Chapter 4:

DISCUSSION
The shikimate pathway is found in bacteria, fungi, and plants but is absent in animals. Chorismate mutase (CM, EC 5.4.99.5) is at the branch point of the shikimate pathway. This enzyme carries out the rearrangement of the enolpyruvyl side chain of chorismate to form prephenate. Prephenate is further processed to form phenylalanine and tyrosine. This property makes CM a potential target for the development of antibacterial and antifungal agents and also herbicides. Chorismate mutase enzymes from different sources have been characterized and structure of many of them have been solved. Recently, three different groups have independently solved the crystal structure of *MtCM. Structurally, *MtCM is an all α-helical dimeric protein which is similar to the AroQ class of chorismate mutases from other sources like *E. coli* and *S. cerevisiae.

*MtCM*, encoded by Rv1885c, has an N-terminal signal sequence by virtue of which it is secreted out of the cell as mature protein (*MtCM*) in which the first 33 amino acid residues are deleted. The mature protein is termed as *MtCM*. The secretory CM from *Erwinia herbicola*, *Salmonella typhimurium* and *Pseudomonas aeruginosa* contribute towards the virulence of the respective pathogens. Based on this, a role in *M. tuberculosis* virulence or intracellular survival has been postulated for *MtCM* (Qamra et al., 2006). The functional significance of the enzymatic activity of secreted *MtCM* is not clear because the shikimate pathway is limited to the cytoplasm of the bacteria and *MtCM* is the only secreted enzyme of this pathway. The other enzymes of the shikimate pathway are not secreted and lack signal sequence. For some reason, *M. tuberculosis* may also secrete chorismate into the phagosomal environment and then reabsorb the product prephenate of CM. One perhaps needs to think on evolutionary lines to understand this. It is also possible that the secretion of CM could have evolved in some ancestral strains of *M. tuberculosis* even before it became human pathogen. The secreted CM protein may have further evolved to have a role in virulence, rather than to a functional significance based on its enzymatic activity. However, we have studied both the aspects i.e. activity and stability in phagosomal environment and designing of novel inhibitors of *MtCM*.

*M. tuberculosis* depends on its ability to survive in the phagocytic vacuole of the macrophage where it encounters an acidic environment. *M. tuberculosis* limits acidification successfully to some extent by utilizing transport systems which exchange protons for cations, resulting in increased levels of these cations within the phagosome.
Because *MtCM is secreted, its activity and stability would be affected by the phagosomal environment. Therefore, we have carried out an extensive and systematic analysis of the effect of pH and cations on the activity, structural stability and dimeric form of the *MtCM protein by using a combination of enzymatic activity assay, CD and fluorescence spectroscopy, and size exclusion chromatography. Our studies conclude that *MtCM is highly stable under physiochemical conditions mimicking the phagosomal environment.

All the crystal structures of *MtCM show it to be an all α-helical dimeric protein containing 6 α-helices and a single α-helical turn (PDB ID 2F6L, 2AO2, 2FP1/2FP2). Each monomer contains only one active site pocket, which is located near to the N-terminus of the protein. The consensus dimer interface from the three crystal structures of *MtCM comprises the C-terminal halves (Residues 103-118) of the H3 helices from the two monomers, which are arranged in an anti-parallel fashion. The extensive hydrophobic interactions, which have been suggested to be primarily the driving force for dimerization, are observed between A107 and Y110 of one monomer with L118 and F113 of the other monomer, respectively. In addition, hydrogen bonding networks are observed involving the guanidino group of R103 of one monomer with the main chain carbonyl oxygen of K117 of the other monomer, and between the side chain carboxyl oxygens of E68 and E106 of one monomer and the $N^\delta$ group of K117 of the other monomer. In addition, there are a significant number of water-mediated interactions near the interface. The dimer interface buries a surface area of 837 Å$^2$ per monomer in one of the structures (PDB ID: 2FL6). These structural aspects are taken into consideration for discussing the observed dependence of activity and stability of *MtCM on pH and cations.

Tryptophan fluorescence emission maximum is a very sensitive marker for the activity and stability of the protein. As described above, two of the buried tryptophan residues (W61 and W116) are positioned close to the active site as well as the dimer interface. We suggest that $\lambda_{\text{max}}$ can be utilized to study the effect of various agents at very low concentrations. This can be confirmed by site directed mutagenesis experiments.
4.1. BIOPHYSICAL AND BIOCHEMICAL CHARACTERIZATION OF *MtCM

4.1.1. Effect of pH

*MtCM has maximum activity in the range of 5 to 6.5. In this range, the protein has highest percentage of secondary structure, thermodynamic stability and compactness of secondary structure.

*MtCM is an acidic protein having an isoelectric point of 4.93 and total charge over a protein monomer of -6. For having an insight into the mechanism of cation- or pH-induced compaction of native conformation of *MtCM, we determined the electrostatic potential of *MtCM by GRASP (Nichollas, 1992) from its crystal structure (PDB ID 2F6L Figure 4.1). The protein surface shows significant clustering of negative charges which could contribute towards repulsive long-range electrostatic interactions. Electrostatic interactions among charged moieties of ionizable surface residues contribute significantly to the conformational stability of proteins. The dependence of protein stability on pH and cations would be affected by its net charge as well as by the positioning of the charged residues. This would hold especially for the highly charged proteins in which long-range electrostatic interactions can contribute to the stability of the native state. In the case of *MtCM, an extended water structure may be involved in shielding of the repulsive electrostatic interactions between negatively charged residues at or close to the dimer interface, for maintaining the relatively open conformation and structural stability of the protein. This would explain the compaction of protein structure observed at mildly acidic pH or at high concentration of cations except for Ca\(^{++}\) and Cu\(^{++}\), where destabilization may be caused because of specific interactions as described below.

The pH optimum for activity of *MtCM observed by us closely matches with the results described by Sasso et al and differs from the results described by Prakash et al. A closely similar dependence of the enzyme activity on pH has been observed for ScCM, which has the highest activity in the range of 5.5 to 6.5. On the other hand, the pH optimal for E. coli is in the range of 6.5 to 8.5. It has been suggested that the pH dependence of ScCM is associated with E246 in the catalytic site. The corresponding residues in the other hand, is S84 and Q88. In comparision *MtCM has E109 at
Figure 4.1: Molecular surface of *MtCM colored according to its electrostatic potential (blue for positive potential and red for negative potential).
corresponding position. All other AroQ class of chorismate mutases with glutamate at corresponding position to *MtCM's E109 also show optimal activity at lower pH, reflecting the requirement of the hydrogen bond from glutamate to the other oxygen of the ligand for efficient catalysis. Therefore, for *MtCM a pH optimum range similar to ScCM, as reported by us and Sasso et al seems to be more reliable.

4.1.2. Effect of Cations

The basis of the effect of NaCl seems to be similar to that of mildly acidic pH, i.e. screening of the repulsion between negatively charged groups. The stabilization most likely comes from the summation of several ‘small’ contributions of $\Delta G$ from the interaction of NaCl at several different sites on the protein surface. Currently, it is difficult to pinpoint it to any one or two particular residues. However, one can think of comparing the high $T_m$ for unfolding for *MtCM in the presence of 1.5 – 2.0 M NaCl with a CM protein from thermophillic bacteria which may have similarly high $T_m$ of unfolding even at low salt concentration. A comparision of protein sequence and structure would provide the clues to understand of the stabilizing effect of NaCl. The NaCl induced compaction of the protein may, in turn, lead to the observed enhanced cooperativity in its thermal stability and stabilization with reference to other denaturants. A similar stabilization and compaction caused by NaCl is observed in the case of other acidic proteins like glucose oxidase (Ahmad et al., 2001 and Akhtar et al., 2002) and RNase T1 (Mayr and Schmid, 1993). Slight decrease in the activity of *MtCM, in the presence of NaCl, may be credited to neutralization of negative charge residues E109 and E106 present in the active site. A similar decrease in activity has been reported for glucose oxidase by Ahmad et al., 2001 and for bovine pancreatic deoxyribonuclease A by Poulos and Price, 1974. Other monovalent cation K$^+$ shows an effect similar to that of Na$^+$ by slightly decreasing the activity.

Divalent cations Mn$^{++}$, Zn$^{++}$ and Co$^{++}$ does not cause compaction of the protein. Ca$^{++}$ and Cu$^{++}$ show characteristic and similar patterns of destabilization. Since other divalent cations donot destabilize *MtCM, it is clear that the effect of cations cannot be explained on the basis of monovalent and divalent differences. It is likely that the
coordination shell of Ca and Cu are interacting with similar sites of protein. Most probably, the destabilizing effect of Ca\textsuperscript{2+} and Cu\textsuperscript{2+} may be because of breaking of the hydrogen bonding interactions involving the glutamic acid residues which are responsible for structure and dimeric form. There also seems to be an electrostatic component for destabilization which is shielded to some extent in the presence of high concentration of NaCl. This needs to be further verified by MD simulation and mutagenesis.

4.1.3. Effect of denaturants

Guanidine hydrochloride and urea are the most common chemical denaturants that are used for protein denaturation. The conformational stability of multimeric proteins can be measured by equilibrium unfolding studies in urea or GdmCl solutions.

*MtCM starts losing its structure and activity at even very low concentration of urea and GdmCl with monophasic transition. C\text{m} of *MtCM for urea is \approx 1.4M and for GdmCl, it is \approx 0.75M and C\text{m} (Urea)/ C\text{m} (GdmCl) is \approx 1.86 which is not very common for dimeric proteins as it has been suggested that multimeric proteins are more susceptible to GdmCl denaturation (Akhtar \textit{et al.}, 2002). This ratio is low in the case of *MtCM because of its similar susceptibility for both urea and GdmCl.

The destabilization of the structure of *MtCM at low concentrations of urea appears to be because of weakening of hydrophobic interactions in the dimer interface. Further, there seems to be a partially unfolded intermediate at 1.5 M urea concentration. On the other hand, GdmCl is fully dissociated into Gdm\textsuperscript{+} and Cl\textsuperscript{−} at pH 6.5. It is likely that at low concentrations, before its chaotropic effect becomes predominating, GdmCl could actually stabilize the structure just like NaCl. However, in GdmCl mediated denaturation, no intermediates are observed.

4.2. INHIBITOR DESIGNING AND VALIDATION

The aza inhibitors and some other inhibitors like endo oxabicyclic dicarboxylic acid are the well known inhibitors of CM enzymes with activity against chorismate mutase from \textit{E. coli} (EcCM) (Hediger, 2004 and Chao and Berchtold, 1982). *MtCM has low
homology with EcCM. We have performed ligand-based virtual screening of compounds from CDRI in-house library based on pharmacophore model and by taking drug likeness properties i.e. hydrogen bond donors are <5, hydrogen acceptors <10, relative molecular weight <500, lipophilicity (logP)<5., and the compound will probably be orally bioavailable. We have identified new lead candidates that possess inhibitory activity against *MtCM. These identified compounds were further screened for in vitro enzymatic activity assay. Chorismate mutase activity assays identified four compounds as inhibitors of *MtCM with low $K_i$ values.

We analyzed the docking modes of active compounds I and II, in the active site of *MtCM. The inhibitors appear to fit well in the active site pocket, showing several hydrogen-bonding and vander Waals interactions with the binding site residues. The binding conformation of two best active compounds is shown in Figure 4.2.

The oxygen atoms of sulfate group of compound I, corresponding to features AA1 and AA2 of pharmacophore model, acts as hydrogen bond acceptor for the side chain N-H of Lys 60 (Distance 2.4 Å) and side chain nitrogen of Gln 76 (Distance 1.956 Å). Side chain carbonyl group of Gln76 also makes hydrogen bond with nitrogen atom of compound 1 (Distance 2.292 Å). Similarly, oxygen atom of nitroate group of compound 1, which maps to AA3 of the pharmacophore model, interacts with guanidium nitrogen of Arg49. The results indicate that the nitro group substituted at 2nd position of phenyl ring plays important role in the tight binding with chorismate mutase, which is also revealed from the binding of compound II, where nitro group interacts extensively with side chain nitrogen of Arg49. However, the oxygen atom of these two inhibitors which corresponds to the AA4 feature of the pharmacophore model was not found to involve in any hydrogen bonding interaction.

Furthermore, most of the hydrophobic interactions between *MtCM and inhibitors were conserved within the docking mode of compounds I and II. Thus, Val56, Arg72, Ile102 and Ile137 were found to form hydrophobic contacts with molecules I and II.
Figure 4.2: Docked conformation of (a) I and (b) II in the catalytic site of MtCM X-ray crystal structure.