# ABSTRACT

X-ray crystallography is a powerful tool to obtain atomic resolution details on the three-dimensional structure of biological macromolecules. Crystallographic techniques find wide applications in understanding macromolecular assembly, enzyme mechanism, mode of activation of enzymes, substrate-specificity, ligand-binding properties, domain movement, etc. Crystallographic studies are essential for understanding the intricate details of the mechanism of action of enzymes. Understanding the subtle differences between the parasite and host enzymes is necessary for the design of inhibitor molecules that specifically inhibit parasite enzymes. This facilitates drug discovery against the pathogen by blocking the action of crucial enzymes.

The current thesis deals with the application of crystallographic techniques for understanding the structure and function of proteins from two different types of pathogenic organisms – a plant virus *Sesbania Mosaic Virus* (SeMV), and the malarial parasite *Plasmodium falciparum* (Pf). The main objectives of the work include structural studies on the major non-structural proteins of the virus, namely protease and VPg domains of the polyprotein, and on two key enzymes of the purine salvage and glycolytic pathways of the parasite, namely hypoxanthine guanine phosphoribosyl transferase (HGPRT) and triosephosphate isomerase (TIM), respectively. The thesis has been divided into eight chapters, with the first four chapters describing the work on the nonstructural proteins of SeMV, while the rest of the chapters deal with the studies on the plasmodial enzymes.

The thesis begins with **Chapter 1** containing an introduction to viral proteins, with special emphasis on viral proteases. After a brief and general introduction about viruses and viral proteins, a description of the work carried out on SeMV till date is described. This is followed by a detailed description of viral proteases and their mechanism of action. SeMV protease is a serine protease of the trypsin fold, and hence a description of the available crystallographic studies on viral proteases of the trypsin fold is included. An overview of the structural studies on the non-structural proteins VPg (viral protein genome-linked) and RdRp (RNA dependent RNA polymerase) also forms part of Chapter 1.

**Chapter 2** deals with the structure determination of the protease domain of the polyprotein of SeMV. The structure was determined by multiple isomorphous replacement and anomalous dispersion (MIRAS) due to the non-availability of a homologous structure as a model for phase determination by molecular replacement. Four isomorphous mercury derivatives were obtained on soaking in four different mercury compounds namely mercuric chloride, mercuric iodide, thiomersal and *p*-chloro mercury benzene sulphonic acid (PCMBS). Both the isomorphous and

anomalous signals obtained from the heavy atom derivatives were useful in phasing. The crystal structure showed that SeMV protease is indeed a serine protease of the trypsin fold with His181, Asp216 and Ser284 as the catalytic triad. A brief description of the steps involved in experimental phasing is also included in Chapter 2.

The structure determination of SeMV protease facilitated its comparison with other viral proteases of the trypsin fold, and identified its unique features, which make it closer to the cellular and bacterial proteases. The presence of an exposed stretch of aromatic residues on the crystal structure of the protease helped in identifying the region of interaction between protease and VPg. This stretch of aromatic residues was implicated to play a role in the activation of the protease by VPg, and explained the special spectral properties observed in the fusion protein of protease and VPg domains. In addition to these, further characterization of the interaction between protease and VPg using NMR (nuclear magnetic resonance) spectroscopy, and the attempts for crystallization of protease-VPg fusion protein are described in **Chapter 3**. The application of the biophysical techniques such as gel filtration and dynamic light scattering (DLS) in monitoring the probability of crystallization of the various mutants of protease-VPg designed by rational mutagenesis for facilitating crystallization is also included in this chapter.

SeMV protease is a glutamate specific protease as it cleaves between glutamate and threonine or glutamate and serine residues occurring in the polyprotein. **Chapter 4** deals with the substrate specificity studies on SeMV protease. The features of the glutamate-binding pocket identified from the crystal structure, and a mutational analysis of the relevant residues are described. The initial attempts on substrate-specificity studies using synthetic peptide substrates were not successful, as the cleavage of the peptide could not be observed by mass spectrometric analysis. The attempts for obtaining crystals of protease complexed with the peptides were also not successful. The lack of binding of the peptides to the protein was demonstrated by isothermal titration calorimetry (ITC) experiments. The length of the peptide on either side of the scissile bond must be optimized by further studies.

Malaria is a global disease infecting several million individuals annually, particularly in the developing countries. Development of drug resistance and absence of a preventive vaccine have led to an immediate necessity for identifying new drug targets to combat malaria. The structural biology approach towards design of novel drug targets against the parasite involves the identification of key differences in the structures of the human and parasite enzymes and the determination of unique protein structures essential for parasite survival. **Chapter 5** begins with a description of the key developments in structural biology of plasmodial proteins in the recent past, and in structure-based drug discovery against the parasite. The enzymes of two key metabolic pathways of the parasite, namely purine salvage and glycolytic pathways, serve as

potential drug targets. The structural information available on these pathways from the parasite is summarized. This is followed by an introduction to the enzyme's hypoxanthine guanine phosphoribosyl transferase and triosephosphate isomerase. The reaction mechanism of the two enzymes and their structural characterization carried out till date, and the objectives of the current study are described.

The crystal structures of both the human and PfHGPRTs are available. Hypoxanthine and guanine act as substrates for the human enzyme, while the Pf enzyme can accept xanthine also as a substrate. Structural studies on a chimera of human and Pf HGPRTs, DS7 (Domain Switch 7), explained in **Chapter 6**, were initiated with the objectives of explaining the differences in the substrate-specificity between the human and Pf enzymes, and the oligomerization properties of the chimera. The crystal structure shows that the dimer interface maintained in the chimera is the canonical AC interface. The disruption of the AB interface could be responsible for the non-optimal binding of GMP, and hence for the decrease in the activity of the chimera. A comparison between the parent HGPRT structures and the chimeric structure provides an explanation for the shift in oligomerization properties of the chimera on the addition of PRPP. Replacement of Tyr197 (human HGPRT) with Ile contributes to the loss of interactions at the AB interface in the chimeric enzyme in the absence of PRPP, while the conservation of the interacting Pro93 and His26 could restore the tetrameric state in the PRPP-bound state. The investigations led to the identification of key differences between the human and Pf enzymes at the catalytically significant AB interface, which could be explored for designing novel inhibitors against the parasite enzyme.

The catalytic tools of PfTIM include the active site residues Glu165, His95 and Lys12 and the loop 6 which closes over the active site to stabilize the reaction intermediate. Structural studies on Phe96 mutants of triosephosphate isomerase from *Plasmodium falciparum* (PfTIM) were carried out in order to understand the role ofPhe96 in determining the open or closed state of loop 6 of PfTIM. It is believed that the higher propensity of the loop open state in the PfTIM crystal structures determined earlier was due to the presence of Phe96 instead of a serine, which is the corresponding residue in TIMs from other organisms. The liganded and unliganded crystal structures of three mutants, F96S, F96H and F96W of PfTIM, are described in **Chapter 7**. Mutation of Phe96 to S or H resulted in decrease of affinity for the ligand at the active site. Decrease in affinity could be due to the differences in the water structure connecting residue 96 to Ser 73 in the vicinity of the active site. The F96W mutant exhibits an increased probability of the loop closed state than the wild type enzyme. The soaking of the crystals in high concentration of the ligand led to the identification of a novel binding site for the ligand at the interface in both the wild type and mutant PfTIMs. An analysis of the crystal contacts is also described in the chapter to consider the role of crystal contacts in directing the position of loop 6.

The dimeric TIM from *Plasmodium falciparum* was compared with the tetrameric TIM from the thermophilic archaea *Methanocaldococcus jannaschii* (Mj) in order to probe the relationship between quaternary structure and stability. **Chapter 8** describes the structure determination of MjTIM from a tetartohedrally twinned crystal. A comparison between the dimeric and tetrameric TIMs shows that there are insertions in the dimeric TIMs in the vicinity of the tetramer interface. MjTIM achieves compactness due to the absence of aromatic residues in the outer barrel and by the shortening of loops. One of the unique features of the crystal structure of MjTIM was the disorder of loop 6, which could be attributed to a missing salt bridge between residues at the N- and C- terminal ends of the loop.

The thesis concludes with a brief statement of the **Future perspectives** of the various studies carried out. The various crystallographic, biochemical, and biophysical techniques employed in the investigations are described briefly in an **Appendix** to the thesis.