investigate what happens in both its activity and low-frequency spectrum. Through such an approach, the idea of activating low-frequency phonons and the presumed vibration mode can be inspected.

References

Birktoft, J. J. & Blow, D. M. (1972) J. Mol. Biol. 68, 187–240

Brown, K. G., Erfurth, S. C., Small, E. W. & Peticolas, W. L. L. (1972) Proc. Natl. Acad. Sci. U.S.A. 69, 1467–1469

Careri, G., Fasella, P. & Gratton, E. (1975) CRC Crit. Rev. Biochem. 3, 141–164

Chou, K. C. & Chen, N. Y. (1977) Sci. Sin. 20, 447–457

Chou, K. C., Chen, N. Y. & Forsén, S. (1981) Chem. Scr. 18, 126–132

Englander, S. W. (1980) Comments Mol. Cell. Biophys. 1, 15–28

Fanconi, B. & Peticolas, W. L. (1971) Biopolymers 10, 2223–2229

Fanconi, B., Small, E. W. & Peticolas, W. L. (1971) Biopolymers 10, 1277–1298

Itoh, K. & Shimanouchi, T. (1970) Biopolymers 9, 383–399

Karplus, M. & McCammon, J. A. (1979) Nature (London) 277, 578

Koshland, D. E., Jr. (1958) Proc. Natl. Acad. Sci. U.S.A. 44, 98–104

Pauling, L., Corey, R. B. & Branson, H. R. (1951) Proc. Natl. Acad. Sci. U.S.A. 37, 205–211

Richardson, J. S. (1981) Adv. Protein Chem. 34, 167–339

Sobell, H. M., Lozanksy, E. D. & Lessen, M. (1979) Cold Spring Harb. Symp. Quant. Biol. 43, 11–19

Sobell, H. M., Lobansky, E. D., Zhou, G. P. & Chou, K. C. (1982) J. Mol. Sci. in the press

Wilson, E. B. (1939) J. Chem. Phys. 7, 1047–1052

Zhou, G. P. (1981) Sheng Wu Hua Hsueh Yu Sheng Wu Wu Li Chin Chan no. 5, 19–22

APPENDIX

Consider the force constant of the spring system shown in Appendix Fig. 1. On application of a unit force at its right end, each constituent spring will
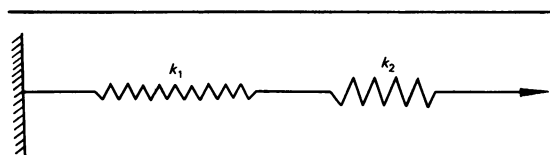

Fig. 1. Spring system formed by a series connection of two individual springs

stretch by an amount $1/k_1$ and $1/k_2$, and the total displacement of the end becomes:

$$\Delta L = \frac{1}{k_1} + \frac{1}{k_2} \quad (A\text{-}1)$$

By definition, the resultant force constant for the system should be:

$$k = \frac{1}{\Delta L} = \frac{1}{\dfrac{1}{k_1} + \dfrac{1}{k_2}} \quad (A\text{-}2)$$

Generally speaking, for a system consisting of n springs in series connection, the resultant force constant k can be written as:

$$k = \cfrac{1}{\cfrac{1}{k_1} + \cfrac{1}{k_2} + \cdots + \cfrac{1}{k_n}} \qquad \text{(A-3)}$$

where $k_1, k_2, \ldots, k_n$ are the force constants of the n individual springs respectively.

However, if a spring system consists of n identical springs in parallel connection, as illustrated in Appendix Fig. 2, a similar derivation will result in:

$$k = n \cdot k_0 \qquad \text{(A-4)}$$

From the above we see that there is actually a crosswise corresponding relation between the formulae for calculating the resultant spring force
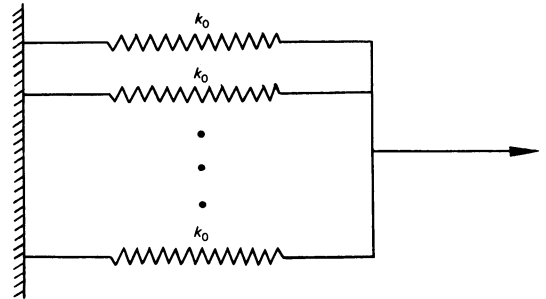

Fig. 2. *Spring system formed by a parallel connection of n identical springs*

constant and those for calculating the resultant resistance as far as the series connection and parallel connection are concerned.