Proc. Int. Symposium on Advances in Neutron Scattering Research, Tokai, 2000
J. Phys. Soc. Jpn. **70** (2001) Suppl. A pp. 420–422

Simulations of Neutron Solvent Contrast Variation of Proteins and Intramolecular Structures

Hiroki IWASE and Mitsuhiro HIRAI

Department of Physics, Gunma University, Maebashi 371-8510, Japan

We calculated theoretical neutron scattering curves of human lysozyme and its 27 mutants, including other 30 proteins, with varying the contrasts. From the theoretical scattering curves, we derived the three basic scattering functions corresponding to the molecular shape, the intramolecular structure and the cross term between them. We found that the basic scattering function corresponding to the intramolecular structure well reflects the packing geometry of polypeptide chains and that the difference between the intramolecular scattering functions of the wild-type protein and the mutant can be used as an index of stability of the protein. The present simulations suggest a high potentiality of a wide-angle neutron scattering method using high-intensity pulsed-neutron source in future researches of protein structures.

KEYWORDS: neutron scattering, human lysozyme, protein structure, structural hierarchy, mutation.

§1. Introduction

X-ray and neutron scattering methods for solutions are well recognized to provide insights of biological macromolecular structures and functions in spite of the low-resolution structural information on solute particles. These methods are usually used to determine the structures of the solute particles, such as overall shapes, radii of gyration, or some large-scale heterogeneity of the particles, mainly by using data sets in small-angle scattering regions.[1] Recently, by using small- and medium-angle X-ray and neutron scattering data, we successfully revealed that the structural transition multiplicity of a globular protein strongly depends on the protein structure hierarchy.[2–5] Our previous reports indicate that the solution scattering methods provide us fruitful information on tertiary and interdomain-correlation structures of proteins, suggesting that these methods have an important role for the studies of a mechanism of folding-and-unfolding of proteins as well as other methods.[6]

On the other hand, various types of neutron contrast variation methods have been developed and shown their high potentialities especially for structural studies of supermacromolecules.[7] The most popular and conventional method is the solvent contrast variation (SCV) method. This method is able to separate an experimental scattering curve $I_0(q)$ of a solute particle into three basic scattering functions and to provide an information of the internal structure of the solute particle, where q, the magnitude of scattering vector defined by $q = (4\pi/\lambda)\sin(\theta/2)$; θ and λ are the scattering angle and the wavelength.[8–12] However, qualitative meanings of the intramolecular structures revealed by the SCV method seem to be somewhat ambiguous. Then, to examine a further availability of the neutron SCV method, we have carried out the SCV simulations for about 60 proteins based on the atomic coordinates from the Protein Data Bank (PDB).

§2. Simulation Method and Selection of Proteins

According to the scheme of the SCV method, the $I_0(q)$ of a particle in a solution is given as follows.[13]

$$I_0(q) = \Delta\rho^2 I_v(q) + \Delta\rho I_{vf}(q) + I_f(q) \qquad (2.1)$$

where $\Delta\rho$ is the excess average scattering density, so-called contrast, of the solute particle in comparison with that of the solvent; $I_v(q)$, $I_f(q)$, and $I_{vf}(q)$ are the basic scattering functions corresponding to the particle shape, the intraparticle structure and the cross term between $I_v(q)$ and $I_f(q)$, respectively. These basic scattering functions can be separated from the scattering curves at least with three different contrasts. We calculated the theoretical neutron scattering curves $I_0(q)$ of a protein at three different contrasts (in 100 %, 80%, 60% v/v D2O) by using the program called CRYSON.[14] This program can well explain experimental scattering curves of proteins in solutions, especially for the q region below ~0.6 Å^{-1}. The concept of "contrast" stands on the physical basis that solvent molecules are regarded to give homogeneous back-ground scattering, therefore, this concept no longer holds in a high-angle scattering limit due to the finite volumes of solvent molecules. Then, after the calculations of $I_0(q)$ of a protein below $q = 1$ Å^{-1} at three different contrasts, we solved three simultaneous equations in the unknown basic scattering functions ($I_v(q)$, $I_f(q)$, and $I_{vf}(q)$) according to the equation (2.1). The proteins treated were selected from the database of Structure Classification Of Proteins called SCOP.[16] This database groups protein folds into four classes depending on the type and organization of secondary structure elements: all-α, all-β, α/β, and $\alpha+\beta$. In the present study we selected 30 proteins classified in all-α, human lysozyme (classified in $\alpha+\beta$) and its 27 mutants whose thermal characteristics have been studied in detail.[17]

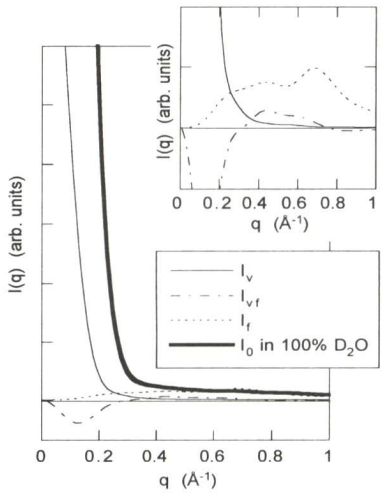

Fig. 1. Theoretical neutron scattering curve $I_0(q)$ of horse myoglobin in 100 % D2O and basic scattering functions ($I_v(q)$, $I_f(q)$, and $I_{vf}(q)$).

§3. Theoretical Scattering Functions and Intramolecular Structures

Figure 1 shows the theoretical scattering curve of horse myoglobin (PDB file code: 1WLA) in 100 % D2O and the three basic scattering functions. Stuhrmann *et al.* separated experimentally the observed scattering curves of whale myoglobin into the basic scattering functions, and showed that those experimental basic scattering functions agree with the theoretical ones calculated from a method using spherical harmonics.[8,9] It should be mentioned that the $I_f(q)$ of myoglobin has a broad peak around q = 0.6-0.8 Å^{-1}. As myoglobin is classified in all-α protein in the SCOP database, we chose other all-α proteins registered in PDB, and separated the scattering curves of these proteins into three basic scattering functions. Table I shows the PDB file codes of the all-α proteins used for the present calculations. Figure 2(a) shows several $I_f(q)$ functions of those proteins. The bars in the theoretical $I_f(q)$ function of 1WLA show the expected magnitudes of the statistical errors when we use a WINK type scattering spectrometer (WINK is a wide-angle neutron diffractometer at the pulsed neutron source KENS of KEK) that will be constructed at a future high-intensity pulsed-neutron source of the joint project between JAERI and KEK on high-intensity proton accelerators. Whereupon we simply assumed that the pulsed-neutron intensity at the joint project will be ca. 100 times stronger than that at KENS, and we used the q dependence of the relative error obtained experimentally from myoglobin at KENS. One of the prominent characteristics of a time-of-flight (TOF) scattering spectrometer at a pulsed-neutron source would be that the q dependence of the relative error does not simply increase with increasing q value since neutrons with a wide-band wavelength can be used. Figure 2(b) shows the number of the positions of the maximum values ($I_f(q)_{max}$) in the $I_f(q)$ functions at the defined q regions with the interval of 0.02 Å^{-1}. Most of $I_f(q)_{max}$ values are in the range of q = 0.55-0.75 Å^{-1}, corresponding to ca. d = 8.4-11.4 Å in real-space distance. Previously, Chothia *et al.* analyzed the packing geometry of pair of helices

Table I. PDB file codes of all-α proteins used for the theoretical scattering curve calculations.

PDB codes				
1ABV	1FHA	1LBU	1PRB	2END
1ADT	1FPS	1LIS	1SIG	2TRT
1AEP	1GRJ	1LPE	1SZT	3ADK
1AK0	1HRC	1NKD	1TNS	3ICB
1COO	1HYP	1NRE	1VIN	4TNC
1CUK	1LBD	1PAX	2ASR	5CPV

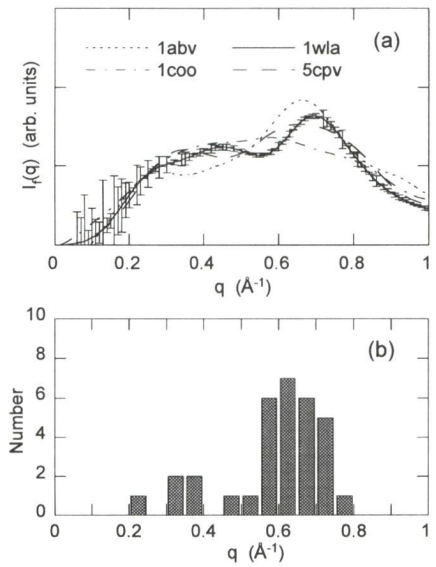

Fig. 2. Comparison of the basic scattering functions $I_f(q)$ of several proteins classified as all-α proteins at SCOP, (a). Distribution of the maximum values ($I_f(q)_{max}$) in $I_f(q)$ at the defined q ranges, (b)

for about 10 proteins. They showed that the average distance of helix-to-helix packing is distributed from 6.8 Å to 12 Å and that the maximum probability of the distance is in the range from 9 Å to 11 Å.[16] Their results agree with our present results in Figure 2(b), suggesting that the basic scattering functions $I_f(q)$ at q = 0.3-0.8 Å^{-1} well reflect the intramolecular correlation in protein structures as pointed elsewhere.[2,4,5]

§4. Intramolecular Structure and Gibbs Free Energy of Protein Mutant

Mutation techniques of proteins have been used for many studies on a mechanism of folding-and-unfolding of proteins since some mutations greatly change stability of proteins under various conditions.[6] The thermal stability and the tertiary structures of human lysozyme and its mutants have been studied intensively in detail.[17] For considering a further use of the neutron SCV method in protein structure studies, it is very interesting to exam-

Table II. PDB file codes a of human lysozyme used for the theoretical scattering curve calculations.

PDB codes				
1LZ1	1LHL	1OUF	1WQN	1YAN
1HNL	1OUA	1OUG	1WQO	1YAO
1LHH	1OUB	1OUH	1WQP	1YAP
1LHI	1OUC	1OUI	1WQQ	1YAQ
1LHJ	1OUD	1OUJ	1WQR	
1LHK	1OUE	1WQM	1YAM	

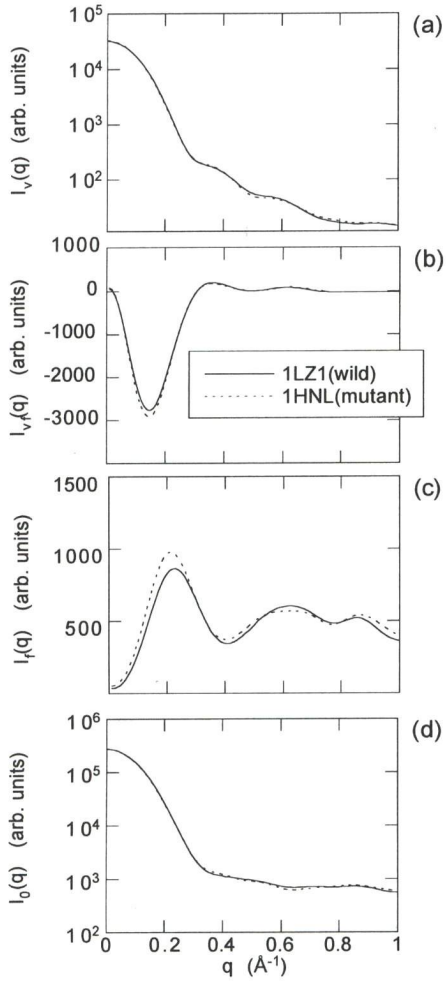

Fig. 3. Theoretical scattering curve $I_0(q)$ at 100 % D_2O and basic scattering functions of human lysozyme (PDB file code: 1LZ1) compared with those of the mutant (PDB file code: 1HNL). (a), (b), (c) and (d) correspond to $I_v(q)$, $I_{vf}(q)$, $I_f(q)$, and $I_0(q)$, respectively.

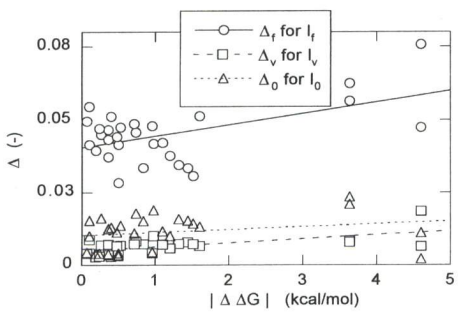

Fig. 4. Relation between the deviation factor Δ and the Gibbs free energy $|\Delta\Delta G|$ of the wild-type and mutant of human lysozyme.

ine an effect of mutation on a intramolecular structure of protein. Then, we treated the wild-type human lysozyme (PDB file code: 1LZ1) and its mutants, and compared those basic scattering functions. Table II shows the PDB file codes of the mutants. We calculated the deviation factors Δ which means the difference between the scattering functions ($I_0(q)$, $I_v(q)$ or $I_f(q)$) of the wild-type and the mutant as defined by

$$\Delta_i = \sum_{q=q_1}^{q_2} |I_i^{wild}(q) - I_i^{mutant}(q)| \Big/ \sum_{q=q_1}^{q_2} I_i^{wild}(q) \quad (4.1)$$

where $i = 0, v, f$. In the calculations of the Δ values we used the scattering functions in the q range of q = 0.01-1 Å^{-1}. The largest Δ values are obtained 0.076 in $I_f(q)$ for 1HNL, 0.019 in $I_v(q)$ for 1HNL, and 0.024 in $I_0(q)$ for 1QUA, respectively. For every mutant, the Δ_f value in $I_f(q)$ is much larger than those in $I_v(q)$ and $I_0(q)$. Figure 3 shows the basic scattering functions and $I_0(q)$ of 1LZ1 and 1HNL. The difference between the scattering functions of the wild-type and the mutant is much larger in $I_f(q)$ compared with those in $I_0(q)$ and $I_v(q)$. Figure 4 shows the Δ value dependence of $|\Delta\Delta G|$ which is defined by the difference in Gibbs free energy between the wild-type and mutant proteins.[17] We got the $|\Delta\Delta G|$ values of the human lysozyme and the mutants from the ProTherm database.[18] In Figure 4 the Δ values tend to become larger with increasing $|\Delta\Delta G|$, which is more evidently seen in Δ_f. Thus, the degradation of thermal stability of the protein in the mutation appears as the deviation of the intramolecular scattering function $I_f(q)$ of the mutant from that of the wild type. In other words, the Δ value in $I_f(q)$ determined by the neutron SCV method would be used as an index of intramolecular structural change and stability of proteins in solutions.

In the case of X-ray scattering, it is essentially difficult to obtain basic scattering functions of proteins. In addition, the covered q range in one measurement is rather limited due to the use of monochromatized X-ray beam. On the other hand, a pulsed-neutron scattering method using the TOF technique can cover a very wide q range. Therefore, a wide-angle neutron scattering using both the SCV and TOF methods at a high intensity pulsed-neutron source would be a very promising technique to study the structural hierarchy of globular proteins.

1) F. A. Feigin and D. I. Svergun: *Structure Analysis by Small Angle X-ray and Neutron Scattering* (Plenum Press, New York, 1987)
2) M. Hirai *et al.*: Physica B **241&243** (1998) 1159 : J. Phys. Chem. B **102** (1998) 1308.
3) H. Iwase, M. Hirai, S. Arai, M. Mitsuya, S. Shimizu, T. Otomo and M. Furusaka: J. Phys.Chem Solids **60** (1999) 1379.
4) S. Arai and M. Hirai: Biophys. J. **76** (1999) 2192.
5) M. Hirai *et al.*: J. Phys. Chem. B **103** (1999) 549: Thermochim. Acta **344** (2000) 95.
6) *Protein Folding* ed. E. E. Creighton: (W. H. Freeman & Co., New York, 1992).
7) H. B. Stuhrmann: Physica B. **156&157** (1989) 444.
8) H. B. Stuhrmann: J. Mol. Biol. **77** (1973) 363.
9) K. Ibel and H. B. Stuhrmann: J. Mol. Biol. **93** (1975) 255.
10) H. B. Stuhrmann and H. Fuess: Acta Cryst. **A32** (1976) 67.
11) H. B. Stuhrmann, J. Haas, K. Ibel, M. H. Koch and R. R. Crichton: J. Mol. Biol. **100** (1976) 399.
12) H. Stuhrmann and A. Miller: J. Appl. Cryst. **11** (1978) 325.
13) H. B. Stuhrmann and R. G. Kirste: Z. Phys. Chem. Frankfurt **46** (1965) 247.
14) D. Svergun *et al.*: J. Appl. Cryst. **26** (1995) 768. : Proc. Natl. Acad. Sci. USA. **95** (1998) 2267.
15) A. G. Murzin, S. E. Brenner, T. Hubbard, C. Chothia: J. Mol. Biol. **247** (1995) 536.
16) C. Chothia, M. Levitt and D. Richardson: J. Mol. Biol. **145** (1981) 215.
17) K. Takano, Y. Yamagata, S. Fujii, and K. Yutani: Biochemistry **36** (1997) 688.
18) M. M. Gromiha, J. An, K. Hino, M. Oobatake, H. Uedaira and A. Sarai: Nucleic Acids Res. **27** (1999) 1.