

SUMMARY OF THE THESIS

The N-terminal methionine excision (NME) in both prokaryotes and eukaryotes is an essential co-translational proteolytic process responsible for the diversity of amino-terminal of proteins. It is an irreversible reaction, which occurs soon after an N-terminal residue of the nascent peptide emerges from the ribosome exit tunnel before the commencement of protein folding. The enzyme involved in this process is known as methionine aminopeptidase (MetAP), which specifically removes the terminal methionine in all organisms, if the penultimate residue is non-bulky and uncharged. In addition to MetAP, in prokaryotes and eukaryotic organelles where the initiator methionine is formylated, NME requires the action of another metalloprotease, peptide deformylase. PDF removes the N-formyl group present on all nascent polypeptides synthesized. In fact, removal of N-formyl group is a prerequisite for the subsequent action of MetAP. Nevertheless, the MetAP action for exclusion of N-terminal methionine is mandatory in 50-70% of nascent proteins. Such an activity is required for proper sub cellular localization, additional processing, and eventually for the degradation of proteins. The work embodied in this thesis is an attempt to get insight about mycobacterial MetAP by performing the structure function analysis and to understand about its interaction with the components of ribosome, the protein synthesizing machinery of the cell.

Although bacteria generally have single MetAP1, mycobacterial genome reveals the existence of 2-4 putative genes for MetAP1; for instance *Mycobacterium tuberculosis* has two genes *mapA* (Rv0734) and *mapB* (Rv2861c) encoding *MtMetAP1a* and *MtMetAP1c* respectively. Although, crystal structure of *MtMetAP1c* was reported as an apoenzyme and in complex form, biochemical characterization of the two *MtMetAPs* has not been carried out in detail despite the presence of their human counterparts.

The presence of a short N-terminal extension in some MetAPs found in Gram positive bacteria is unusual. Therefore, it is interesting to know their phylogenetic placement. The phylogenetic analysis done with protein sequences from archaea, prokaryotes and eukaryotes revealed that sequences which can be categorised as MetAP1c (MetAPs with short N-terminal extension) group together and close to *E. coli* MetAP1a (MetAPs without extension). In fact, they have a very distinct clade

organization. Interestingly, MetAP1a from various Gram positive bacteria formed a separate cluster.

Further, to get better understanding about the *map* genes from *M. tuberculosis* their promoter analysis was performed *in silico*. Accordingly, the upstream regions of these genes were cloned in plasmid pSC301, for the construction of transcriptional fusions, transformed in *M. smegmatis* strain mc<sup>2</sup>155 and their activity was determined by monitoring GFP expression. Western blotting with anti-GFP antibody and measurement of relative fluorescence intensity of GFP in the cultures revealed that *mapA* promoter has higher activity compared to that of the *mapB* at all the stages of growth. This indicates that probably *in vivo* expression level of *MtMetAP1a* would be higher than that of *MtMetAP1c*. *Mycobacterium* resides in the macrophages upon infection and *in vivo* it is exposed to various stresses. To have an idea if *MtMetAPs* are sensitive to such conditions, early log phase cells were subjected to different conditions of stress. Monitoring the promoter activity revealed that under oxidative stress (cultures exposed to H<sub>2</sub>O<sub>2</sub>) activity of both the promoters enhances marginally indicating the two *MtMetAPs* could be modulated by oxidative stress conditions. It is well known that, *M. tuberculosis* while invading the host cells is exposed to slight decrease in the pH. Therefore, the effect of pH on the transcriptional activity of the *mapA* and *mapB* promoters was evaluated by transferring cells growing in media of pH 7 to that of pH 4 and subsequently monitoring the relative GFP level at different stages. However, the activity of both the promoters did not exhibit any alteration on exposure to acidic conditions. Interestingly, following exposure to heat stress, while there was no significant change in the promoter activity of *mapA*, the *mapB* displayed a 50% reduction in its GFP level compared to the one at 37°C. These results, therefore argue that the GFP level regulated by *mapB* promoter is affected *in vitro* in response to temperature stress. Thus it is apparent that both *mapA* and *mapB* promoters are active. This led to the assessment of enzymatic activities of both *MtMetAP1a* and *MtMetAP1c*.

The *mapA* and *mapB* genes were further cloned in pET-28c, histidine-tagged proteins were expressed and purified to carry out biochemical studies. An enzyme coupled assay system was employed to monitor the functionality of the two proteins *in vitro*. Different peptide substrates were used to assess the most preferred substrate.

It was observed that a tetra-peptide Met-Gly-Met-Met is the preferred substrate for *MtMetAP1a* while the tri-peptide Met-Ala-Ser was cleaved more efficiently by *MtMetAP1c*. Both *MtMetAP1a* and *MtMetAP1c* did not show any activity with substrate Met-Gly indicating that the dipeptide tested was insufficient for efficient hydrolysis by the mycobacterial MetAP1s. Further, on use of Gly-Gly-Ala as the substrate none of the *MtMetAPs* exhibited any activity, thereby validating the specificity of the two enzymes for the N-terminal methionine excision.

The evaluation of enzyme activity at different temperatures revealed that optimum temperature for the activity of both the metalloproteases is 30-37°C. Interestingly, *MtMetAP1a* was active even at 50°C while *MtMetAP1c* hardly retained any activity at this temperature. The far-UV CD spectra for both the enzymes were determined at 25°C and 50°C to monitor the effect of temperature on their secondary structures. It was observed that with increase in temperature *MtMetAP1a* was able to retain its secondary structure but *MtMetAP1c* was completely unfolded and as a result did not show activity at higher temperature. The far-UV CD studies thus suggested that *MtMetAP1a* is structurally more stable at high temperature compared to that of the *MtMetAP1c*. Thermal unfolding of the two proteins revealed that  $T_m$  of *MtMetAP1a* is 53.6°C while that of *MtMetAP1c* is 42.7°C. Such an observation also corroborates well with the findings that the decrease in *mapB* promoter activity compared to that of *mapA* upon exposure to heat stress. The kinetic analysis of *MtMetAP1c* revealed that it has ~180-fold more enzyme turnover rate with the same substrate compared to *MtMetAP1a*. It is interesting to mention at this juncture that the magnitude of *mapA* promoter activity was at least 9 fold higher compared to that of the *mapB*.

To get an insight into the cause of such differential behaviour of the two enzymes, structure of *MtMetAP1a* was modelled. Analysis of the modelled structure revealed that the active site residues in both the *MtMetAP1s* were essentially same (His-88, His-193, Asp-117, Glu-219, Trp-236 for *MtMetAP1a* and His-114, His-212, Asp-131, Glu-238, Trp-255 for *MtMetAP1c*), though their arrangement in space is different. Furthermore, these amino acids in the active site are crucial because any alteration in them for both the enzymes yielded inactive protein. Thus it is logical to

presume that this difference in the arrangement may have reflected in their enzymatic activity/functionality.

Although, the active site residues identified in both the *MtMetAPs* are similar, one very interesting feature present in *MtMetAP1c* is that of ~40 residues long unstructured N-terminal extension. To elucidate the role of N-terminal extension on the activity of the enzyme, a series of deletion mutants were generated in which 9, 14, 19, 29 and 39 residues were progressively deleted from the N-terminal end (resulting in constructs:  $\Delta 2-10$ ,  $\Delta 2-15$ ,  $\Delta 2-20$ ,  $\Delta 2-30$  and  $\Delta 2-40$ ). Western blotting with anti-His antibody was performed to confirm the authenticity of the mutant proteins. A drastic decrease in enzyme activity was observed as a result of any deletion. Even on using 10 fold excess of enzyme (total protein) while  $\Delta 2-10$  and  $\Delta 2-15$  depicted only ~30% activity relative to the wild-type other deletion mutants ( $\Delta 2-20$ ,  $\Delta 2-30$  and  $\Delta 2-40$ ) did not show any activity. In fact, SDS-PAGE analysis of these constructs revealed altered purification profiles. Such alteration in the purification profile was presumably because of the involvement of N-terminal extension in wrapping around the protein which prevents exposure of the hydrophobic surface of *MtMetAP1c* to the solvent and thereby has a role in its stability. To know the effect of deletions on secondary and tertiary structures of the protein far-UV and near-UV CD spectroscopic studies were carried out. They revealed that the secondary structure was very much similar to the wild-type though, the near-UV CD spectra of the  $\Delta 2-15$  showed some alteration which is possible since deleting these residues can alter the local tryptophan environment. Further, to understand the possible reason for the sudden alteration in purification profile of  $\Delta 2-20$  mutant sequence analysis was carried out. The protein sequences from different Gram positive bacteria spanning the region of N-terminal extension were analyzed to evaluate presence of any typical feature in them. Interestingly, it was noticed that Val-18 and Pro-19 present in *MtMetAP1c* are conserved amongst all the sequences compared. Therefore, single and double mutants of *MtMetAP1c* were generated for these residues to decipher their role in enzyme activity.

The examination of the enzyme activity of different mutants of these two conserved residues revealed that while V<sub>18</sub>A and P19A were partially active, the double mutant V18AP19A did not show any activity. On the other hand, P19G

exhibited partial activity but neither V18G nor V18GP19G were active. These results therefore, highlight the contribution of Val-18 and Pro-19 present in the N-terminus of MtMetAP1c towards its enzymatic activity. Further, to assess if these mutations resulted in any structural alterations, far-UV CD studies were carried out. In the CD spectra of MtMetAP1c, the 205-208 nm negative band is significantly more intense than the 220-225 nm band. In the crystal structure of MtMetAP1c, it has been reported that there is some PP-II structure in the N-terminal extension region surrounding residues 14-17, which have the sequence (N)-PTRP-(C). It appears from the CD spectrum of the wild-type form of MtMetAP1c protein that the PP-II content could be even higher, involving a longer region within the N-terminal extension, in solution. PP-II structures are characterized by a much higher negative mean residue ellipticity signal in the region of 200-210 nm than in the region of 220-225 nm, with the result that any protein with PP-II structures is likely to have a significantly more intense negative band with a band minimum in the region of 205-208 nm than in the region of 222 nm. In the mutants (V18A, P19A and V18AP19A) although there was no reduction in the 220-225 nm band, but there was a significant reduction in the 205-208 nm band suggesting that these mutations disturb the PP-II structure of the region. Further, in the P19G mutant, there was a reduction in the intensity of the 220-225 nm band without any significant change in the 205-208 nm band relative to what is seen with the mutants (V18A, P19A and V18AP19A), indicating that the PP-II structure is affected to the same extent but with an additional reduction of the helical content. On the other hand, V18G mutant displayed a significant reduction in PP-II content as well as helical content. Thus, these results suggested the importance of these two residues (Val-18 and Pro-19) exclusively and mutually towards the enzymatic activity of MtMetAP1c presumably because of the structural alterations having occurred due to mutation at Val-18 and Pro-19.

To comprehend how these residues situated far away from the active site can affect the activity of MtMetAP1c, MD simulations were executed with protein structures of wild-type, double (V18GP19G) and single mutants (V18G and P19A). The MD simulations displayed that the double (V18GP19G) and single (V18G) mutants, which were found to be inactive in the biochemical assays exhibited remarkable differences in the movement of the active site residues. In fact, these

findings unequivocally emphasizes that N-terminal extension of *MtMetAP1c* contributes towards the functionality of the enzyme by regulating active site residues through “action-at-a-distance” mechanism and this presumably is a unique function of *MtMetAP1c*.

MetAPs are involved in the co-translational removal of initiator methionine *in vivo*. Thus, MetAP needs to interact with the ribosome to perform its function efficiently. *MtMetAP1c*'s N-terminal extension especially the PTRP motif has been proposed to facilitate this interaction in Gram positive bacteria. It has been suggested that the PTRP motif present in it can interact with the SH3 domain containing ribosomal protein L24. Therefore, all the four residues of the PTRP motif present in the extension were mutated to Ala to evaluate if the mutant can interact with ribosomal protein L24. For this purpose *in vitro* pull down assay was performed with purified GST-*MtMetAP1c* and His-*MtL24*. The results of the assay indicated that both GST-*MtMetAP1c* and GST-P14A/T15A/R16A/P17A (multiple mutant, all residues mutated to alanine) were able to interact with His-*MtL24*. In addition to the *in vitro* pull down of the two purified proteins, interaction was confirmed by performing expression pull down where co-purification of His-*MtL24* with either GST-*MtMetAP1c* or GST-P14A/T15A/R16A/P17A was evident. Thus, these two lines of evidence argue that PTRP motif does not play any significant role in direct protein-protein interaction of *MtMetAP1c* and *MtL24*. Interestingly, MetAPs found in Gram negative bacteria do not have any such motif. Therefore, it is conceivable that interaction of MetAP and ribosome through PTRP motif may not be a universal mechanism as has been proposed earlier by others.