Shape of the conformational energy surface near the global minimum and low-frequency vibrations in the native conformation of globular proteins

ARTICLE *in* BIOPOLYMERS · JANUARY 2004

Impact Factor: 2.39 · DOI: 10.1002/bip.1978.360170520

1 AUTHOR:

Nobuhiro Go
Japan Atomic Energy Agency
238 PUBLICATIONS **9,961** CITATIONS

Shape of the Conformational Energy Surface near the Global Minimum and Low-Frequency Vibrations in the Native Conformation of Globular Proteins

NOBUHIRO GŌ, *Department of Physics, Faculty of Science, Kyushu University, Fukuoka 812, Japan*

Synopsis

Based on the assumption that the conformational energy surface of a protein molecule can be approximated near the global minimum point by a multidimensional parabola, conformational fluctuations in the native state are discussed. In this approximation the conformational fluctuations can be viewed as excitations of coupled harmonic oscillations of dihedral angles. For the purpose of estimating the range of frequencies of such vibrations, globular proteins are assumed to be made of homogeneous continuous elastic material. The number of vibrational modes in such an elastic body, with the wavelength no less than the characteristic length of an amino acid residue, are estimated roughly to be three times the number of amino acid residues in a protein, which is slightly less than the number of variable dihedral angles in a protein. Their frequencies, when converted to the wavenumber of corresponding light, are found to range from 1.8×10 cm^{-1} to 2.1×10^2 cm^{-1} for a protein with diameter $d = 40$ Å, when Young's modulus $E = 10^{11}$ dyne/cm^2 is assumed. A significant fraction of the coupled vibrations of dihedral angles in real globular proteins are collective ones, i.e., those involving the whole protein molecule. Based on these results, it is concluded that the depth of the global minimum is at least 150 kcal/mol.

INTRODUCTION

Experimentally observed, spontaneous folding of proteins has been interpreted to imply that the native conformation of a protein corresponds to the local minimum of the conformational energy which has the lowest free energy.[1] Let us call this local minimum simply the global minimum. Many theoretical attempts to predict the native conformation of a protein from its amino acid sequence are based, directly or indirectly, on this correspondence between the native conformation and the global minimum. The naive strategy of minimizing the conformational energy by starting from arbitrarily chosen points in the conformational space cannot attain the final goal of determining the global minimum because of the existence of an enormous number of local minima in the multidimensional conformational energy surface. It is, however, interesting to have an idea about the general shape of the energy surface near the global minimum. Energy minimization within the well containing the global minimum point has been used to refine x-ray coordinates of proteins. How much reduction of the conformational energy is expected depends on how deep the well is.

Biopolymers, Vol. 17, 1373–1379 (1978)

0006-3525/78/0017-1373$01.00

In ordinary calculations of conformational energy, dihedral angles around rotatable single bonds are taken as independent variables. Each point in the multidimensional conformational space corresponds to a microscopic conformational state specified by a set of corresponding values of dihedral angles. The native state of a protein is rather a macroscopic concept corresponding to a collection of microscopic conformational states, which correspond to the global minimum point and the neighboring points around it. One can alternatively view the native state of a protein as the microscopic conformational state of a single protein molecule fluctuating within the well of the conformational energy surface containing the global minimum. The nature of this fluctuation is governed by the shape of the well.

In the present paper it is assumed that the conformational energy surface can be approximated by a multidimensional parabola in the range of conformational fluctuations in the native state. Based on this assumption, it is pointed out that the conformational fluctuations in the native state can be viewed as excitation of coupled normal vibrational modes of dihedral angles. It is shown that a significant fraction of these vibrational modes are collective. Here, vibrational modes are called collective when vibrations are not localized in a certain part of a protein molecule but involve the whole protein molecule.

The depth of the well containing the global minimum is then estimated. The estimation is again based on the parabolic approximation. Energy reduction in the course of refinement of x-ray coordinates of a protein is discussed from this point of view.

COLLECTIVE MODES OF FLUCTUATION OF DIHEDRAL ANGLES

Discussions in the present paper will be based on the assumption that the conformational energy surface near the global minimum point can be approximated by a multidimensional parabola in the range of conformational fluctuations. No clear evidence to support this assumption seems to exist. The attractive, but not yet established, migrating defect model of Nakanishi et al.[2] and of Lumry,[3] which was assumed primarily to explain the low-activation-enthalpy process for hydrogen exchange in native proteins, seems possible only when the conformational energy surface near the global minimum point is more complex than a simple multidimensional parabola. Therefore, it is not an intention of this paper to assert the validity of the parabolic assumption, but to discuss a few consequences of the assumption. We believe that this assumption is warranted for the present because no clear counter-evidence exists.

A multidimensional parabola is characterized by a second derivative matrix, **F**, at the minimum point. Dihedral angles, which describe the conformation of a molecule, fluctuate within the well of this multidimensional parabola. The second derivative matrices, **F**, were actually

calculated for α-helices of a few homopolypeptides[4–6] in order to calculate the Zimm-Bragg parameters, s and σ, for the helix-coil transition. For this purpose, the free energy of an α-helix was calculated. It is essential in this calculation to take into account conformational fluctuations of an α-helix, determined by the second derivative matrix, **F**. Later, it was shown that the inverse matrix of **F** is proportional to the correlation of fluctuations in dihedral angles.[7] Then, by using this property, geometrical fluctuations of α-helices were calculated.[8] These geometrical fluctuations were compared with those of a rod made of homogeneous continuous material. Mechanical properties of the rod determine the latter. From the comparison, mechanical properties of rods equivalent to the α-helices were determined.[8] What will be discussed in the present paper on globular proteins are, in a sense, extensions of the ideas and calculations developed previously for α-helices.

Let us denote a set of dihedral angles, which are treated as independent variables to describe conformations of a protein molecule, by $Q = (q_1, q_2, \ldots, q_m)$, where m is the number of dihedral angles treated as independent variables. Once the energy surface, $F = F(Q)$, is approximated by a multidimensional parabola, i.e., by a quadratic expression

$$F = F(Q_0) + \frac{1}{2}\sum_{i,j=1}^{m} f_{ij}(q_i - q_{i,0})(q_j - q_{j,0}) \tag{1}$$

around the global minimum point, $Q = Q_0$, the system is formally equivalent to a collection of harmonic oscillators. Here, f_{ij} is an element of the second derivative matrix, **F**. In order to determine frequencies of the harmonic oscillators, we need an expression of the kinetic energy of a protein molecule. It is a quadratic function of generalized momenta $P_r = (p_1, p_2, \ldots, p_m)$ conjugate to the independent dihedral angles Q:

$$K_r = \frac{1}{2}\sum_{i,j=1}^{m} g_{ij}p_i p_j \tag{2}$$

Here, suffix r indicates that only dihedral angles, not bond lengths and bond angles, are treated as variables. This equation is the same as Eq. (9) of the recent paper by Gō and Scheraga.[9] An expression for the coefficient g_{ij} is given there. Eigenvalues of the product **FG** of the two matrices **F** and **G**, where **G** is a matrix whose i,j element is g_{ij} in Eq. (2), give $(2\pi\nu_i)^2$, $i = 1, 2, \ldots, m$, where the ν_i are the frequencies of the normal harmonic oscillators. Conformational fluctuations of a native protein within the parabolic well of the conformational energy surface can thus be viewed as excitation of these normal vibrational modes.

It is, however, not practical actually to calculate the two matrices **F** and **G** for a whole protein and then to determine the eigenvalues of the product **FG**. A more fundamental difficulty than this practical one may exist in calculating the frequencies ν_i from matrices **F** and **G**. It has been formulated in the preceding paragraph that dihedral angles couple each other to form normal vibrational modes. However, other degrees of freedom

belonging to (1) solvent molecules and (2) bond lengths and bond angles in a protein molecule may also be involved in these vibrational modes. If this is the case, the frequencies ν_i can not be determined accurately in a treatment in which only dihedral angles are independent variables. The purpose of the preceding paragraph is only to show that dihedral angles couple eachother to form normal modes, but not to advocate using matrices **F** and **G** for calculating the frequencies.

Because it is difficult to calculate the frequencies ν_i from real molecular structures of proteins, we will take the steps of the calculations in reverse order in this paper from that used for α-helices. In α-helices, geometrical fluctuations were calculated from the second derivative matrix, **F**, and then mechanical properties were determined from the geometrical fluctuations. In this paper we assume, as in an earlier paper,[10] that a protein molecule is an elastic body made of continuous material of known mechanical properties. Then, we will calculate the frequencies of vibrations of this elastic body. By assuming reasonable geometrical dimensions and mechanical properties, calculated frequencies of vibrations of the elastic body are expected to give us a qualitatively correct picture about the vibrational modes in real proteins.

Globular proteins would behave as continuous elastic bodies for low-frequency motions whose wavelengths are longer than the size of individual residues.[10] Let us consider the number of such vibrational modes. For this purpose we assume that a protein molecule is an elastic cuboid with three mutually perpendicular edges of lengths L_x, L_y, and L_z. For the purpose of simplifying the analysis, we assume that such elastic cuboids are arranged regularly to pack the space. Thus, we can apply the periodic boundary condition[11] to the discussion of elastic waves. Then, we count the number of plane-wave solutions consistent with the periodic boundary condition and the x, y, and z components of the wavenumber vectors up to π/r, where r is the characteristic length of a residue defined by $v^{1/3}$, with v being an average volume of a residue. By the standard method,[11] this number is calculated to be $3L_xL_yL_z/r^3$ or $3n_a$, where n_a is the number of amino acid residues in a protein molecule. The factor 3 comes from the fact that there are three types (one longitudinal and two transverse) of plane waves in an infinite elastic body. If we remove the periodic boundary condition and treat each cuboid isolated in space, different plane-wave solutions mix together to form a normal vibrational mode because of the free boundary condition. However, mixing occurs mainly between plane waves whose frequencies are close to each other. And in this mixing the number of plane-wave solutions is conserved as the number of normal vibrational modes in an isolated cuboid. Therefore, in a rough calculation, the number of normal vibrational modes and the range of distribution of their frequencies can be estimated by artificially applying the periodic boundary condition.

Because the number of vibrational modes whose wavelength is longer than the size of a residue is not expected to be a strong function of its shape,

the conclusion obtained in the preceding paragraph, namely, that the number is roughly $3n_a$, is expected to hold for more general conformations of globular proteins.

Let us then estimate the range of frequencies of these vibrational modes. The frequencies are given by c/λ, when c is the velocity of the elastic wave and λ is its wavelength. As maximum and minimum values of λ we should take the diameter of a protein, d, and the characteristic length of a residue, r, respectively. The velocities of the transverse and longitudinal elastic waves are given by $c_t = [E/2\rho(1 + \sigma)]^{1/2}$ and $c_l = [E(1 - \sigma)/\rho(1 + \sigma)(1 - 2\sigma)]^{1/2}$, respectively.[12] Here, E is Young's modulus, σ is Poisson's ratio, and ρ is mass density. Poisson's ratio assumes a value between -1 and $\frac{1}{2}$. If we assume, as before,[10] $E = 10^{11}$ dyne/cm^2, $\rho = 1$ g/cm^3, and take arbitrarily $\sigma = 0$, then, we have $c_t = 2.2 \times 10^5$ cm sec^{-1}, and $c_l = 3.2 \times 10^5$ cm sec^{-1}. As d and r, let us take values of 40 and 5 Å, respectively. Then, as the range of frequencies, we have $\nu = 5.5 \times 10^{11}$ sec^{-1} to 6.4×10^{12} sec^{-1}. By converting to wavelengths of corresponding light, we have the range of wavelengths of $(1/\lambda) = 1.8 \times 10$ cm^{-1} to 2.1×10^2 cm^{-1}. The above range of frequencies corresponds to $x = 0.09$–1.0, where x is the ratio of the energy quantum of vibration, $h\nu$, to the thermal energy, kT, with $T = 300°$K. This means that most of these vibrations can be treated classically for usual thermodynamic discussions.

In the parabolic approximation of the energy surface, variable dihedral angles couple to each other to form normal vibrational modes. They range from low-frequency collective modes that involve the whole protein molecule to localized modes involving one or a few variable dihedral angles. The number of variable dihedral angles in a protein molecule is only slightly larger than $3n_a$, the number of vibrational modes in an elastic protein model. Therefore, most of the normal modes of coupled vibrations of variable dihedral angles are those in which the protein molecule behaves as a continuous elastic body. Mechanical inhomogeneity within a protein molecule, neglected in the analysis of the present paper, tends to make vibrational modes with higher frequencies more localized than those in a homogeneous elastic protein model. Yet, a significant fraction of the normal modes should still be collective ones, i.e., those involving the whole protein molecule.

Brown et al.[13] reported the existence of a peak at the low-frequency region around 30 cm^{-1} in the laser Raman spectra from the native α-chymotrypsin and pepsin. The peak disappears when the proteins are denatured. It is likely that the experimentally observed peak corresponds to the vibrational modes in which a protein molecule behaves as an elastic body. In the previous paper[10] we attributed the width of the experimentally observed peak to the damping of a vibrational mode. However, the analysis in the present paper indicates that distribution of many vibrational modes in this frequency range should also be considered.

DEPTH OF THE GLOBAL MINIMUM

Let us calculate the mean energy of conformational fluctuations in the native state of a protein. We again assume that the conformational energy surface near the global minimum point can be approximated by a multidimensional parabola in the range of conformational fluctuations. Then, the mean excitation of the conformational energy above the minimum value can be easily calculated as follows:

$$\langle F - F(Q_0) \rangle = -\frac{\partial}{\partial \beta} \ln \int \exp\left(-\frac{\beta}{2} \sum_{i,j=1}^{m} f_{ij}(q_i - q_{i0})(q_j - q_{j0})\right) = mkT/2 \quad (3)$$

In this calculation, β is used for $1/kT$. This result is in a sense obvious from what was concluded in the preceding section, i.e., conformational fluctuations in the native state of a protein can be described by a set of m harmonic oscillators, most of which are in the classical frequency range, because mean potential energy of a single classical harmonic oscillator is $kT/2$.

By assuming $m \simeq 3n_a \simeq 500$, we have $\langle F - F(Q_0) \rangle \simeq 150$ kcal/mol at room temperatures. This means that the depth of the global minimum is at least 150 kcal/mol for globular proteins whose native conformations are stable at room temperatures. If there is a rim with conformational energy lower than 150 kcal/mol of the minimum value, the conformational state of the protein does not stay within the well of this minimum because of the thermal fluctuations.

This value of the depth of the global minimum is relevant to how much reduction of the conformational energy is expected in a process of refinement of x-ray coordinates of a protein. In the usual process of refinement, unreasonable bond lengths, bond angles, and steric overlaps are first eliminated. Energy reduction associated with these eliminations is large. The resulting conformation may be regarded as one of instantaneous fluctuating conformations within the well of the global minimum. Because the average energy of such fluctuating conformations is roughly 150 kcal/mol, we can still expect a reduction of the conformational energy of this order in a further refinement after the elimination of unreasonable bond lengths, bond angles, and steric overlaps. This contribution should come mainly from such soft energy terms as the attractive parts of nonbonded interactions and the rotational potential energies. In fact, in the refinement of atomic coordinates of lysozyme by Warme and Scheraga,[14] a reduction of 167 kcal/mol is attained in the final stage of refinement. Severe steric overlaps are already removed in the previous stage. Of 167 kcal/mol, contributions from nonbonded and rotational energy terms are 81 kcal/mol and 63 kcal/mol, respectively, with the rest being from other terms. It should be noted here that any conformation in the process of refinement serves as an example of an instantaneous fluctuating state, if its conformational energy is within 150 kcal/mol of the global minmum value. It is interesting to obtain an idea about the amplitude of fluctuations in the dihedral angles from the process of refinement of coordinates by the use of the conformational energy.

CONCLUSION

Small-amplitude conformational fluctuations in native globular proteins are discussed from two different points of view. If it is assumed that the conformational energy surface near the global minimum point can be approximated by a multidimensional parabola, the conformational fluctuations can be viewed as excitation of coupled normal vibrational modes of variable dihedral angles. If we regard a globular protein as an approximate elastic body made of continuous elastic material, the conformational fluctuations can be viewed as excitation of normal elastic vibrational modes. It is concluded that a significant fraction of the coupled vibrational modes of variable dihedral angles is made up of collective ones, i.e., those involving the whole protein molecule, and can be regarded as vibrational modes in an elastic protein model. From this point of view, the depth of the global minimum is discussed and concluded to be at least 150 kcal/mol.

References

1. Anfinsen, C. B. & Scheraga, H. A. (1975) *Adv. Protein Chem.* **29,** 205–300.
2. Nakanishi, M., Tsuboi, M. & Ikegami, A. (1973) *J. Mol. Biol.* **75,** 673–682.
3. Lumry, R. (1974) *Proc. Natl. Symp. Sickle Cell Disease,* DHEW Pub. No. (NIH) 75-723, pp. 165–166.
4. Gō, M., Gō, N. & Scheraga, H. A. (1970) *J. Chem. Phys.* **52,** 2060–2079.
5. Gō, M., Gō, N. & Scheraga, H. A. (1971) *J. Chem. Phys.* **54,** 4489–4503.
6. Gō, M., Hesselink, F. T., Gō, N. & Scheraga, N. (1974) *Macromolecules* **7,** 459–467.
7. Gō, M. & Gō, N. (1976) *Biopolymers* **15,** 1119–1127.
8. Suezaki, Y. & Gō, N. (1976) *Biopolymers* **15,** 2137–2153.
9. Gō, N. & Scheraga, H. A. (1976) *Macromolecules* **9,** 535–542.
10. Suezaki, Y. & Gō, N. (1975) *Int. J. Peptide Protein Res.* **7,** 333–334.
11. Ziman, J. M. (1965) in *Principles of the Theory of Solids,* Cambridge University Press, Cambridge, Chap. 1.
12. Landau, L. D. & Lifshitz, E. M. (1959) in *Theory of Elasticity,* Pergamon, New York, p. 99.
13. Brown, K. G., Erfurth, S. C., Small, E. W. & Peticolas, W. L. (1972) *Proc. Natl. Acad. Sci. USA* **69,** 1467–1469.
14. Warme, P. K. & Scheraga, H. A. (1974) *Biochemistry* **13,** 757–767.

Received August 10, 1977
Accepted October 31, 1977