AINO JOURNAL, Vol. 5, 2006, pp. 17－23

Review

Three-dimensional structure analyses of protein interactions and their application to medicine

Masaaki MATSUSHIMA *Professor, Faculty of Nursing and Rehabilitation, Aino University*

Abstract

The three-dimensional structures of proteins are published in many scientific journals and by the mass media. Although the charts displayed in these publications look like animation figures constructed using computer graphics, considerable research has contributed determining those molecular structures. Identifying the three-dimensional structures of proteins is indispensable to understanding their functional interactions with other molecules. Therefore, determining these structures greatly contributes to basic life sciences and application of this knowledge contributes to the creation of effective medicines. This report will show three-dimensional structures of the proteins determined by this author and the interactions between macromolecules and small compounds, then discuss application to medicine.

Key words: proteins, three-dimensional structures, protein interactions, simulations, medicine

Introduction

The human genome was accurately determined by International Human Genome Sequencing Consortium and published in "Nature" in 2004, and the total expected number of genes, about 22 thousand, is less than that previously expected. Since the number of proteins, of which all of amino acid sequences may be coded by the genes, has been thought to be more than 100 thousand in the human body. Although more studies about splicing or protein processing are necessary to clarify RNA processing and protein expression, the total range of genes in the human body has been investigated by various research projects.

Funding from many nations has been injected in the life sciences over the recent several years, because of the human desire to be cured from diseases and to lead healthy lives. In order to respond these demands, nations have agreed to cooperatively promote research into life science. One such scientific research subject is the analysis of the proteins that play very important roles in the living body, such as enzymes in metabolism and catabolism, or proteins in development, regeneration, apoptosis, and signal transduction or conduction by proteins to maintain homeostasis (National Project on Protein Structural and Functional Analysis by the Ministry of Education, Culture, Sports, Science and Technology has focused on 3,000 fundamental structures and functions of the major proteins, 'Protein 3000', in which many scientists in Japan have contributed since 2002). Identifying the three-dimensional protein structures is thought to contribute to increasing basic knowledge about the functions, resulting in efficient development of new medicines in their application. Especially, membrane proteins are expected to be the objects of receptors, which can be targets for medicines. When a disease arises from a disorder in a particular protein function, the protein(s) related to the disease should be studied with regard to their functions as well as interactions in the functional network of the body. Disordered functions could finally be controlled by the patients themselves supported by medication or medical treatments.

To study a disease, proteins related to the disease should at least be identified in laboratory animals allowing a medical target to be selected from those proteins. It is possible that

the target protein was analyzed in human genome, and that the gene can be obtained and the protein over-expressed in a suitable system. Once the molecular protein is obtained, huge screening system in a chemical compound library is applied to inhibit or enhance the protein function. Such a pharmaceutical system is time-consuming and expensive to operate. Other useful information can be obtained from the three-dimensional structure of the protein. The function of the protein can be explained by the structure of the functional site, since proteins with a similar function have similar three-dimensional structures at the active site, even if they are constructed of different amino acid sequences.

The three-dimensional structural studies are started by examining the expression and purification of the identified proteins. When the gene of the protein is available, it is possible to over-express the protein, even if there is limited protein expressed in the body, or very rarely expressed at a particular stage of life. Such proteins can be rather easily purified, and then significant amount of the protein molecules can be supplied for scientific studies.

The main foci of research are the interactions between a protein and an oligo-nucleic acid molecule (including ribo-nucleic acid), and a protein and compounds, among proteins. Protein-nucleic acid interactions can initiate gene duplication or repair, and protein expression. Protein-protein interactions relate to their functions, such as those between antigens and antibodies. And an interaction between a protein and a small molecule can involve a transduction of a signal, or a functional control by a medicine against the protein.

Here, the benefits and limitations of protein structure analyses are being introduced to medical researchers. The author hopes that they will develop ideas regarding the interactions of protein molecules in the human body from these basic scientific results of structure studies.

Bio-molecular Structure determinations

There are three methods of analyzing three-dimensional structures of proteins at present. Those are X-ray crystallographic analysis, Nuclear Magnetic Resonance analysis, and Electron Microscopic analysis.

X-ray crystallographic analysis is a method of determining very accurate atomic positions in a unit cell of a crystal with observed X-ray diffraction intensities. Once a single protein crystal can be obtained, the volume of information from the crystallographic analyses is far greater than those by other methods with regard to their accuracy and application to other researches.

X-ray diffraction analysis of protein crystals (text books, Drenth, 1994; Kyogoku et al., 1997; Turk et al., 1999; Messerschmidt, 2007):

The same as crystal structure analysis of a small compound, a protein structure can be analyzed by observed diffraction intensities and unobservable phases of the diffractions from a protein crystal. Once the initial atomic parameters, which consist of element and positional and several other parameters, are obtained, the unobserved phases can be computed. Comparisons between those observed and calculated intensities, the atomic parameters can mathematically be refined within the limits of statistical error. The accuracy of the atomic positions is thought to be less than 0.1 Å (1Å=0.1 nanometer) after the results have been refined very carefully. To obtain the initial phases of diffractions by experiments, several heavy atoms are introduced in the crystal, or seleno-methionines in the protein molecule. The theoretical and fundamental methodology of protein crystallography was established in the 1980's. Thereafter, the studies have been advanced by the development of many experimental instruments and practical software. Using the shared gigantic facilities for the synchrotron X-ray radiation, such as SPring-8 in Hyogo prefecture (Homepage SPring-8) and Photon Factory in Ibaragi prefecture (Homepage Photon Factory) where enormously strong X-ray beams are used to irradiate various scientific targets, the three-dimensional structure analyses of proteins have further progressed (Home-page Protein Data Bank). For example, huge protein complexes, such as virus structures, can be examined, and the functional interactions of those proteins can be discussed together with biochemical experimental data.

Merits of X-ray crystal structure analyses

This method can be applied to all of crystallized molecules from small compounds to biomacromolecules, such as viruses or ribosomes. Protein molecules seem to be floated in

the solvent in crystals, except that the molecules interact at several points on the molecular surface with neighboring molecules in crystals. Although the molecules are packed in a soft crystal, their deformations are minor and local around points of the interaction. The protein functions have satisfactorily been explained together with the findings on crystallographic and biochemical experiments.

The parameters of small molecules around a protein, such as molecules of water, ions, products or inhibitors of enzymes and biologically interactive chemical factors, are also obtained by this analysis. These molecules are often quite important contributors to protein functions. All of these accurate geometrical parameters greatly support simulation calculations of the protein molecule in water. The results of such simulation calculations can open discussions about the dynamic motion and forces in the protein molecule and between the protein molecule and small molecules.

Demerits of X-ray crystallographic analyses

In order to apply crystallography, the objects must be crystallized (McPherson, 1999; Miki et al., 1995). The size of a crystal is ideally a 0.1–0.2 mm cube. A protein target must be biochemically purified before crystallization. In order to crystallize protein molecules, the total amount of a sample protein over-expressed by gene engineering is expected to more than one hundred milligram. There is no theory about protein crystallization, and researchers describe crystallization of bio-macromolecules as an art. If a reliably effective method can be identified, protein and life science would progress 10-fold. Such three-dimensional crystallization of bio-macromolecules is one of the great problems in this method, especially for membrane protein analyses.

Another problem is that the parameters related to 'time' are averaged in the diffraction intensity measurements, though these parameters could be obtained by simulation calculations of the protein structures.

Results by X-ray crystallographic structure analyses:

1) **An example of very accurate determination of structure**

Three-dimensional atomic co-ordinates of human lysozyme, water molecules, and sodium and chloride ions have been very accurately studied by X-ray crystallography (Fig. 1) (Matsushima et al., 1990). The conformations of the amino-acid side chains and some hydrogen atoms of amino acid side chains, as well as the water molecules around the protein, were clearly shown in the results of high-resolution analyses (Higo et al., 2002). The three-dimensional structure of the complex between human lysozyme and synthetic substrates of tetra- and hexa-N-acetyl-glucose was also analyzed (Fig. 2) (Matsushima et al., 1990; Song et al., 1994).

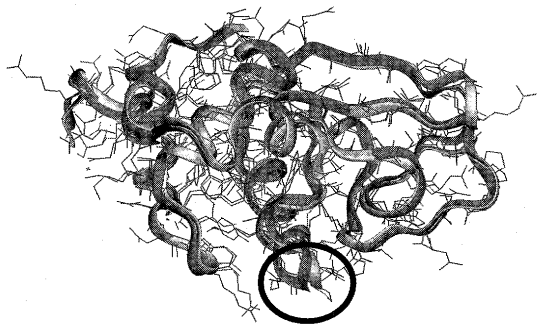

Fig. 1 Over-lay of hen egg-white lysozyme (HEWL) human lysozyme (HLZM)
The main chain of HEWL is painted by gray and that of HLZM is painted by dark. The ellipsoid circle in this figure indicates the part which mainly interacts with the antibody of HyHEL10

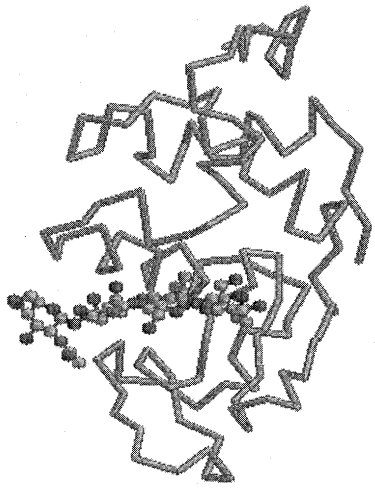

Fig. 2 The three-dimensional structure of the complex between human lysozyme (HLZM) and its product of tetra-saccharide
The product (shown by ball and stick) is bound in the active site cleft of HLZM

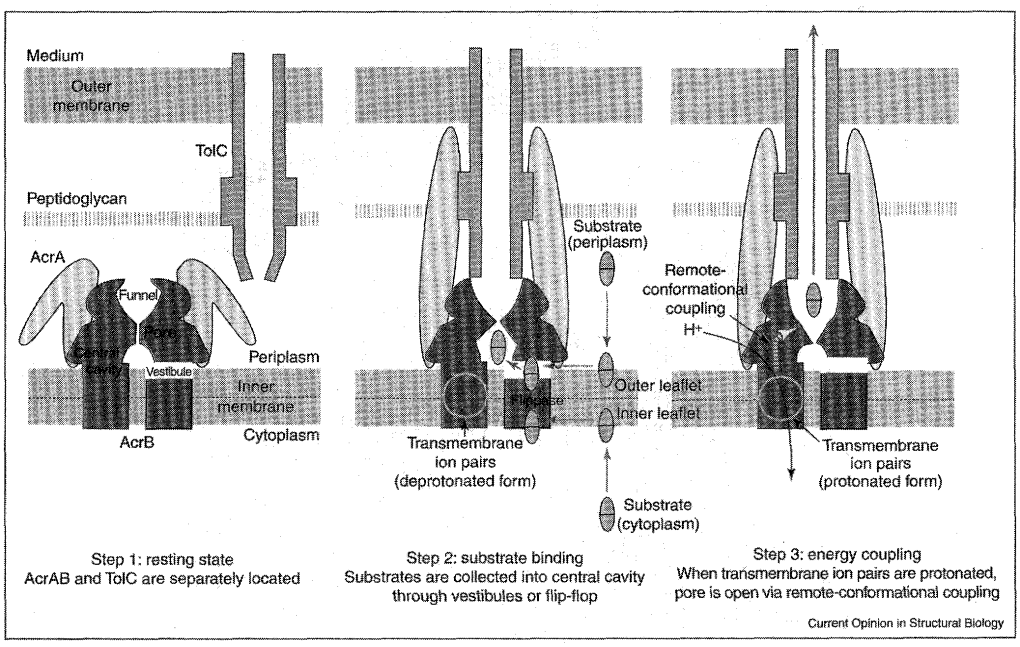

Fig. 3 The proposed mechanism of the drug transportation by AcrB, TolC, and AcrA in periplasmic space (copied the Fig. 5 in the referred report (Murakami et al., 2003 with permission)

2) Large membrane protein structures of multi-drug transporters

It is clear that structure analyses of proteins very much contribute to the study of molecular functions of very large molecular systems, such as multi-drug transporters (Yamaguchi et al., 2003; Murakami et al., 2003a and 2003b). The structural study of bacterial multi-drug transporter sheds light on multidrug resistance by the bacteria Pseudomonas aeruginosa (Fig. 3). The protein molecular units are in periplasmic space. Those are AcrB, TolC, and AcrA, respectively. The total sizes of AcrB are about 100 Å diameter and 50 Å thick at the trans-membrane domain of the inner membrane, and protrusions into periplasmic space are about 40 Å thick at the pore domain and about 30 Å thick at the TolC docking domain. The TolC is a tube about 100 Å long forming α-helixes in periplasmic space (Koronakis et al., 2000). The trimer of AcrA, which has never determined crystallographically, is a connector between AcrB and TolC molecules. A pore through the centers of these molecules forms a tunnel spanning from cytoplasm via periplasm to outside the bacteria. Drugs are pumped out by the energy of the H^+ gradient between the concentrations in periplasm and cytoplasm. The drug-binding site is located in the transmembrane domain of AcrB. The reporters explained the multi-drug exportation by charged side chains of two asparatic acids, a lysine and an arginine, which are responsible for nonspecific binding of drug molecules and proposed the functionally rotating mechanism (Murakami et al., 2006)

Applications of Protein Molecular Structures

1) Simulations of protein thermo-stabilities

The enzymatic function of lysozyme is discussed in relation to stereo-chemical structures of the protein and substrate atoms on the bases of these experimental findings. The atomic parameters of human lysozyme were used to develop computer simulations of protein structures in solution. The results of molecular dynamic simulations could explain the thermo-stability of the protein, and for example, the computed free-energy of an α-lactalbumin mutant (2.19 kcal/mol) agrees very well with the experimentally determined finding (1.86 kcal/mol) (Horii et al., 2001).

2) Simulations of protein-protein interactions

The computer simulation method could also predict the comparative stability of the interaction between proteins in solution. Computations of the simulation were based on three accurately determined three-dimensional structures of two antigens, hen egg-white lysozyme (HEWL) and human lysozyme (HLZM), which were very similar stereo-chem-

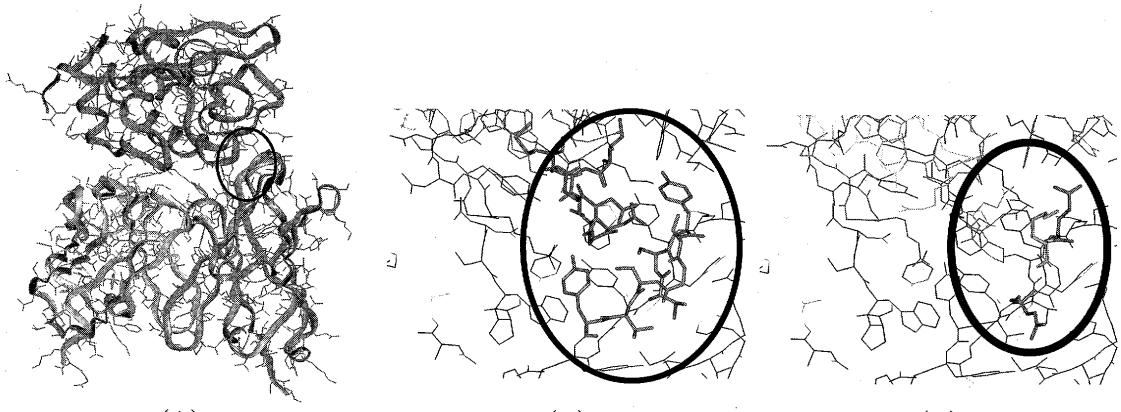

Fig. 4 The interactions of an antigen (HEWL and HLZM)
(A) The analyzed structure of the complex between HEWL and its antibody HyHEL10
The whole structure of HEWL and the hyper variable part, Fv, of the antibody are shown in this figure. The circle shows the main interactions of the antigen and the antibody
(B) The details of the main interactions between HEWL and HyHEL10
The amino acids painted dark in the upper part of an ellipsoid circle is those of HEWL and the amino acids painted gray belongs in the molecule of HyHEL10 in the lower part of the circle
(C) The amino acids mutated by the simulation calculations
The amino acids sequence of HLZM is shown by lines in upper left of an circle, and three amino acids of HyHEL10 shown by stick lines are predicted as a high affinity sequence against HLZM by the simulation calculations

ical structures, and of the complex structure of HEWL and antibody molecules (HyHEL10) (Fig. 4 (A)). A few amino acids of the antibody in the interface of the complex can be replaced in the calculations (Fig. 4 (B) and (C)). Using of powerful computer resources, this method might be developed to predict a few antibodies against a forth-coming antigen before infection.

3) Medicine Developments

Tens of candidate compounds have been experimentally screened from an enormous number of compounds stocked in the chemical libraries of pharmaceutical companies.

It might be suspected that there is a far remote relation between the three-dimensional structures of proteins and medical practice, but basic researches has greatly progressed since the human genome project, and the results are being applied to the pharmaceutical investigations of medicines.

The three-dimensional structures of enzymes related to diseases are exceptionally useful in the creation of new medicines. When the structure has been identified, the three-dimensional structures of millions of actual chemical compounds can be used to develop virtual complexes on a computer, and promising chemical candidates can be selected from indices indicating the stability of these complexes. Supposing that several hundreds of these chemicals are obtained, a fudge machinery system of a pharmaceutical company will be activated, a few chemical compounds will be refined for test application to human disease, and only one of those best compounds will be developed as a commercial medicine on market after several ten billion yen are invested over a period of 10 years or more.

One of the well-known examples is the development of the anti-viral medicine, anti--neuraminidase compound (Fig. 5). X-ray crystallographers determined the complex of an inhibitor and influenza virus neuraminidase. The inhibitor molecule looks very much like the enzyme substrate molecule, N-acetylneuramic acid (Varghese et al., 1995; Lawrence et al., 2004). Identifying the three-dimensional structure of the enzyme was an important contribution to designing the inhibitors, and allowing improved inhibitors to be proposed at a stage of the development (Varghese et al., 1997). The compound of 4-guanidino-Neu5Ac2en in the complex was finally developed in medicines, marketed as Zanamivir and Oseltamivir (Fig. 6).

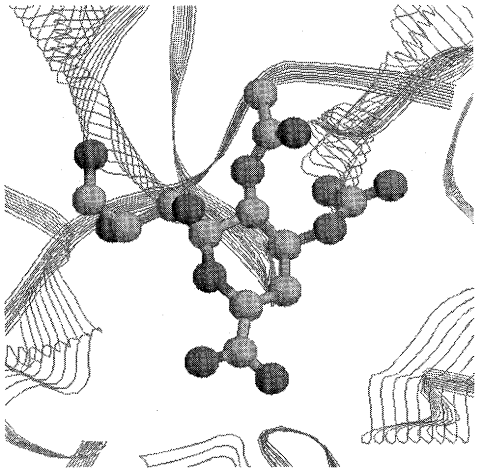

Fig. 5 The three-dimensional structure of the complex between the influenza virus neuraminidase and its inhibitor compound
The main chain of neuraminidase was shown by ribbon and the inhibitor molecule by ball and stick. All of the protein side-chains were erased in order to be clarified

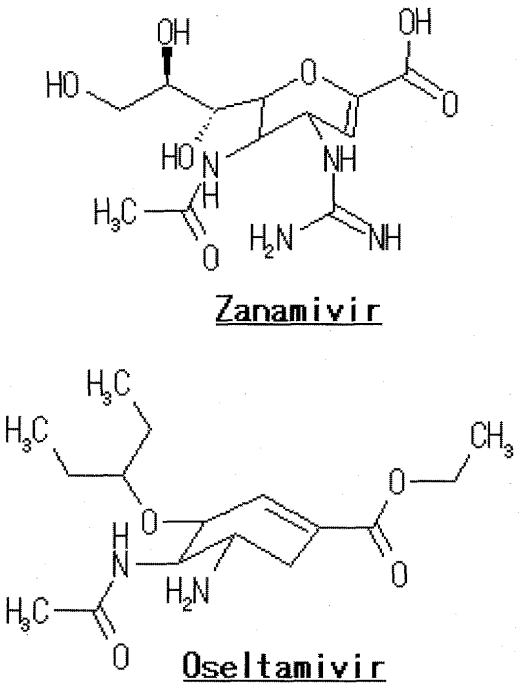

Fig. 6 The chemical structures of Zanamivir; 'Relenza', and Oseltamivir; 'Tamiflu'

Conclusion

Anatomy is quite important in medical education and surgery. Physiology is also very important to understand the physical condition of a patient in clinical medicine. Those subjects focus on understanding the human body and organs. Studies of protein structures and their functions are very similar to those of anatomy and physiology. In order to understand proteins, the three-dimensional structures of the protein molecules must be analyzed, similar to understanding human structures in anatomy, then the protein functions must be clarified, similar to clarifying organ activity in physiology. It is possible that the three-dimensional structures of a target protein molecule can assist in selecting or designing several molecular candidates to cure specific diseases.

References

Drenth J: Principles of Protein X-ray Crystallography. Springer-Verlag, 1994, translated by Takenaka A, Katsube Y, Sasada Y. Tokyo, Springer-Verlag, 1998

Higo J, Nakasako M: Hydration structure of human lysozyme investigated by molecular dynamics simulation and cryogenic X-ray crystal structure analyses: on the correlation between and crystal-water sites, solvent-density and solvent-dipole. J Comput Chem 23: 1323–1336, 2002

Home-page SPring-8: http://www.spring8.or.jp/

Home-page Photon Factory: http://pfwww.kek.jp/

Home-page Protein Data Bank: http://www.pdb.org/index-j.html

Horii K, Saito M, Yoda T, Tsumoto K, Matsushima M, Kuwajima K, et al.: Contribution of Thr29 to the Thermodynamic Stability of Goat α-Lactalbumin as Determined by Experimental and Theoretical Approaches. Proteins 45: 16–29, 2001

Koronakis V, Shariff A, Koronakis E, Luisi B, Hughes C: Crystal structure of the bacterial membrane protein TolC central multidrug efflux and protein export. Nature 405: 914–919, 2000

Kyogoku M, Tsukihara T, eds.: シリーズ・ニューバイオフィジックス③ 構造生物学とその解析法. Tokyo, Kyoritsu Shuppan, 1997 (in Japanese)

Lawrence MC, Borg NA, Streltsov VA, Pilling PA, Epa VC, Varghese JN, et al.: Structure of the haemagglutinin-neuraminidase from human parainfluenza virus type III. J. Mol Biol. 335 (5): 1343–1357, 2004

Matsushima M, Inaka K, Danno M: 蛋白質結晶の回折X線強度測定と精密構造解析の実際. Journal of the Crystallographic Society of Japan 32: 268–273, 1990 (in Japanese)

McPherson A: Crystallization of Biological Macromolecules. New York, Cold Spring Harbor Laboratory Press, 1999

Messerschmidt A: X-ray crystallography of biomacromolecules: A practical guide. Weinheim, Wiley-VCH, 2007

Miki K, Ataka M, Furuyama K, Higuchi Y, Miyashita T, eds: Crystallization of Biological Macromolecules. Proceedings: The sixth international conference on crystallization of biological macromolecules at Hiroshima, Japan on Nov. 12–17. Elsevier Science B. V., 1996

Murakami S, Yamaguchi A: Crystal structure of bacterial multidrug efflux transporter AcrB. Protein, Nucleic acid and Enzyme 48: 26–32, 2003a

Murakami S, Yamaguchi A: Multidrug-Exporting Secondary Transporters. Curr Opin Struct Biol 13: 443-452, 2003b

Murakami S, Nakashima R, Yamashita E, Matsumoto T, Yamaguchi A: Crystal structures of a multidrug transporter reveal a functionally rotating mechanism. Nature 443 (7108): 173-179, 2006

Turk D, Johnson L, eds.: Methods in Macromolecular Crystallography. IOS Press, 2001

Song H, Inaka K, Maenaka K, Matsushima M: Structural changes of active site cleft and different saccharide binding modes in human lysozyme co-crystallized with hexa-N-acetyl-chitohexaose at pH 4.0. J Mol Biol 244: 522-540, 1994

Varghese JN, Epa VC, Colman PM: Three-dimensional structure of the complex of 4-guanidino-Neu5Ac2en and influenza virus neuraminidase. Protein Sci 4: 1081-1087, 1995

Varghese JN: The design of anti-influenza virus drugs from the X-ray molecular structure of influenza virus neuraminidase. In: Vererapandian P, ed. Structure-Based Drug Design. New York, Marcel Dekker, 1997

Yamaguchi A, Murakami S, Nakashima R, Yamashita E: Crystal Structure of Bacterial Multidrug Efflux Transporter AcrB. FASEB J. A1185 (part2) suppl.: S 17, 2003