Crystallography and Drug Design

K Suguna

Drugs treat diseases by blocking or modifying the functions of molecules such as proteins. Knowledge of the crystal structures of proteins is of tremendous help in designing effective drugs as the structures provide precise details of the mechanism of action and inhibition of the proteins at the atomic level. The impact of X-ray crystallography, by which the 3D structures are elucidated, on the efforts to combat diseases has been presented with a few examples.

K Suguna is a Professor at the Molecular Biophysics Unit, IISc, Bangalore. She is interested in determining the structures of biological macromolecules by X-ray crystallography to understand their structure-function relationships.

Structures of biological macromolecules, as determined by X-ray crystallography, not only reveal their shapes and sizes which define their functional properties, but also provide precise details of the intrinsic arrangement of atoms in the molecules. The structural information is invaluable in understanding the fundamental biological processes at the atomic level. This knowledge is of immense help in developing drugs for specific diseases by targeting molecules responsible for causing the diseases. 3D structures of several macromolecules, especially those of proteins, have been determined, enabling structure-based drug design for the treatment of several diseases. Structures, thus, play important roles in healthcare.

The first step in the structure-guided drug design process is to identify the target with a known structure. The target is a biological molecule, such as a protein, involved in crucial processes related to the disease, or in the case of pathogenic infections, the target should be present either in the pathogen or in the host. Proteins, such as enzymes, are appropriate targets as they have well-defined binding pockets where the substrates bind, and these can be used to fit small drug molecules and block their functions. Blocking the functions of the targets inhibits reactions leading to harmful effects in the body or prevents the pathogens from infecting the host. To bring in such effects, the inhibitor mol-

Keywords
Rational drug design, X-ray crystallography.

Figure 1. A cartoon representation of the interactions between a target molecule (yellow) and the designed inhibitor (green). Black lines represent possible interactions. Charges are shown as blue (+ve charge) and red (-ve charge) circles.

ecules should have better affinity than the natural substrates while interacting with the target. Strong binding can be achieved if there is perfect shape complementarity between the inhibitor and the binding pocket on the surface of the target (*Figure* 1), similar to inserting a missing piece in a jigsaw puzzle. Additional interactions through hydrogen bonding, van der Waals interactions, hydrophobic interactions and charge compensation will result in tighter binding. Molecules can be designed to satisfy these conditions, or selected from a large pool of available libraries and the binding strengths can be experimentally measured. As an alternative or in conjunction with the experimental methods, computational methods can be used to design and calculate the binding energies. Sophisticated computer graphics programs are available for visual analysis of the target–ligand interactions. Computational methods will also estimate the pharmacological properties of various designed molecules. The inhibitory efficiency can be progressively improved by altering or adding atoms or functional groups to the inhibitor. This process is guided by the available structural information of the targets. Thus one or more potential drug molecules can be designed using the structure of the target molecule as the scaffold.

Traditional drug design was either mechanism-based or an outcome of the experience of using natural products for treatment. Combination of intuition, ingenuity and attempts to find the potential of available molecules as drugs and chance discoveries such as that of penicillin, contributed to the availability of several drugs. Later methods involved high throughput screening of several thousand compounds against single targets. Structure-based drug design began more than 30 years ago and has flourished since then. The initial lead compounds are usually natural substrates or reaction intermediates or other molecules known to interact with target proteins.

Traditional drug design was either mechanism-based or an outcome of the experience of using natural products for treatment.

Molecules that are designed based on structures cannot be used directly as drugs; but several other properties need to be optimized, such as bioavailabilty and solubility. Drugs should have the ability to pass through membranes to reach the target site and should

recognize and specifically bind to the target and modulate its function. The drug should be chemically synthesizable and cost effective, should be accepted by the body and retained. The drug should not bind to host molecules and interfere with essential cellular processes of the host which leads to toxic side effects. The suitability needs to be investigated by rigorous clinical trials before the drugs are made available in the markets. Structure-based design speeds up the initial process of selecting the lead compounds to be tested, in contrast to traditional methods which require random screening of thousands of compounds. A few examples of how crystal structures helped in the development of drugs are presented here.

The molecules designed based on structures cannot be used directly as drugs but several other properties need to be optimized, such as bioavailabilty and solubility.

HIV Drugs

A previously unknown disease which appeared in the early 1980s and was called AIDS (acquired immunodeficiency syndrome), was identified to be caused by a virus named later as the human immunodeficiency virus (HIV). Since it caused severe infections by destroying the body's immune system with no available medication, patients of this deadly disease did not escape death until treatment was made available a few years later. The life cycle of the virus and the mode of infecting humans were worked out soon after the identification of the virus. Sequence of its genome revealed the proteins present in the virus. Structure investigations of these proteins were taken up in quick succession.

The first structure to be determined was that of an aspartyl protease (in 1989) which is essential for the replication and maturity of the virus. Cleavage of the newly transcribed polypeptide into functional proteins is performed by the protease. Inhibition of the viral protease (HIV PR) results in immature virus particles that are incapable of infection. The structure of HIV PR is similar to that of pepsins though it is smaller in size with a very low sequence similarity. The fold and the catalytic site architecture are remarkably similar to those of pepsins. The function of HIV PR was worked out and the substrates identified. Also available by that time were the structures of several pepsins and

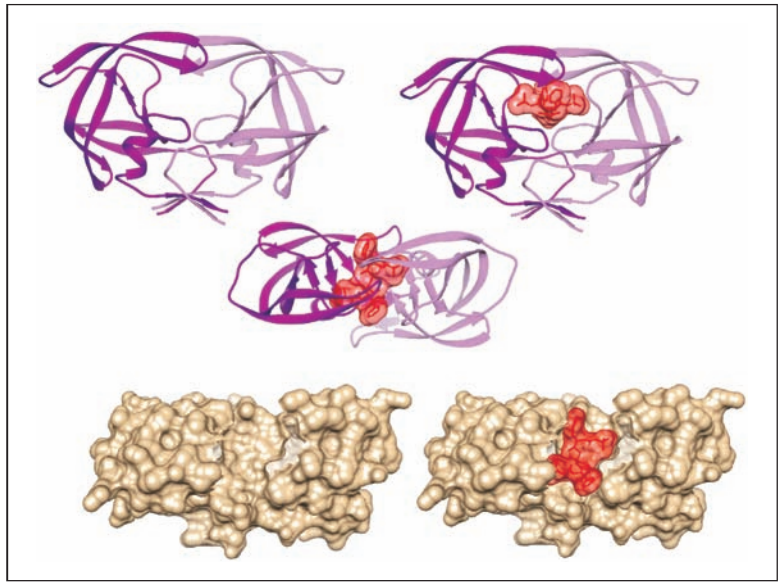

Figure 2. HIV protease in complex with its inhibitor in different views and representations (PDB code: 1odw).

their complexes with inhibitors. This prior knowledge was of immense help in designing inhibitors to HIV PR and developing them into effective drugs in a significantly short time frame [1]. Within a few years, crystal structures of several complexes of HIV PR with a variety of inhibitors were determined (*Figure* 2), each structure providing new insights to design a better drug. These drugs were made available to AIDS patients with the first approved drug, saquinavir, coming in 1995; and by 1998, four drugs were made available and several followed later. The drugs were approved in just about 6 years after the structure of HIV PR was determined, whereas, it usually takes 10–15 years for the entire process. Though structure-based drug design in this case started with lead compounds based on the natural substrates (containing Phe-Pro, Tyr-Pro sequences), they had to be significantly modified before effective drugs could be made. Careful examination of enzyme–ligand interactions, subsites and water structure in the binding site led to the modification/introduction of several types of functional groups including a 7-membered ring. At each stage of the development, crystallography, binding energy measurements and computations were carried out extensively. In a treatment with 'cocktail drugs', HIV PR inhibitors were given in combination with drugs which are inhibitors to another enzyme of the virus, the

reverse transcriptase. This combination therapy called HAART (highly active antiretroviral therapy) with these antiretroviral drugs not only prolonged the life of AIDS patients but also improved the quality of their lives. This remarkable success of drug development for the treatment of AIDS paved the way for many similar future endeavours.

The structure of reverse transcriptase was determined in 1992. Though the treatment targeting this enzyme had started much earlier with an anti-cancer drug AZT (azidothymidine), the structure provided the platform for many drugs developed later. The antiretroviral treatment, though not a complete cure, certainly changed the fate of many patients from certain death to a life expectancy close to that of unaffected people.

Side effects and multidrug resistance (discussed later) that appeared during the treatment with these drugs called for improved efficiency of the existing drugs as well as search for new strategies for alternate drug development. This led to ways to intervene with fusion and entry of the virus to the host cells. By this approach, the virus is inhibited extracellularly, i.e., before entering the host cell. One of the surface glycoproteins on the virus, gp120, gets attached to the cell, enabling the release and fusion of another glycoprotein gp41 with the host cell. The advantage of targeting gp41 is that there is no human analogue to this protein. After fusion, gp41 attains a novel 6-helical bundle form whose structure was determined. A peptide-based drug molecule, enfuvirtid, designed based on the gp41 sequence to prevent the formation of the helical structure, acts as a successful fusion inhibitor. Gp120 binds to CD4 T-cells and coreceptors CCR5 and CXCR4 on the cell surface. A drug, maraviroc, designed to bind to CCR5 prevents attachment of gp120 to the cell by blocking the entry of the virus to the cell. The crystal structure of the complex between CCR5 and maraviroc reveals the structural basis for inhibition of the viral entry. A successful drug, raltegravir, was also developed to inhibit the action of the third enzyme of the virus, integrase, guided by the structure of integrase. At present, there are more than 20 drugs approved by FDA, mostly HIV PR

Side effects and multi-drug resistance call for improved efficiency of drugs as well as search for new strategies for alternate drug development.

New improved drugs have better tissue penetration and stay longer in the body.

and reverse transcriptase inhibitors, and one each to inhibit integrase, viral entry and fusion.

Hypertension

Hypertension is a major health concern all over the world as it leads to cardiovascular diseases, with one in 3 adults being affected at present. Two enzymes, renin and angiotensin-I converting enzyme (ACE), of the renin-angiotensin system of the reaction pathway that regulates blood pressure in the body are being probed as drug targets. Renin catalyzes the cleavage of the protein angiotensinogen to yield a decapeptide angiotensin-I, which, in turn, is cleaved by the enzyme ACE, which produces an octapeptide angiotensin-II, a potent vasoconstrictor involved in controlling the blood pressure in the body. Blocking the functions of renin or ACE maintains blood pressure and prevents hypertension. Renin is an aspartic proteinase similar to pepsins. Since a wealth of information was available on pepsin structures in their native form and with a wide variety of inhibitors, significant work was carried out to design inhibitors based on its structure. Though many ligands were shown to bind to the enzyme, there was limited success in developing them as effective drugs. After many years of extensive efforts, one drug, aliskiren, has made it to the market.

Structural information of ACE has become available since 2003 [2]. ACE is a zinc-dependent dipeptidyl carboxypeptidase. Even before the structure of ACE was determined, ACE inhibitors were designed based on the reported structures of carboxypeptidase A and thermolysin. Hypertension medication such as captopril which resulted from these efforts and has been in use for decades. Since captopril treatment caused side effects like lack of taste and skin rashes, non-sulphurous compounds such as enapril and lisinopril were developed. The structures of ACE provided insights into the hydrolytic mechanism of the enzyme and showed the basis for binding of these first generation drugs (*Figure* 3). ACE has two domains, both of which bind to inhibitors but with minor differences. Examination of the binding interactions led to modifications

resulting in drugs with less side effects. Further modification with the incorporation of selenium produced drugs which act as, antioxidants in addition to being antihypertensive, as hypertension and oxidative stress were found to be related. Differences between the domains were exploited to design domain-specific inhibitors. All these efforts resulted in about ten drugs which are being used extensively. The new improved drugs have better tissue penetration and stay longer in the body.

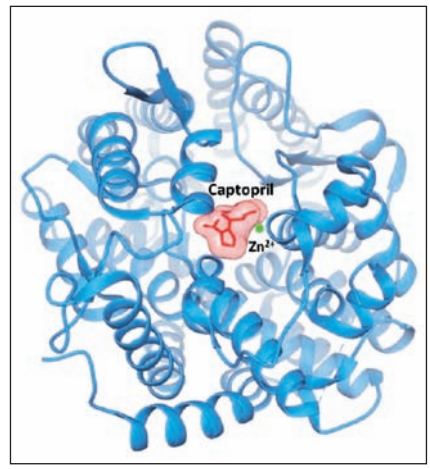

Figure 3. Angiotensin converting enzyme in complex with captopril (PDB code: 1uzf).

Targets of Aspirin

The story of aspirin dates back to around 400 BC. The bark and leaves of the willow tree were prescribed by Hippocrates to treat pain and fever and were later found to contain salicin which was responsible for the pain relief. The compound was modified to acetylsalicylic acid which is called aspirin and has been in use for the treatment of pain, inflammation and fever since the late 1890s. Is spite of its long known existence as a successful pain reliever, the mechanism of its reaction was not known until its receptor enzymes– cyclooxygenases – COX-1 and COX-2 were identified in 1970s. Cyclooxygenases, (also called prostaglandin endoperoxide H_2 synthases), catalyze the biosynthesis of prostaglandins which are involved in the signalling pathway related to pain and inflammation response. The structure of the enzyme has been determined in the presence of aspirin (3). The acetyl group that gets transferred from aspirin and covalently attached to Ser530 of the enzyme and also the bound leaving group, salicylic acid, were clearly visible in the structure (*Figure* 4). The entry to the active site through the tunnel gets blocked in the complex due to the presence of the product (salicylic acid) and by the conformational change of the side chain of Ser530. Serine is not a catalytic residue as a Ser $\rightarrow$ Ala mutation does not affect the activity but has no aspirin sensitivity. A number of drugs have been developed from the structural information with a view to reduce side effects such as gastric bleeding which were caused by some of the drugs already in use. Some of the new drugs are also effective against arthritis.

The story of aspirin dates back to around 400 BC. The bark and leaves of the willow tree were prescribed by Hippocrates to treat pain and fever.

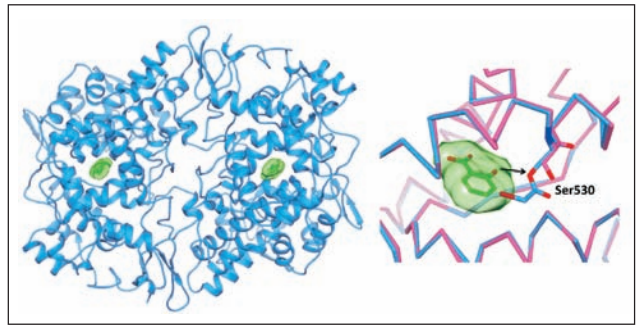

Figure 4. Cox-1 structure with bound salicylic acid. A close-up view of Ser530 in the native structure (pink, PDB code: 1prh) and the acetylated Ser530 in the complex (blue, PDB code:1pth) are shown.

Several structures subsequently determined showed small but significant differences in the structures of Cox-1 and Cox-2. Cox-1 is constitutively expressed in the body, whereas, Cox-2 is expressed at the site of inflammation. Although, aspirin binds to both, it has a higher affinity to Cox-1 and also has side effects such as gastric bleeding. It was hypothesized that Cox-2 specific drugs would have significant advantages over aspirin. The structures revealed an additional cavity in Cox-2 leading to the development of Cox-2 specific inhibitors which are now available in the market. A setback to these efforts was seen when some of these inhibitors showed severe side effects such as increasing incidences of heart attacks which lead to the withdrawal of some of these drugs. The overlap of Cox-1 and Cox-2 functions and the observation that Cox-2 is also constitutively expressed in other tissues may be responsible for the adverse effects. Design of alternate drugs and with better properties is underway.

Influenza Treatment

Influenza virus causes flu such as bird flu and swine flu. Major pandemics have occurred a few times, the most recent ones being the 1997 bird flu and the 2009 swine flu threats. Seasonal outbreaks of influenza occur every year. Thus, both seasonal and pandemic diseases are prevalent, resulting in 250,000 to 500,000 deaths annually because of the many variants of the virus such as H1N1 and H5N1. Neuraminidase is a surface glycoprotein of the influenza virus. The virus gets attached to the host cell surface by the interaction of its heamagglutinin with sialic acid present on the surface proteins of the host cell. Neuraminidase catalyzes the

Design of alternate drugs and with better properties is underway.

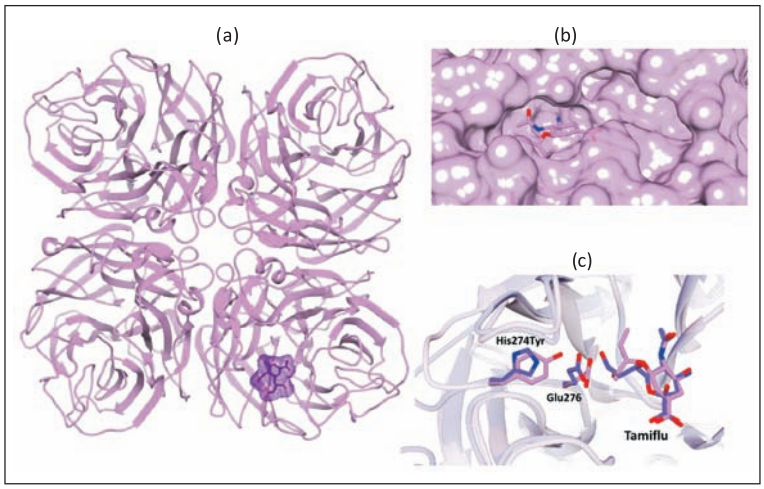

Figure 5. (a) Neuraminidase in complex with tamiflu. (b) Additional cavity near the binding site of tamiflu.(c) Native (blue, PDB code: 2hu0) and His274Tyr mutant (pink, PDB code: 3cl0) structures

cleavage of sialic acid from the surface of the host cell, thus releasing the virus and enabling it to spread in the host. Its structure elucidation in 1983 (4) and subsequent design of drugs based on its substrate/reaction intermediate led to the successful development of the drugs, tamiflu and zanzamivir. In the swine flu pandemic caused in 2009 by H1N1 virus, several lives were saved by these drugs. The enzyme is a box-like homotetrameric molecule with 4-fold symmetry (*Figure* 5a). Structures of complexes reveal the mode of inhibition and show an additional cavity around the inhibitors (*Figure* 5b). Later drugs were designed by addition of groups that occupy this cavity, known as the '150-cavity', located at the binding site.

Multidrug Resistance

Finding an effective drug to treat any disease involves enormous efforts. Once the drug is released to the market and positive results are observed, one may think that the task is successfully completed and the drug will be used by generations to come and the war against the pathogen has been won by humans. But, what happens in reality is different. Frequent use of drugs leads to resistance by the pathogen. Also, resistance is developed not just for one drug, but for a number of them simultaneously. This phenomenon is called multi drug resistance (MDR). Reports of MDR against several drugs used for the treatment of malaria, TB,

Finding an effective drug to treat any disease involves enormous efforts.

Drug design assisted by structures showed remarkable success in the last 25 years and will continue to contribute for the development of safe and affordable drugs to treat several diseases.

HIV, flu have appeared, indicating that it is not restricted to any particular drug or organism but is all-pervasive. It appears as if the pathogens have decided to fight back. Multiple mechanisms for MDR such as overexpression of efflux pumps, induction of alternate pathways, mutation of key residues on the target molecule, are known. A brief account of one of the strategies, i.e., mutations in the target protein will be presented here. The mutations happen in such a way that the proteins retain their activity but reject the drug even though a majority of the drugs bind at the active site. How they devise such intelligent strategies to evade being inactivated, is a marvel, though an annoying one, and poses a tough challenge as the drugs become ineffective, calling for the need for new targets and/or new strategies.

Crystallography again is helpful in providing the structural basis for MDR. One example is shown in *Figure* 5c. The structure of neuraminidase of a drug- resistant strain of influenza virus with His274Tyr mutation was determined as a complex with tamiflu. Due to the larger size of the side chain of the mutated residue, there is a small but significant shift of tamiflu from the interacting glutamate residue. This movement results in a 300-fold drop in the affinity of tamiflu to the enzyme. Several such structures of target proteins from MDR strains are available like those of HIV PR. The structures certainly provide clues for improving or redesigning the inhibitors. In the case of HIV PR, about half of its residues were found to be mutated. Another interesting observation is that a number of them were located away from the active site, not in direct contact with the inhibitor. Superposition of all the structures with inhibitors and substrates led to the delineation of the 'substrate envelope'. This is being considered for future drug design as the envelope surrounds the substrates but the inhibitors appear to be protruding out of it, at the site where the virus is mutating (5).

Conclusions

Several crystal structures are available and a large number of drug targets for many diseases such as cancer, Alzheimer's and diabe-

tes have been identified. Debilitating diseases like malaria and tuberculosis kill hundreds of thousands of people every year but no new drugs have been produced except for one tuberculosis drug which was approved in 2012 after 40 years, in spite of serious efforts at drug development. Recent structures of ribosomes in complex with antibiotics such as streptomycin and tetracyclin and the new structural information on G-protein coupled receptors have opened avenues for designing new drugs.

New structural information on G-protein coupled receptors have opened avenues for designing new drugs.

Drug design is a complex, expensive and time-consuming process. Structure-based drug design helps in improving the efficacy of the lead compounds from micromolar to subnanomolar levels and reduces the number of compounds to be screened tremendously, thus saving time and resources, but is by no means the complete process for bringing out the final product. Drug design process is a multidisciplinary approach. Many obstacles have to be crossed to convert the lead molecules to usable drugs. Multidrug resistance and side effects should be dealt with for any drug being used for treatment. Alternate approaches to overcome the problems have to be developed. Drug design assisted by structures has shown remarkable success so far and will continue to contribute for the development of safe and affordable drugs to treat several diseases.

Suggested Reading

[1] **A Wlodawer and J Vondrasek, Inhibitors of HIV-1 protease: A major success of structure-assisted drug design, *Annu. Rev. Biophys. Biomol. Struct.*, Vol.27, pp.249–84, 1998.**

[2] **J F Riordan and M R Ehlers, K R Acharya and E D Sturrock, Ace revisited: a new target for structure-based drug design, *Nat. Rev. Drug Discov.*, Vol.2, pp.891–902, 2003.**

[3] **P J Loll, D Picot and R M Garavito, The structural basis of aspirin activity inferred from the crystal structure of inactivated prostaglandin H_2 synthase, *Nature struct. Biol.*, Vol.2, pp.637–643, 1995.**

[4] **J N Varghese, W G Laver and P M Colman, Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 Å resolution, *Nature*, Vol.303, pp.35–40, 1983.**

[5] **M N L Nalam and C A Schiffer,New approaches to HIV protease inhibitor drug design II: Testing the substrate envelope hypothesis to avoid drug resistance and discover robust inhibitors, *Curr. Opin. HIV AIDS*, Vol.3, pp.642–646, 2008.**

Address for Correspondence
K Suguna
Molecular Biophysics Unit
Indian Institute of Science
Bangalore
Email:
suguna@mbu.iisc.ernet.in