CROATIAN MEDICAL JOURNAL
CMJ
42(4):375-378,2001

REVIEW

Structural Biology and Its Applications to the Health Sciences

Aaron J. Oakley, Albert Ketterman[1], Matthew CJ. Wilce

Department of Pharmacology/Crystallography Centre, University of Western Australia, Perth, Australia; and [1]Institute of Molecular Biology and Genetics, Mahidol University, Salaya Campus, Thailand

Part of the decipherment of genomic information lies in understanding the structure and function of the protein products of these genes. Protein structure is of further importance because of the molecular basis of many diseases. Structural biology is the field of research focusing on the experimental determination of the structure of biological molecules. We review the field of structural biology and its application to medical research and drug discovery, and describe the structural results recently obtained in our laboratory for the detoxifying enzyme glutathione S-transferase from the Asian mosquito *Anopheles dirus* species B, an important malaria vector. These enzymes have detoxifying activity toward pesticides and thus contribute to pesticide resistance in insects.

Key words: *anopheles; crystallography, X-ray; drug design; glutathione transferase; insect vectors; magnetic resonance spectroscopy; malaria*

Since the first protein structure (of sperm whale myoglobin) was determined (1) and Watson and Crick discovered the double helix structure of DNA (2), there has been an ever-increasing research effort in the field of structural biology. Broadly, structural biology is defined as the investigation of the structure and function of biological systems at the molecular level. The significance of this field of research in part derives from the fact that macromolecular structure is important to many disease states. Sickle-cell anemia was recognized to be a result of a mutation in hemoglobin, causing it to polymerize into long rod-shape complexes that distort and destroy red blood cells (3). The structural consequence of this mutation is now understood (4). Today, many molecular diseases, such as cancer, are known to result from mutations of genes that affect the gene product, altering its function. This is invariably caused by changes in the structure and function of the protein product of the gene. Understanding these processes can be important for treating the disease. Information about protein structure can be used in so-called structure-based drug design, where a macromolecular structure is used as a template for drug design. This review examines the field of structural biology and its relevance to medicine and drug discovery.

Typically, protein crystallography and nuclear magnetic resonance (NMR) spectroscopy have been the tools of choice for the structural biologist. Briefly described here are the two main techniques for macromolecular structure determination.

X-ray Crystallography

X-ray crystallography can give atomic resolution structure of proteins and other macromolecules, such as DNA and their complexes. The technique requires the availability of milligram quantities of >99% pure macromolecule, usually produced through cloning and overexpression in bacterial plasmids, and then purified through the standard techniques of biochemistry, such as gel filtration, affinity chromatography, and ion exchange chromatography. The macromolecule must then be crystalized. This is usually achieved through the addition of the protein to a precipitant, such as ammonium sulfate or polyethylene glycol. Extensive trials of numerous potential conditions are often required to find a condition that gives crystals of sufficiently high quality for X-ray analysis. The next step is to place the crystal in an X-ray beam produced by a laboratory source or at a synchrotron radiation source, such as those located in Trieste (Italy) or Argonne (USA). Crystals diffract X-rays, and the resulting pattern of scattered X-rays is processed computationally to reveal the electron density of the molecule subject. This technique is reviewed in detail elsewhere (5).

Nuclear Magnetic Resonance Spectroscopy

NMR is another important tool for probing the structure of biological molecules. Like X-ray crystallography, this technique can provide three-dimensional structural information, but the underlying method is completely different. Like X-ray crystallography, the technique requires milligram quantities of

protein. This protein is subjected to a strong magnetic field (~1 Tesla) that causes atomic nuclei with unpaired protons or neutrons to "spin" with their magnetic moment aligned with the applied magnetic field. Suitable nuclei are ^{1}H, ^{15}N, and ^{13}C. The nuclei can be flipped by radio waves. Depending on nuclei chemical environment, the amount of energy in the applied radio waves required to "flip" the nuclei will change. As the nuclei "relax" to their ground state, which is in alignment with the magnetic field, they re-emit detectable radio waves. By analyzing this emission spectrum, information about the chemical environments of the various atoms can be obtained, leading to information about the locations of the various chemical groups in the structure. An excellent review of the NMR technique in structural biology is given by Wider (6).

The imaging of bio-molecules is not limited to the above techniques. Electron microscopy also provides structural information for molecules, but not to the same resolution as the above techniques.

Structure-based Drug Design

One of the most important consequences of the ability to determine macromolecular structures is the potential to design drugs based on those structures. There are number of successful examples of this application, and two of them are discussed here.

Viral Neuraminidase

Neuraminidase is a glycohydrolase that cleaves terminal sialic acid units from glycoproteins, glycolipids, and oligo-saccharides. It is found on the surface of the influenza virus, where it is thought to facilitate the escape of progeny virions from infected cells. The crystal structure of the sialic acid-enzyme complex (7) and subsequent computational analysis of this structure led to predictions of favorable substitutions to an existing lead compound (2-deoxy-2,3-didehydro-D-N-acetylneuramic acid, Neu5Ac2en; a non-selective neuraminidase inhibitor), and the synthesis of a superior inhibitor, 4-guanidino-Neu-5Ac2en (8), today marketed as Relenza (Glaxo Wellcome Australia Ltd, Boronia, Vic., Australia).

Human Immunodeficiency Virus Protease

Human immunodeficiency virus, HIV-1, which causes acquired immuno-deficiency syndrome (AIDS), is the subject of much structural research. When the genome of the virus was sequenced over a decade ago (9), it was suggested that a retroviral protease gene was present. The discovery that inactivation of retroviral protease by chemical inhibition or mutation resulted in immature, impotent virions (10). The first crystallographic studies showed that the protease is a dimer with a pepsin-like active site (11). Subsequent research resulted in the determination of hundreds of retroviral protease-inhibitor complexes. The retroviral protease inhibitors, Saquinavir (Roche Laboratories, Nutley, NJ, USA), Nelfinavir (Viracept, Agouron Pharmaceuticals, La Jolla, CA, USA), Indinavir (Merck Pharmaceuticals, Whitehouse Station, NJ, USA), and Ritonavir (Abbott Laboratories, Abbott Park, IL, USA) have been approved for human use in the treatment of HIV infection (12).

Discovery of Molecular Function

The following example aims to highlight how structural biology serves to facilitate the discovery of function of biological mechanisms. The antibiotic geldanamycin is known for its anti-proliferative and anti-tumor activities. The cellular target for geldanamycin was found to be the molecular chaperone heat shock protein 90 (Hsp90) (13). Heat shock proteins are known for their ability to bind and assist in the folding of many polypeptides. The reason this interaction results in anti-proliferative activity is that Hsp90 binds to and is required for the folding and activation of a range of client proteins involved in cell cycle regulation. Using X-ray crystallography, Stebbins and co-workers (14) determined the structure of the geldanamycin-Hsp 90 complex. They proposed that the geldanamycin-binding pocket is where Hsp 90 binds to polypeptides substrates. It was the determination of a similar crystal structure (15) that showed that geldanamycin was actually blocking an adenosine triphosphate binding site. Up until this structural revelation, the role of adenosine triphosphate in the biochemical functions of Hsp90 had been controversial.

Glutathione S-transferases

Our research focuses on detoxifying enzymes from the mosquito *Anopheles dirus* species B. These enzymes, called glutathione S-transferases, conjugate environmental and endogenous toxins to glutathione (GSH, γ–glutamyl-cysteinyl-glycine). Glutathione S-transferases are found in eukarya and bacteria, but not archea. They are dimeric proteins with a molecular weight of about 50 kDa. Glutathione S-transferases catalyze the attack of the sulfhydryl group of GSH upon electrophilic groups in the substrate (16). The active sites of glutathione S-transferases are described as being composed of a G-site (the GSH binding site) and the H-site (for binding hydrophobic co-substrates). A general form of the reaction is expressed thus:

$$GSH + RX \rightarrow GS\text{-}R + X^{-} + H^{+}$$

In this scheme, R is the substrate with electrophilic group X.

The glutathione S-transferase isozymes from *Anopheles* species are known to have detoxifying activity toward dichlorodiphenyltrichloroethane (DDT) and other pesticides (17), as they catalyze the conversion of DDT to dichlorodiphenyldichloroethane (Fig. 1). Evidence shows that these enzymes are in part responsible for pesticide resistance. In a DDT-resistant strain of the closely related mosquito *Anopheles gambiae*, an increase in the synthesis of glutathione S-transferase isozymes possessing a greater dehydrochlorination activity was found (18,19).

To date, we have determined the structure of two glutathione S-transferase isozymes from *Anopheles dirus* species B. These isozymes, designated as adGST1- 3 and adGST1-4, are produced from alter-

nate splicing and share a common N-terminal domain (20,21). The isoforms have the same overall structure (Fig. 2). The structure reveals the overall fold of the glutathione S-transferases, as well as the details of the active site. The G-sites of *Anopheles* glutathione S-transferases are similar to those in previously determined glutathione S-transferase isozymes, such as the sheep blowfly glutathione S-transferase (22). It comprises of several hydrogen-bond donors and acceptors that complement glutathione S-transferase. The greatest variability in the structure of the glutathione S-transferases is in the H-site. This structural heterogeneity is in part responsible for the specificity of the glutathione S-transferase isozymes for different hydrophobic substrates. The structure of adGST1-3 and

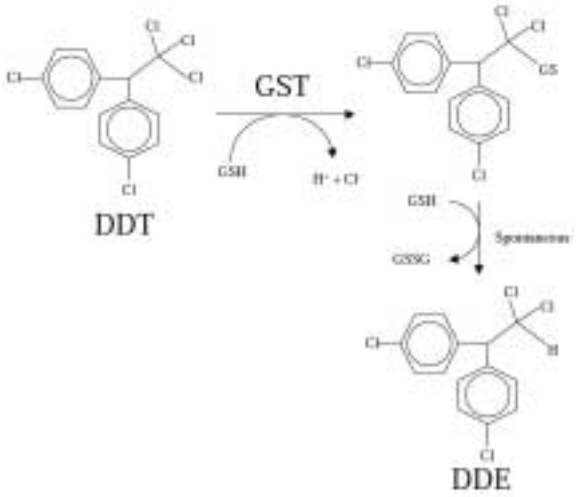

Figure 1. The chemical decomposition of dichlorodiphenyltrichloroethane (DDT) to dichlorodiphenyldichloroethane (DDE) as catalyzed by glutathione S-transferases. GSSG represents oxidized glutathione.

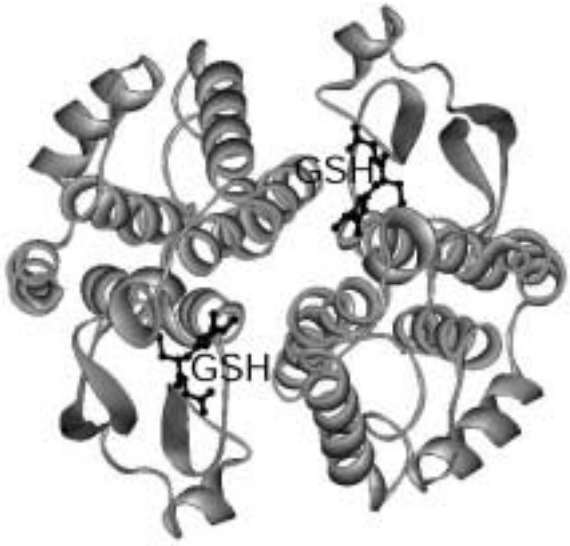

Figure 2. The three-dimensional structure of adGST1-4 as shown in ribbon form, where the polypeptide chain is represented as a single ribbon. The two identical polypeptide chains composing the GST dimer are represented. The substrate, glutathione (GSH), is shown in black.

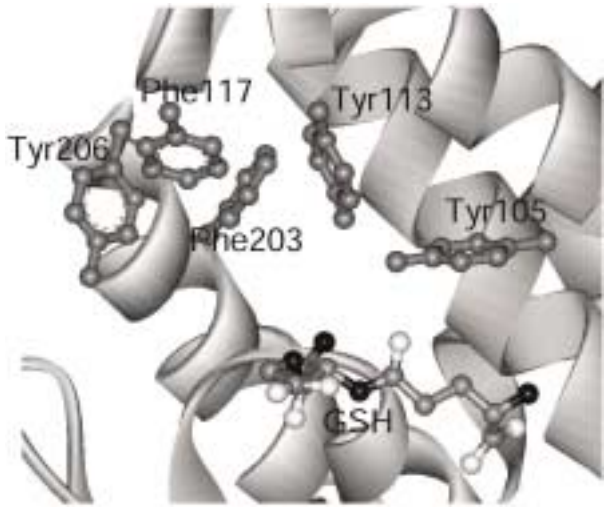

Figure 3. The active site of adGST1-4. The substrate, glutathione (GSH), is shown in ball-and-stick form. The polypeptide backbone is shown as a ribbon. Residues likely to be involved in binding dichlorodiphenyltrichloroethane (DDT) and other substrates are shown in ball-and-stick and are labeled.

adGST1-4 are similar, and consist of a pattern of hydrophobic amino acids in the H-site that are most likely responsible for binding DDT and other pesticides. These residues are tyrosine or phenylalanine, which are arranged to create a hydrophobic patch on the surface of the protein (Fig. 3).

Knowing the structures of these enzymes will help understanding the mechanism by which they break down pesticides. Future work will focus on trying to obtain the complex between the *Anopheles* glutathione S-transferase isoforms and pesticides, and determining their three-dimensional structures. The structures may be useful in the design of inhibitors that could be used in conjunction with pesticides to improve their efficacy. Such an approach may lead to decreased overall pesticide use.

The Future

In the post genomic era, protein structure will continue to be an integral part of basic biological research. As basic research into gene function continues, the molecular basis of many diseases will come to be understood. Consequently, revealed protein structure will increasingly provide the basis for drug discovery. The main challenges in the structure biology field will be in the investigation of multi-subunit complexes (e.g., ribosomal subunits) and integral membrane proteins. Structure-based drug design, in combination with other drug discovery strategies, such as combinatorial chemistry, will continue to grow in importance.

Acknowledgment

This research is supported by a grant from The Thailand Research Fund (AJK) and a UWA/ARC Small Research Grant (AJO). AJO is a recipient of a UWA postdoctoral research fellowship and

ARC fellowship. MCJW is supported by a Raine Medical Research Foundation Fellowship.

References

1 Kendrew JC, Bodo G, Dintzis HM, Parrish RG, Wyckoff H, Phillips DC. A three-dimensional model of the myoglobin molecule obtained by X-ray analysis. Nature 1958;181:662-6.

2 Watson JD, Crick FH. A structure of deoxyribose nucleic acid. Nature 1953;171:737.

3 Ingram VM. Gene mutation in human haemoglobin: the chemical difference between normal and sickle cell haemoglobin. Nature 1957;180:326-8.

4 Watowich SJ, Gross LJ, Josephs R. Analysis of the intermolecular contacts within sickle hemoglobin fibers: effect of site-specific substitutions, fiber pitch, and double-strand disorder. J Struct Biol 1993;111:161-79.

5 Oakley AJ, Wilce MCJ. Macromolecular crystallography as a tool for investigating drug, enzyme and receptor interactions. Clin Exp Pharmacol Physiol 2000;27: 145-51.

6 Wider G. Structure determination of biological macromolecules in solution using nuclear magnetic resonance spectroscopy. Biotechniques 2000;29:1278-92.

7 Varghese JN, Laver WG, Colman PM. Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 A resolution. Nature 1983;303:35-40.

8 von Itzstein M, Wu WY, Kok GB, Pegg MS, Dyason JC, Jin B, et al. Rational design of potent sialidase-based inhibitors of influenza virus replication. Nature 1993; 363:418-23.

9 Ratner L, Haseltine W, Patarca R, Livak KJ, Starcich B, Josephs SF, et al. Complete nucleotide sequence of the aids virus HTLV-III. Nature 1985;313:277-94.

10 Seelmeier S, Schmidt H, Turk V, von der Helm K. Human immuno-deficiency virus has an aspartic-type protease that can be inhibited by pepstatin A. Proc Natl Acad Sci USA 1988;85:6612-6.

11 Lapatto R, Blundell T, Hemmings A, Overington J, Wilderspin A, Wood S, et al. X-ray analysis of HIV-1 proteinase at 2.7 A resolution confirms structural homology amongst retroviral enzymes. Nature 1989;342: 299-302.

12 Wlodawer A, Vondrasek J. Inhibitors of HIV-1 protease: a major success of structure assisted drug design. Annu Rev Biophys Biomol Struct 1998;27:249-84.

13 Whitesell L, Mimnaugh EG, De Costa B, Myers CE, Neckers LM. Inhibition of heat shock protein HSP90–pp60v-src heteroprotein complex formation by benzoquinone ansamycins: essential role for stress proteins in oncogenic transformation. Proc Natl Acad Sci USA 1994;91:8324-8.

14 Stebbins CE, Russo AA, Schneider C, Rosen N, Hartl FU, Pavletich NP. Crystal structure of an Hsp90- geldanamycin complex: targeting of a protein chaperone by an antitumor agent. Cell 1997;89:239-50.

15 Prodromou C, Roe SM, O'Brien R, Ladbury JE, Piper PW, Pearl LH. Identification and structural characterization of the ATP/ADP-binding site in the Hsp90 molecular chaperone. Cell 1997;90:65-75.

16 Hayes JD, Pulford DJ. The glutathione S-transferase supergene family: regulation of GST and the contribution of the isoenzymes to cancer chemoprotection and drug resistance. Crit Rev Biochem Mol Biol 1995;30: 445-600.

17 Prapanthadara L, Promtet N, Koottathep S, Somboon P, Ketterman AJ. Isoenzymes of glutathione S-transferase from the mosquito Anopheles dirus species B: the purification, partial characterization and interaction with various insecticides. Insect Biochem Mol Biol 2000; 30:395-403.

18 Prapanthadara LA, Hemingway J, Ketterman AJ. Partial purification and characterization of glutathione S-transferases involved in DDT resistance from the mosquito *Anopheles gambiae*. Pestic Biochem Physiol 1993;47: 119-33.

19 Prapanthadara L, Ketterman AJ, Hemingway J. DDT-resistance in *Anopheles gambiae* Giles from Zanzibar Tanzania based on increased DDT-dehydrochlorinase activity of glutathione S-transferases. Bulletin of Entomological Research 1995;85:267-74.

20 Pongjaroenkit S, Jirajaroenrat K, Boonchauy C, Chanama U, Leetachewa S, Prapanthadara L, et al. Genomic organization and putative promoters of highly conserved glutathione S-transferases originating by alternative splicing in Anopheles dirus. Insect Biochem Mol Biol 2001;31:75-85.

21 Jirajaroenrat K, Pongjaroenkit S, Krittanai C, Prapanthadara L, Ketterman AJ. Heterologous expression and characterization of alternatively spliced glutathione S-transferases from a single *Anopheles* gene. Insect Biochem Molec Biol 2001;31:867-75.

22 Wilce MC, Board PG, Feil SC, Parker MW. Crystal structure of a theta-class glutathione transferase. EMBO J 1995;14:2133-43.

Received: June 9, 2001

Accepted: July 9, 2001

Correspondence to:

Matthew Wilce

Department of Pharmacology

University of Western Australia

35 Stirling Highway

Crawley WA 6009

Australia