Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Molecular Graphics and Modelling 23 (2005) 355–366
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Comparative protein modeling of methionine *S*-adenosyltransferase (MAT) enzyme from *Mycobacterium tuberculosis*: a potential target for antituberculosis drug discovery

Santosh A. Khedkar, Alpeshkumar K. Malde, Evans C. Coutinho*

Department of Pharmaceutical Chemistry, Bombay College of Pharmacy, Kalina, Santacruz (E), Mumbai 400098, India

Received 28 July 2004; received in revised form 27 October 2004; accepted 3 November 2004
Available online 8 December 2004

Abstract

Mycobacterium tuberculosis (Mtb) is a successful pathogen that overcomes the numerous challenges presented by the immune system of the host. In the last 40 years few anti-TB drugs have been developed, while the drug-resistance problem is increasing; there is thus a pressing need to develop new anti-TB drugs active against both the acute and chronic growth phases of the mycobacterium. Methionine *S*-adenosyltransferase (MAT) is an enzyme involved in the synthesis of *S*-adenosylmethionine (SAM), a methyl donor essential for mycolipid biosynthesis. As an anti-TB drug target, Mtb–MAT has been well validated. A homology model of MAT has been constructed using the X-ray structures of *E. coli* MAT (PDB code: 1MXA) and rat MAT (PDB code: 1QM4) as templates, by comparative protein modeling principles. The resulting model has the correct stereochemistry as gauged from the Ramachandran plot and good three-dimensional (3D) structure compatibility as assessed by the *Profiles-3D* score. The structurally and functionally important residues (active site) of Mtb–MAT have been identified using the *E. coli* and rat MAT crystal structures and the reported point mutation data. The homology model conserves the topological and active site features of the MAT family of proteins. The differences in the molecular electrostatic potentials (MEP) of Mtb and human MAT provide evidences that selective and specific Mtb–MAT inhibitors can be designed using the homology model, by the structure-based drug design approaches.

Keywords: Methionine *S*-adenosyltransferase (MAT); *Mycobacterium tuberculosis*; Comparative (homology) modeling; Molecular electrostatic potentials (MEP)

1. Introduction

Tuberculosis (TB) remains one of the world's greatest causes of mortality and morbidity, with approximately 8 million new infections and 2 million deaths per year [1]. More adults die due to TB every year than AIDS and malaria together [2]. *Mycobacterium tuberculosis* (Mtb), the main causative organism of TB, is a successful pathogen that overcomes the numerous challenges presented by the immune system of the host. Mtb is difficult to kill for a number of reasons such as its slow growth, dormancy, complex cell envelope, intracellular pathogenesis and genetic homogeneity [3]. So far, 50 million people have been infected with drug-resistant Mtb strains, and very few drugs have been developed in the past 40 years [4,5]. The situation regarding control of tuberculosis has significantly worsened over the last decade with the spread of strains resistant to multiple antimycobacterial agents [6]. There is thus a profound need for the identification and development of novel chemotherapeutic compounds active against tuberculosis. These compounds should have an ability to penetrate the mycobacterial cell wall, and be active against both the acute and the chronic growth phases.

The complete sequencing of the genome of the best-characterized strain of *M. tuberculosis* H37Rv [7,8] has provided a better understanding of the biology of this slow growing pathogen and also unraveled potential targets that may be of utility in prophylactic and therapeutic interventions. One such target is methionine *S*-adenosyltransferase

* Corresponding author. Tel.: +91 22 26670871; fax: +91 22 26670816.
E-mail address: evans-im@eth.net (E.C. Coutinho).

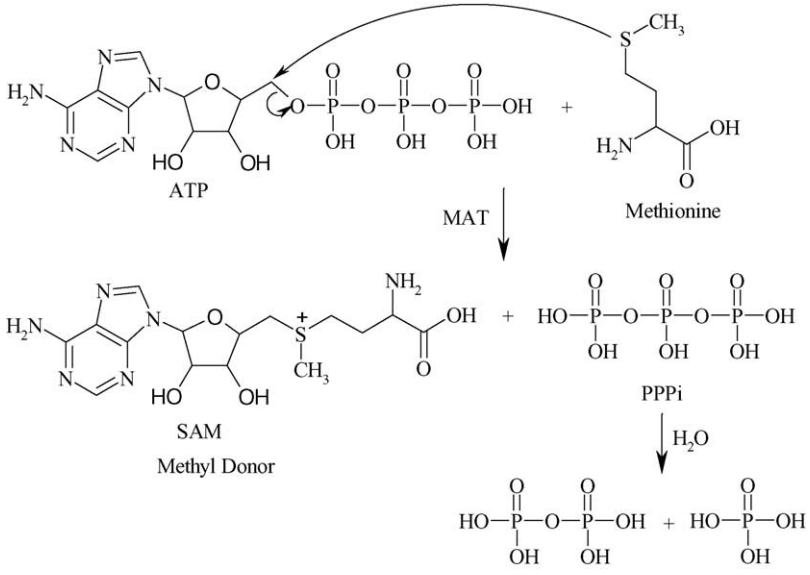
1093-3263/$ – see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2004.11.003

(MAT) [9], which belongs to a family of cytosolic enzymes that synthesize *S*-adenosylmethionine (SAM) using L-methionine and ATP as substrates. The mechanism of MAT catalyzed synthesis of SAM [10,11] involves an S_N2 attack of the sulfur atom of L-methionine on the C5′ atom of ATP, with simultaneous cleavage of the PPP_i chain (Scheme 1). In the second step, PPP_i is hydrolyzed yielding pyrophosphate (PP_i) and orthophosphate (P_i) as side products. SAM is the most important methyl donor in *M. tuberculosis* [12], and the number of reactions that involve SAM has been calculated to be as large as those using ATP [13]. Decarboxylated SAM acts as an aminopropyl donor in the synthesis of polyamines such as spermidine and spermine [13–15]. It has been hypothesized that the positively charged spermidine and spermine stabilize DNA during unwinding and separation of the DNA strands [16], and in mycobacteria may also play a role in transcriptional regulation [17]. Because of its crucial role in polyamine biosynthesis, which occurs during active growth, and in the methylation and cyclopropylation of mycolipids that are necessary for survival during the chronic (dormant) phase, Mtb–MAT has been categorized as an essential enzyme, and is a good target for development of new antitubercular therapy. Recently, Berger and Knodel have characterized the MAT enzyme from *M. tuberculosis* (Mtb) and *M. smegmatis* (Ms), and simultaneously validated MAT as an anti-TB target by inhibiting its activity with some nonselective methionine and purine analogs [9]. Maximum inhibition of Mtb–MAT was found with the purine analogs 8-azaguanine (81.7%) and azathioprine (75.5%) at 10 mM inhibitor concentration; the latter compound, azathioprine was found to inhibit *M. smegmatis* growth in vitro with an MIC of 500 μM. Among the methionine series, the best inhibitors are the epithio and epoxy analogs of L-methionine. There is thus a great potential for the design of more potent and selective inhibitors of mycobacterial MAT. We are interested in constructing a 3D-structural model of this enzyme, which will be useful in the design of more potent and selective inhibitors of Mtb–MAT.

In this paper, we report a homology model of Mtb–MAT using the crystal structures of MAT from *E. coli* (cMAT) [18,19] and rat (rMAT) [20–22]. Using the three-dimensional (3D) model, we have identified the residues important for MAT activity, i.e. the substrate and inhibitor-binding site, which can be exploited in the design of more potent inhibitors. In addition, as a first step to guide selectivity studies, we have calculated and compared molecular electrostatic potential (MEP) surface of Mtb–MAT with its counterpart in human.

2. Methodology

All computations and molecular modeling of Mtb–MAT were carried out on a Silicon Graphics O2 workstation (R4400 MIPS processor) using the *INSIGHT II* molecular modeling package (Accelrys Inc., USA) [23]. The *HOMOLOGY* program (Accelrys Inc., USA) was used for the Comparative protein modeling. The amino acid sequence of Mtb–MAT was obtained from the protein database from NCBI [24] (GenBank accession no. P77899; gi: 3915763). The PSI-BLAST [25] algorithm was used to identify homologous structures for Mtb–MAT by searching the structural database of protein sequences in the Protein Data Bank (PDB) [26]. The MAT crystal structures of *E. coli* (PDB code 1MXA) [18,19] and rat (PDB code 1QM4) [20–22] were selected as template structures for homology modeling of the Mtb–MAT enzyme.



Scheme 1. Mechanism of SAM synthesis involving catalysis by MAT.

2.1. Sequence alignment

The sequences of cMAT and rMAT, extracted from their crystal structures 1MXA and 1QM4, respectively, were aligned using pairwise sequence alignment and further modified manually to optimize the alignment. The sequence identity between cMAT and rMAT is ∼58%, whereas Mtb–MAT shares ∼63% sequence identity with cMAT and ∼54% with rMAT. The structurally conserved regions (SCRs) between reference proteins were then identified. The SCRs were superimposed and the Mtb–MAT sequence was then aligned to the SCRs using the *alignment* module, and finally

```
Ecol  01:    AKHLFTSESVSEGHPDKIADQISDAVLDAILEQDPKARVACETYVKT
Rat   17:    GAFMFTSESVGEGHPDKICDQISDAVLDAHLKQDPNAKVACETVCKT
Hliv  13: LSEGVFMFTSESVGEGHPDKICDQISDAVLDAHLKQDPNAKVACETVCKT
Mtb   01: MSEKGRLFTSESVTEGHPDKICDAISDSVLDALLAADPRSRVAVETLVTT
Msm   01: MS-KGRLFTSESVTEGHPDKICDAISDSVLDALLEQDPKSRVAVETLVTT
           *     ****** ******* * *** **** *  **   ** **   *
Ecol  48: GMVLVGGEITTSAWVDIEEITRNTVREIGYVHS---DMGFDANSCAVLSA
Rat   64: GMVLLCGEITSMAMIDYQRVVRDTIKHIGYDDS---AKGFDFKTCNVLVA
Hliv  63: GMVLLCGEITSMAMVDYQRVVRDTIKHIGYDDS---AKGFDFKTCNVLVA
Mtb   51: GQVHVVGEVTTSAKEAFADITNTVRARILEIGYDSSDKGFDGATCGVNIG
Msm   50: GQVHVAGEVTTTAY---ADIPKIVRDRILDIGYDSSTKGFDGASCGVNVA
          *     ** *                  *         ****   * *
Ecol 095: IGKQSPDI-------|------RADPLE-QGAGDQGLMFGYATNETDVLM
Rat  111: LEQQSP-----|-ED--------------VGAGDQGLMFGYATDETEECM
Hliv 110: LEQQSPDIAQCVHLDRNEED---------VGAGDQGLMFGYATDETEECM
Mtb  101: IGAQSPDIAQGVDTAHEARVEGAADPLDSQGAGDQGLMFGYAINATPELM
Msm  097: IGAQSPDIAQGVDTAHETRVEGKADPLDLQGAGDQGLMFGYAIGDTPELM
          *  ******              ****   ************   *   *
Ecol 135: PAPITYAHRLVQRQAEVRKNGTLPWLRPDAKSQVTFQY--DDGKIV--GI
Rat  152: PLTIVLAHKLNTRMADLRRSGVLPWLRPDSKTQVTVQYVQDNGAVIPVRV
Hliv 151: PLTIILAHKLNARMADLRRSGLLPWLRPDSKTQVTVQYMQDNGAVIPVRI
Mtb  151: PLPIALAHRLSRRLTEVRKNGVLPYLRPDGKTQVTIAY--EDNVPV--RL
Msm  147: PLPIALAHRLARRLTEVRKNGVLDYLRPDGKTQVTIQY--DGTTPV--RL
         *  *  ** *  *    *  ***  **** * ***  *
Ecol 181: DAVVLSTQHSEEIDQKSLQEAVMEEIIKPILPA-----EWLTSATKFFIN
Rat  202: HTIVISVQHNEDITLEAMREALKEQVIKAVVPA-----KYLDEDTIYHLQ
Hliv 201: HTIVISVQHNEDITLEEMRRALKEQVIRAVVPA-----KYLDEDTVYHLQ
Mtb  197: DTVVISTQHAADIDLEKTLDPDIREKVLNTVLDDLAHETLDASTVRVLVN
Msm  193: DTVVLSTQHADGIDLEGTLTPDIREKVVNTVLADLGHETLDTSDYRLLVN
             * * **   *
Ecol 226: PTGRFVIGGPMGDCGLTGRKIIVDTYGGMARHGGGAFSGKDPSKVDRSAA
Rat  247: PSGRFVIGGPQGDAGVTGRKIIVDTYGGWGAHGGGAFSGKDYTKVDRSAA
Hliv 246: PSGRFVIGGPQGDAGVTGRKIIVDTYGGWGAHGGGAFSGKDYTKVDRSAA
Mtb  247: PTGKFVLGGPMGDAGLTGRKIIVDTYGGWARHGGGAFSGKDPSKVDRSAA
Msm  243: PTGKFVLGGPMGDAGLTGRKIIVDTYGGWARHGGGAFSGKDPSKVDRSAA
          * * ** *** ** * ************   **********  *******
Ecol 276: YAARYVAKNIVAAGLADRCEIQVSYAIGVAEPTSIMVETFGTEKVPSEQL
Rat  297: YAARWVAKSLVKAGLCRRVLVQVSYAIGVAEPLSISIFTYGTSKKTEREL
Hliv 296: YAARWVAKSLVKAGLCRRVLVQVSYAIGVAEPLSISIFTYGTSQKTEREL
Mtb  297: YAMRWVAKNVVAAGLAERVEVQVAYAIGKAAPVGLFVETFGTETEDPVKI
Msm  293: YAMRWVAKNVVAAGLAERVEVQVAYAIGKAAPVGLFVETFGSETVDPAKI
          ** * ***  * ***  *   ** **** * *      * *
Ecol 326: TLLVREFFDLRPYGLIQMLDLLHPIYKETAAYGHFGREHF--PWEKTDKA
Rat  347: LEVVNKNFDLRPGVIVRDLDLKKPIYQKTACYGHFGRSEF--PWEVPKKL
Hliv 346: LDVVH-NFDLRPGVIVRDLDLKKPIYQKTACYGHFGRSEF--PWEVPRKL
Mtb  347: EKAIGEVFDLRPGAIIRDLNLLRPIYAPTAAYGHFGRTDVELPWEQLDKV
Msm  343: EKAIGEVFDLRPAAIVRDLDLLRPIYAPTAAYGHFGRTDIELPWEQTNKV
                 *****      * *  ***  ** ******     ***   *
```

Fig. 1. Sequence alignment of MAT enzymes from *E. coli* (Ecol), *R. norvegicus* (Rat), human liver (Hliv), *M. tuberculosis* (Mtb), and *M. smegmatis* (Msm) based on secondary structure and sequence homology. The identical residues in all aligned sequences are indicated with an asterisk (*), whereas, active site residues colored magenta indicate methionine binding site, blue ATP binding site, and red phosphate binding residues. The active site ‘‘mobile loop’’ residues are colored green (Mtb: 107–129, Msm: 103–125, Hliv: 116–129) and the corresponding residues of Ecol (101–108) and Rat (117–128) which are missing in the crystal structures are indicated by ‘‘|’’.

fine-tuned manually, which is shown in Fig. 1. Any one of the two segments for each SCR could be chosen to assign coordinates to Mtb–MAT model. The choice was based on the mutation matrix score of sequence similarity between the target sequence and one of the template SCRs. The SCR segment of either 1MXA or 1QM4, with higher similarity score, was chosen for coordinate assignment.

2.2. Loop and end region modeling

The three regions of Mtb–MAT, i.e. the variable regions: D^{84} to D^{87}, D^{107} to S^{129}, and D^{230} to T^{235}, have very little homology to the template proteins. These were modeled with the de novo *loop generate* algorithm [27], since no similar sequences were recovered from the PDB database with the *loop search* algorithm [28]. The N-terminal (M^1 to K^4) and C-terminal (D^{397} to I^{403}) residues were placed in an extended conformation using the *end repair* command. These regions were refined using a simulated annealing (SA) protocol, after the 'side chain rotamer search' and 'splice-repair' procedures as described in the following sections.

2.3. Side chain rotamer search and splice repair

The possible range of stable conformations for a certain residue's side chain or those of a set of interacting residues were explored with the *rotamer search* module. The *auto rotamer* method was used to monitor any possible steric clashes, which were then relieved appropriately. With the *splice repair* command, the junction or splice points of segments taken from different templates were smoothened out. This entailed an energy minimization protocol using the consistent valence forcefield (CVFF) [29] as implemented in the *Discover* program (v 98 Accelrys Inc., USA). The splice optimization was carried out using steepest descents and conjugate gradients methods to a convergence criterion of 0.001 kcal/mol Å as the maximum derivative.

2.4. Simulated annealing and restrained minimizations

The structures of the N-terminal (M^1 to K^4), the ''active mobile loop'' (D^{107} to S^{129}), and the C-terminal (D^{397} to I^{403}) were refined by an initial minimization step (steepest descents 20,000 steps; conjugate gradients 10,000 steps), followed by simulated annealing (SA) wherein all degrees of freedom for these three regions were allowed to relax, but the heavy atoms of all other residues were held rigid. The protocol used for SA involved a slow heating to 600 K in steps of 100 K, followed by a slow cooling to 300 K, for a period of 25 ps at every temperature. The lowest energy structure from the 300 K trajectory was then subjected to a final round of minimization, with all heavy atoms tethered by a force constant of 100 kcal/mol $Å^2$. The minimization was carried out with 10,000

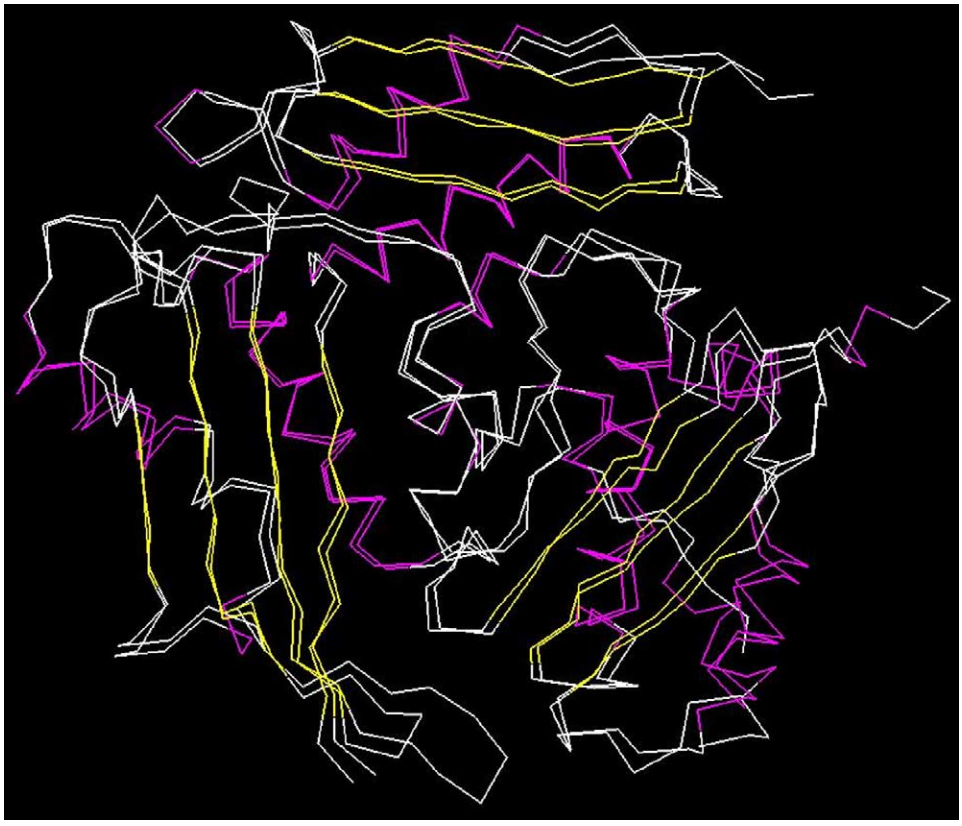


Fig. 2. Superimposition of SCRs of the reference proteins 1MXA and 1QM4 showing only the trace (Cα) along with the secondary structure.

steps each of steepest descents and conjugate gradients. The arrangement of the homodimer was then generated from this optimized structure.

2.5. Validation of the Mtb–MAT 3D-model

The homodimer of Mtb–MAT was generated with the *Pdbset* module of *CCP4* (Windows Version 5.0) [30] using the *Biomt* transformations. The hydrogens were added corresponding to pH 7.4, which resulted in a +1 charge for arginines, and lysines, while the aspartates and glutamates were assigned a formal charge of −1. The dimer structure was further optimized to relieve bad contacts between the side chains of the residues at the dimer interface. The bond lengths, bond angles, torsions and chirality of the Cα atoms in the homodimer were analyzed with the *ProStat* module. The accuracy and validity of the model was tested with *Profiles-3D* (Accelrys Inc., USA) [31], which calculates a 3D to 1D compatibility score, and graphically portrays the properly folded and misfolded region(s) in the protein structure (Fig. 5) by performing an Eisenberg analysis [32,33] of the model.

The homology model of human MAT (hMAT) [34] was generated using the same protocol as described for Mtb–MAT using only the rat MAT (1QM4) crystal structure as the reference protein.

3. Molecular electrostatic potential (MEP)

All MEP calculations and visualization were carried out using the *MOLCAD* program implemented in the *SYBYL* molecular modeling package (Tripos Inc., USA) [35]. The hydrogens to the side chains of both Mtb–MAT and hMAT homology models were added using the *Insight II Biopolymer* module at pH 7.4, resulting in a +1 charge for arginines and lysines, and −1 charge for the aspartates and glutamates. The homology models of Mtb–MAT and hMAT were superimposed on to the rat MAT crystal structure (1QM4) and the corresponding residues within a 7 Å radius of the inhibitor (L-*cis*-AMB) in the 1QM4 crystal structure (residues shown magenta in Fig. 1) were identified. The Gasteiger–Hückel charges were assigned to the atoms of both structures (Mtb and hMAT models). The electron density isosurfaces for both proteins were calculated with a cut value of 0.003 and step width of 0.5 Å, over which the MEP surfaces were generated and visualized.

4. Results and discussion

The aspects that affect the quality of the homology model such as overall sequence identity, extent of binding

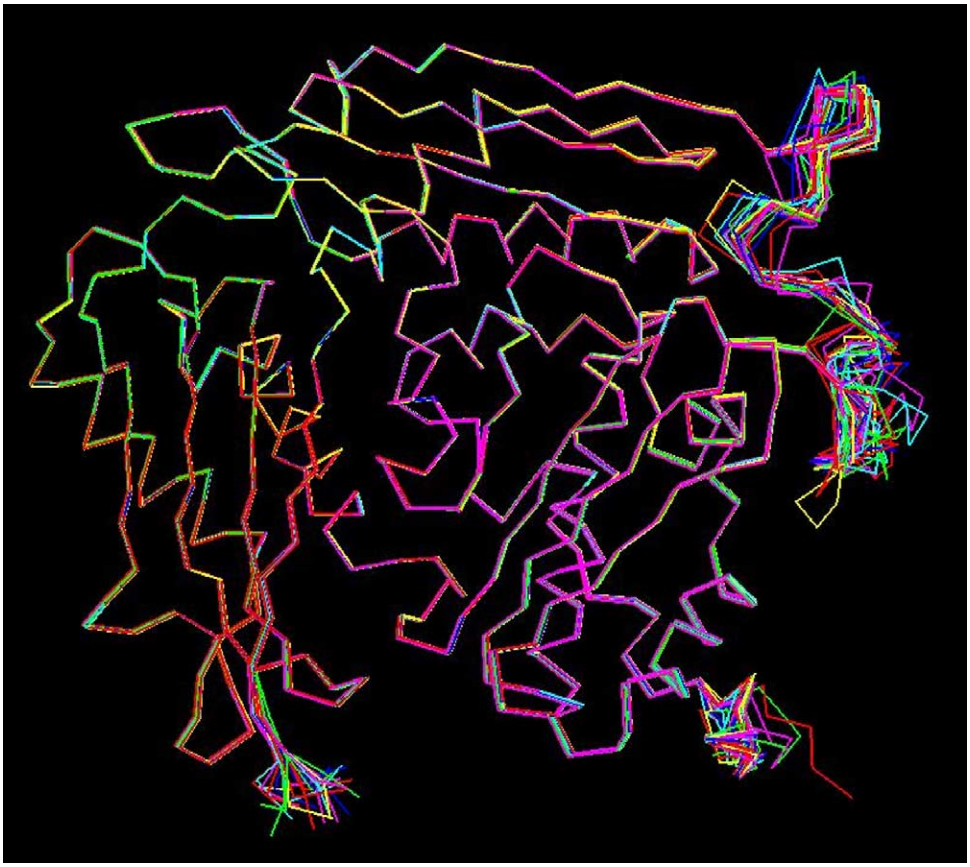


Fig. 3. A display of the 25 structures from the simulated annealing trajectory of the ‘‘active site mobile loop’’.

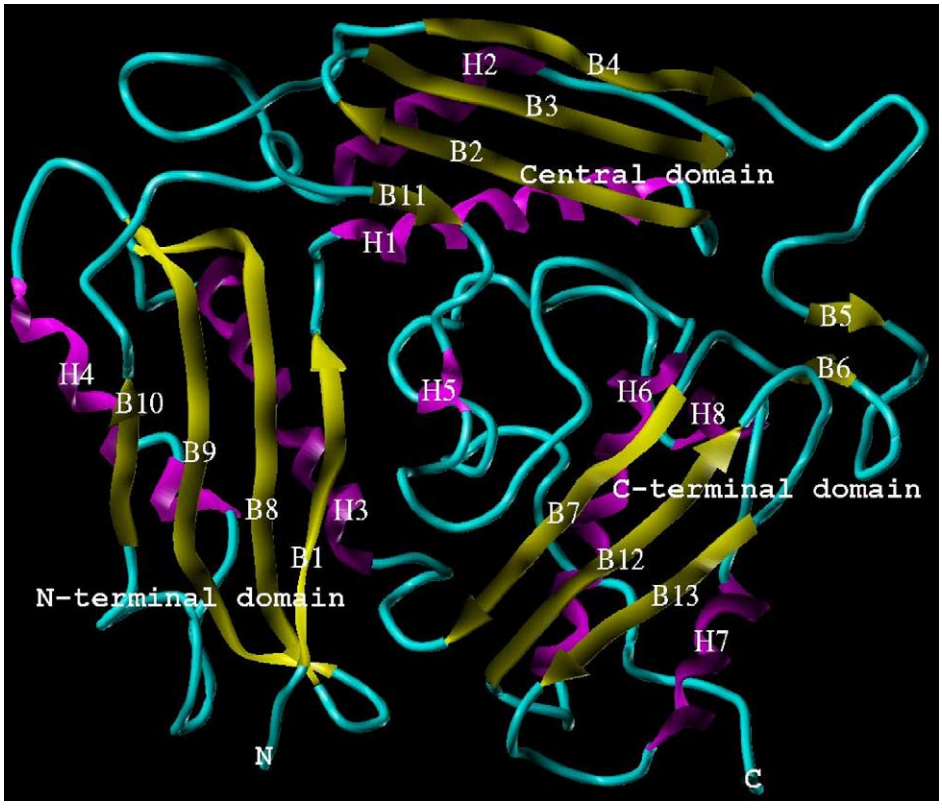


Fig. 4. Structural topology of Mtb–MAT as obtained by comparative modeling. Helices (H) are colored magenta and β-strands (B) are colored yellow. N and C indicate the N-terminal and C-terminal regions of the protein. H's and B's are numbered along the amino acid sequence from the N- to C-terminal.

site conservation, size and location of insertions/deletions, and resolution [36] were actively considered while choosing the template structures for the comparative modeling of Mtb–MAT. A PSI-BLAST [25] search of the PDB database revealed 12 crystal structures of the MAT enzyme either native or co-crystallized with a substrate or inhibitor. Of these, seven structures were from *E. coli* (PDB codes: 1MXA, 1MXB, 1MXC, 1XRA, 1XRB, 1XRC, and 1FUG) [18,19] and five structures from the common rat (*Rattus norvegicus*; PDB codes: 1QM4, 1O90, 1O92, 1O93, 1O9T) [20–22]. Two crystal structures (1MXA and 1QM4) with the highest resolution (2.80 and 2.66 Å, respectively) were chosen as templates for the alignment. Of the two templates, the one chosen for co-ordinate assignment was based on percentage identity to the target protein sequence, as judged by the similarity scores. Since two proteins were being used as templates and coordinates from either of them were to be used, it was necessary to first align the SCRs in the same coordinate frame by superimposition on to each other (Fig. 2). This helps to smoothen out the splice junctions [37]. After alignment, the coordinates of the side chains and the

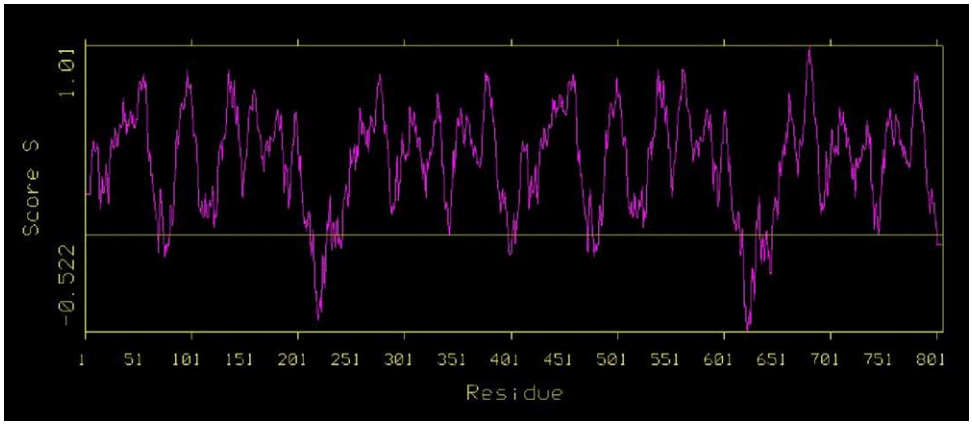


Fig. 5. *Profiles-3D* plot of the homology model of Mtb–MAT dimer.

backbone atoms were copied to the target sequence only if, identical amino acids were found at corresponding positions in the sequence; for 'similar' but not 'identical' amino acids, only the common side chain atoms were copied, while for the rest, amino acid conformations from the library were used. After assigning the coordinates from the SCRs to the target sequence, loops were modeled by searching for similar sequences in the PDB database, and in the case when no sequences were found, coordinates were assigned using the *loop generate* algorithm. The N-terminal and C-terminal regions were placed in an extended conformation. The conformations of every side chain for amino acids in both the SCRs and loop regions were explored, and those with minimum steric clashes retained. A stringent minimization procedure using steepest descents followed by conjugate gradients, with a gradient convergence of 0.001 kcal/mol Å was used to optimize the splice junctions. For modeling less conserved regions, such as the ''active mobile loop'' (amino acids 107–129; which are missing in crystal structures), and the N- and the C-terminal regions, simulated annealing (SA) was used to find out the best possible conformation. It was done by a slow heating of the system to 600 K, followed by a slow cooling to 300 K, to generate the possible lowest energy conformation. The complete structure was finally refined by minimization to remove any steric clashes of the side chains with each other and/or with backbone atoms.

The ''active site mobile loop'', as referred to in the literature [18–22,38], controls the entry of the substrate/inhibitor into the active site and has not been detected in the electron density maps of the X-ray structures. The molecular

Fig. 6. The structure of the homodimer of Mtb–MAT. Residues involved in the methionine (or inhibitor) binding are shown in stick form.

dynamics simulations of this loop shows that it can adopt a wide range of low energy conformations as seen in Fig. 3, which accounts for its flexibility and hence absence in the X-ray diffraction electron density maps [20–22]. The ''active site mobile loop'' after SA is found in a discontinuous β-sheet structure (labeled B5 and B6) in C-terminal domain (Fig. 4).

The ϕ and ψ dihedral angles in the homodimer lie within the allowed regions of the Ramachandran plot. The overall 3D-structure compatibility (*Profiles-3D*) score for the dimer was 287.21, which was far higher than the threshold score of 166.35 for a protein of this size, expressing a strong confidence in the 3D model (Fig. 5).

5. Topology of the Mtb–MAT enzyme

The active site of MAT resides at the interface of two monomers (Fig. 6), the residues of one monomer constituting the methionine binding site, while residues from the second monomer are involved in ATP binding. Biologically, MAT of *E. coli* and Mtb exists as a homodimer [9] and that of rats and humans exists as a dimer (called *MAT III*) as well as a tetramer (*MAT I*) [21]. Further, *MAT I* and *MAT III* remain in equilibrium; the predominant form is dependent on the methionine concentration.

There are three major domains in the MAT protein—the N-terminal domain, the central domain, and the C-terminal domain, so termed because of their location in the tertiary structure. Each domain is made up of two α-helices and three or four β strands. A long α-helix, followed by a pair of anti-parallel β-strands and a short α-helix (e.g. H1–β2–β3–H2 in the central domain, Fig. 4) is a common secondary structural motif in each domain. Correspondingly, the N-terminal and C-terminal domains are formed by H3–β8–β9–H4 and H6–β12–β13–H7 motifs, respectively. It must be noted that coordinates for the H5 helix were assigned from the rMAT structure (1QM4) in which it appears as a helix, but occurs as a loop in the cMAT (1MXA) structure, despite having a 100% sequence identity. The strands β4 and β10 form parallel β-sheets with β3 and β9 in the central and N-terminal domains, respectively. On the other hand, β7 forms an antiparallel sheet with β12 in the C-terminal domain. The

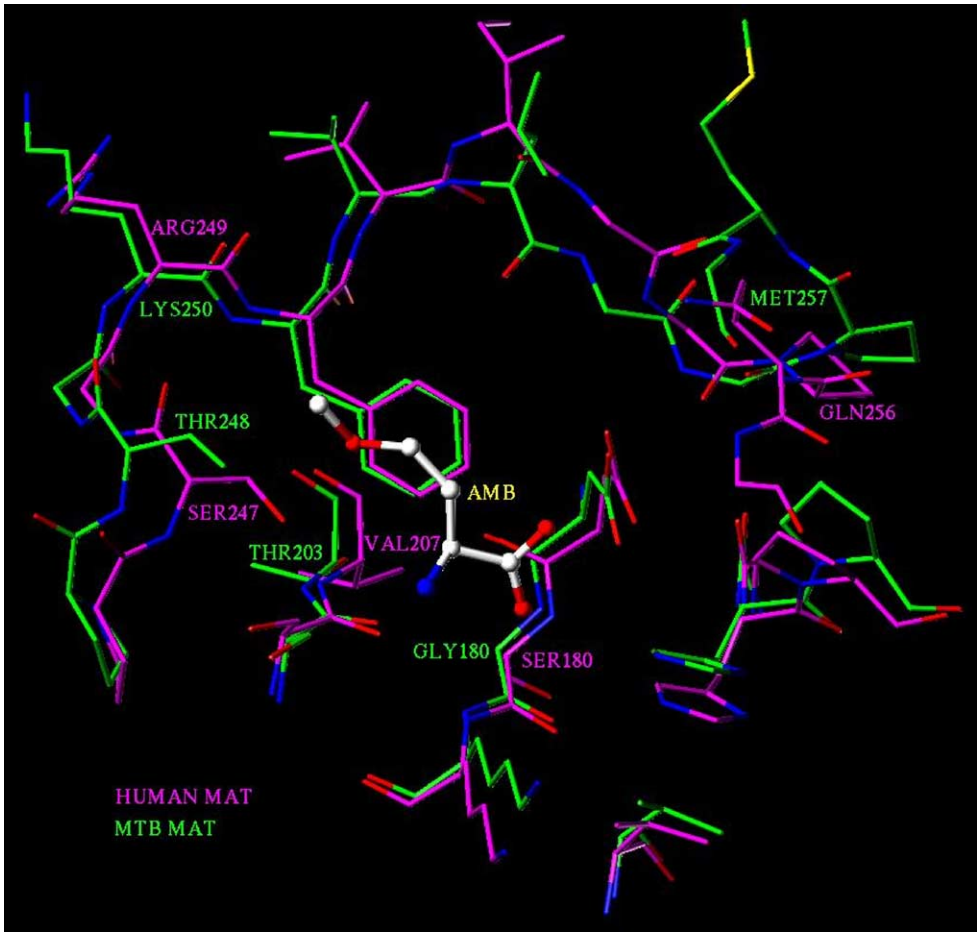


Fig. 7. The active site residues of Mtb–MAT (magenta) and hMAT (green) after the superposition onto the rMAT crystal structure (1QM4). The residues of rMAT have not been shown as it shares 100% sequence identity to human MAT, in the active sites. Only the inhibitor L-2-amino-4-methoxy-*cis*-but-3-enoic acid (L-*cis*-AMB) from the 1QM4 X-ray structure has been shown to indicate its relative position in the Mtb and hMAT active sites.

fourth strand, β11 in the central domain and β1 in the N-terminal domain form antiparallel sheets with β2 and β8 strands, respectively. The ''active mobile loop'' which forms a bridge between the central and the C-terminal domains is found to form a short antiparallel sheet with strands β5 and β6. A close inspection of the Mtb–MAT model reveals that it has all the topological features of this family of proteins.

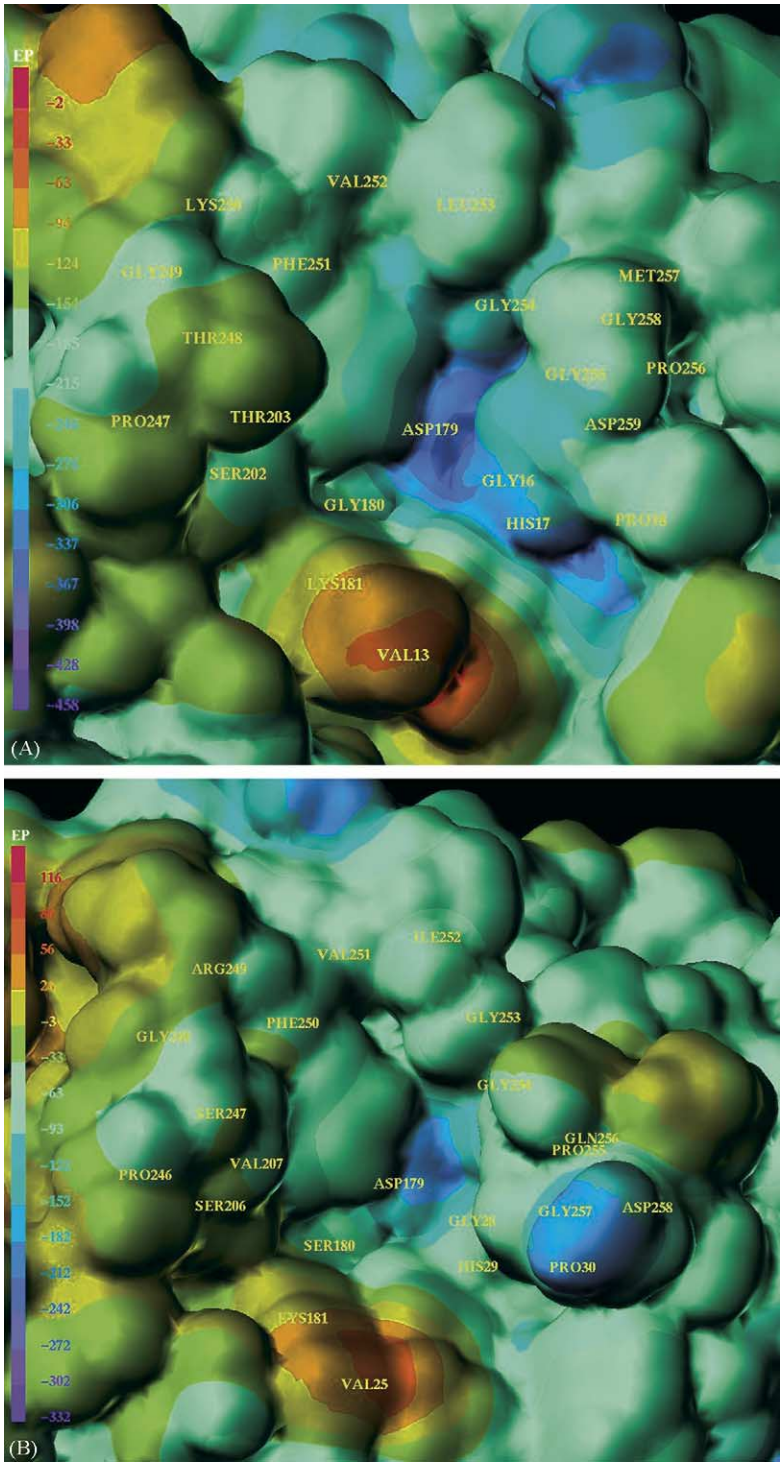


Fig. 8. The absolute values of the molecular electrostatic potential (MEP) displayed for the active sites of Mtb–MAT (A) and hMAT (B). The deep blue color indicates the highest negative potential whereas the most positive (the lowest negative) potential is seen as a deep red color. The color spectrum shown to the left shows the gradation of electrostatic potential in the respective enzymes. In order to make valid comparisons between figures (A) and (B), the electrostatic potentials have been put on the same scale, which are shown in (C) for Mtb–MAT and (D) for hMAT enzymes.

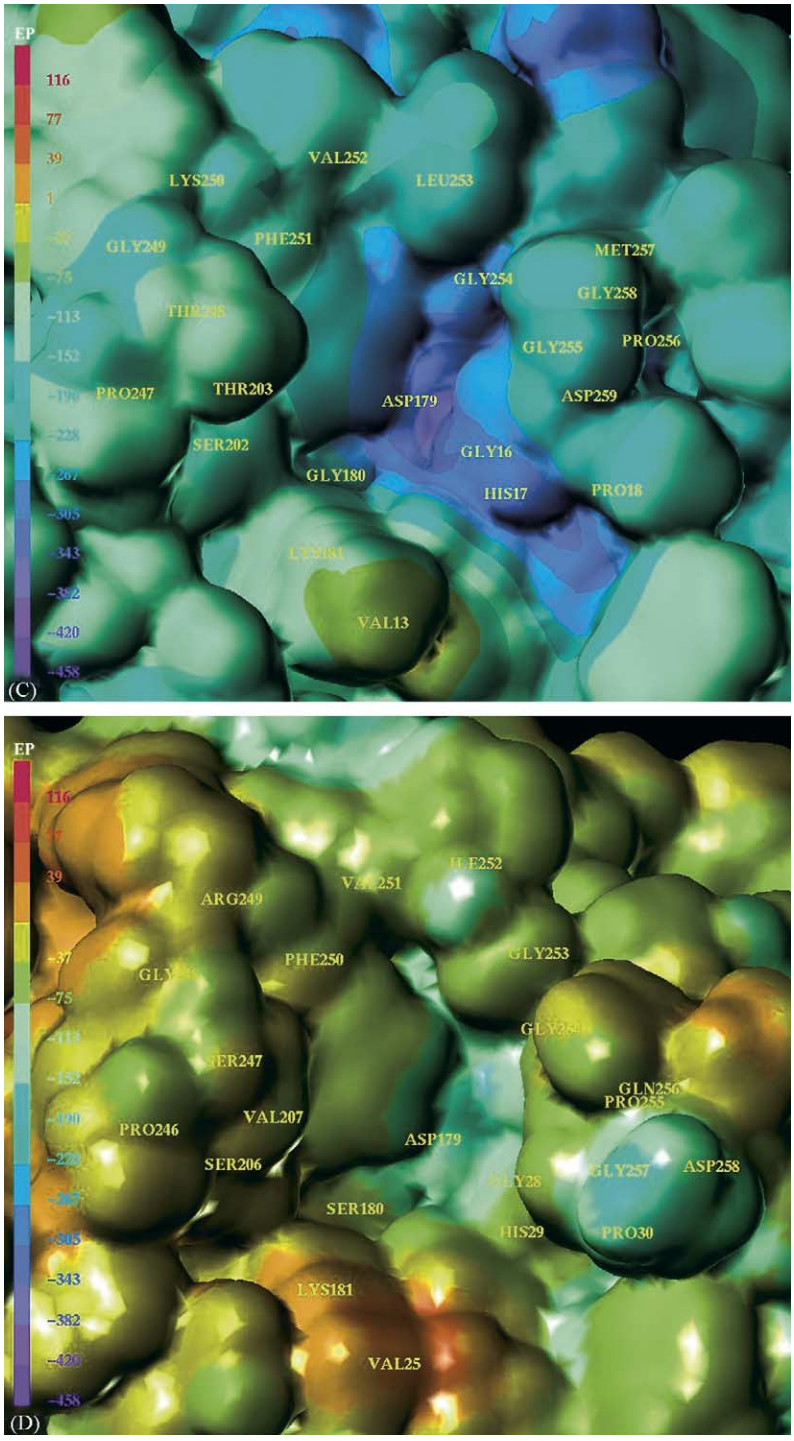


Fig. 8. (*Continued*).

6. Mtb–MAT active site

The residues forming the active site of Mtb–MAT have a good correspondence to other MATs, except for a few mutations (Fig. 1). The Mtb–MAT active site is located in a broad cavity (Fig. 6) composed of the dimer interface with residues from both subunits involved. The ATP-binding region is composed of G^{131}AGDQG^{136} residues, while the segment H^{278}GGGAFSGKD^{287} is the phosphate-binding P-loop. The methionine-binding motif is formed by residues F^{251}VLGGPMGDA^{260} in the loop which connects the β10 strand with β11, as seen in Fig. 4. While the two former regions are well conserved in the three MATs, mutations are seen in the methionine-binding region of Mtb–MAT. A superposition of MAT models of human and Mtb reveals the differences in the active site (Fig. 7). The

residues Gly[180], Thr[203], Thr[248], Lys[250] and Met[257] in Mtb–MAT are replaced by Ser[180], Val[207], Ser[247], Arg[249], and Gln[256] in hMAT, respectively (Fig. 7). These relative differences in the active site will undoubtedly affect the steric and electrostatic interactions with inhibitors, and this can be exploited to design selective MAT inhibitors. It is important to mention that the ATP and methionine binding sites lie in different chains of the dimer, and together form the full active site where the biochemical reaction (Scheme 1) [11] leading to the synthesis of SAM takes place. It is known that potassium and magnesium ions confer rigidity to the active site by forming a salt bridge between residues of the two chains [20]. A magnesium ion coordinates with the carboxylate oxygen atom of the inhibitor (L-*cis*-AMB) and the side chain carboxylate oxygen atom of Asp[179] [21,22]. Site-directed mutagenesis studies carried out for rMAT have led to the conclusion that Phe[251], Asp[180], and Lys[182], are critical for SAM synthesis either by binding methionine or possibly by being involved in subsequent steps of the mechanism. These conclusions were based on the fact that their mutations resulted in an absolute loss of MAT activity when compared to the native enzyme [20].

7. Molecular electrostatic potential (MEP)

The MEP is a popular indicator of electrophilic and nucleophilic centers, which governs the strength of bonds, the strength of non-bonded interactions and molecular reactivity. It affects the strength of the interaction of the ligand with the receptor protein. Bhattacharrjee and Karle have used the MEP to relate the antimalarial potency of carbinolamine analogs [39] and neurotoxicity of artemisinin analogs [40]. In case of a ligand–protein interaction, at the active site, the ligand experiences a unique environment in terms of electrostatic, steric and hydrophobic properties. Variations in these properties in the active sites of different proteins can contribute to selective/specific or tighter binding of the ligand between the different proteins. Thus, a comparison of the MEPs of Mtb–MAT and hMAT [34] is one-step towards understanding selectivity in MAT inhibition.

The MEP surfaces are shown with absolute values of the electrostatic potential for Mtb–MAT (−2 to −458 kcal/mol Å; Fig. 8A) and hMAT (+116 to −332 kcal/mol Å; Fig. 8B). Since the range of potentials in the two proteins differ greatly, meaningful comparisons can only be made when the electrostatic potential of the two proteins are placed on the same scale +116 to −458 kcal/mol Å, as shown in Fig. 8C and D for Mtb–MAT and hMAT respectively. As is evident in Figs. 8C and D, the electrostatic potential covering the active site residues has sharply differing values, with the active site of Mtb–MAT being more deeply negative. Within, hMAT itself, the surface around Asp[179], Gly[257] and Asp[258] residues has a relatively more electronegative potential (Fig. 8D) when compared with other regions in the active site. This same contrasting feature is visible for the region encompassing residues Asp[179], Gly[16] and His[17] in the active site of Mtb–MAT (Fig. 8C). This means that the two proteins will bind inhibitors in the same relative sense [21,22], but the binding affinities of a given inhibitor will not be the same for the two proteins, due to the differences in the absolute values of the electrostatic potential. Further, the three-fold difference in the value of the potential at Asp[179] in Mtb–MAT (−382 kcal/mol Å) and hMAT (−113 kcal/mol Å) will have a deep impact on the selective binding of methionine analogs via coordination of the Mg^{2+} ion [21,22].

8. Conclusions

Methionine *S*-adenosyltransferase is one of the 256 enzymes deemed necessary for life [14,15,41]. The critical role played by this enzyme in both the active and dormant phases of the mycobacterial life cycle makes it an attractive antitubercular target. A model of Mtb–MAT has been constructed from the crystal coordinates of two MAT structures of *E. coli* and the common rat by comparative protein modeling principles. The overall topology of Mtb–MAT is similar to the MAT family of enzymes. The functionally and biochemically important residues comprising the active site of Mtb–MAT, i.e., the substrate binding sites appear to be conserved, but there are some differences in the active sites of Mtb and human MAT, which can be exploited to design specific and selective inhibitors of Mtb–MAT. These differences become apparent when one compares the molecular electrostatic potential of mycobacterial and human MAT. Further studies to explore the selective binding profile of Mtb–MAT inhibitors are under way.

Acknowledgements

This work was made possible by grants from the All India Council of Technical Education (F. No. 8022/RID/NPROJ/RPS-5/2003-04) and the Department of Science and Technology (SR/FST/LSI-083/2003) to E. Coutinho. SAK thanks the Lady Tata Memorial Trust, Mumbai and AKM the University Grants Commission (UGC, New Delhi) for financial support.

References

[1] C. Dye, S. Scheele, P. Dolin, V. Pathania, M.C. Raviglione, Consensus statement. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. WHO Global Surveillance and Monitoring Project, JAMA 282 (1999) 677–686.

[2] E.L. Corbett, C.J. Watt, N. Walker, D. Maher, B.G. Williams, M.C. Raviglione, C. Dye, The growing burden of tuberculosis: global trends and interactions with the HIV epidemic, Arch. Int. Med. 163 (2003) 1009–1021.

[3] G.S. Besra, D. Chatterjee, Lipids and carbohydrates of *Mycobacterium tuberculosis*, in: B.R. Bloom (Ed.), Tuberculosis. Pathogenesis, Protection, and Control, ASM Press, Washington, 1994, pp. 285–306.
[4] World Health Organization (WHO), Tuberculosis Fact Sheet (No. 104), 2000. http://www.who.int/gtb/.
[5] P. Chopra, L.S. Meena, Y. Sing, New drug targets for *Mycobacterium tuberculosis*, Indian J. Med. Res. 117 (2003) 1–9.
[6] R.J. O'Brien, P.P. Nunn, The need for new drugs against tuberculosis. Obstacles, opportunities, and next steps, Am. J. Respir. Crit. Care Med. 163 (2001) 1055–1058.
[7] W.J. Philipp, S. Poulet, K. Eiglmeier, L. Pascopella, V. Balasubramanian, B. Heym, S. Bergh, B.R. Bloom, W.R. Jacobs Jr., S.T. Cole, An integrated map of the genome of the tubercle bacillus, *Mycobacterium tuberculosis* H37Rv, and comparison with *Mycobacterium leprae*, Proc. Natl. Acad. Sci. U.S.A. 93 (1996) 3132–3137.
[8] S.T. Cole, R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S.V. Gordon, K. Eiglmeier, S. Gas, C.E. Barry III, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Delvin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, A. Krofh, J. Mclean, S. Moule, L. Murphy, K. Oliver, J. Osborne, M.A. Quill, M.-A. Rajendream, J. Rogers, S. Rutter, K. Seeger, J. Skelton, R. Squares, S. Squares, J.E. Suelston, K. Taylor, S. Whitehead, B.G. Barrell, Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence, Nature 393 (1998) 537–544.
[9] B.J. Berger, M.H. Knodel, Characterization of methionine adenosyltransferase from *Mycobacterium smegmatis and M. tuberculosis*, B.M.C. Microbiol. 3 (2003) 12–24.
[10] S.H. Mudd, The Enzymes, vol. 8, 3rd ed., Academic Press, New York, 1973, pp. 121–154.
[11] R. Parry, A. Minta, Studies of enzyme stereochemistry. Elucidation of the stereochemistry of *S*-adenosylmethionine formation by yeast methionine adenosyltransferase, J. Am. Chem. Soc. 104 (1982) 871–872.
[12] G.L. Cantoni, Biological methylation: selected aspects, Annu. Rev. Biochem. 44 (1975) 435–451.
[13] J.M. Mato, L. Alvarez, P. Ortiz, M.A. Pajares, *S*-adenosylmethionine synthesis: molecular mechanisms and clinical implications, Pharmacol. Ther. 73 (1997) 265–280.
[14] C.W. Tabor, H. Tabor, Methionine adenosyltransferase (*S*-adenosylmethionine synthetase) and *S*-adenosylmethionine decarboxylase, Adv. Enzymol. 56 (1984) 251–282.
[15] M. Kotb, A.M. Geller, Methionine adenosyltransferase: structure and function, Pharmacol. Ther. 59 (1993) 125–143.
[16] L.J. Marton, A.E. Pegg, Polyamines as targets for therapeutic intervention, Ann. Rev. Pharmacol. Toxicol. 35. (1995) 55–91.
[17] N.K. Sarkar, S. Shankar, A.K. Tyagi, Polyamines exert regulatory control on mycobacterial transcription: a study using RNA polymerase from *Mycobacterium phlei*, Biochem. Mol. Biol. Int. 35 (1995) 1189–1198.
[18] F. Takusagawa, S. Kamitori, G.D. Markham, Structure and function of *S*-adenosylmethionine synthetase: crystal structures of *S*-adenosylmethionine synthetase with ADP, BrADP, and PPi at 2.8 angstroms resolution, Biochemistry 35 (1996) 2586–2596.
[19] F. Takusagawa, S. Kamitori, S. Misaki, G.D. Markham, Crystal structure of *S*-adenosylmethionine synthetase, J. Biol. Chem. 271 (1996) 136–147.
[20] Z. Fu, Y. Hu, G.D. Markham, F. Takusagawa, Flexible loop in the structure of *S*-adenosylmethionine synthetase crystallized in the tetragonal modification, J. Biomol. Struct. Dyn. 13 (1996) 727–739.
[21] B. Gonzales, M.A. Pajares, J.A. Hermoso, L. Alvarez, F. Garrido, J.R. Sufrin, J. Sanz-Aparicio, The crystal structure of tetrameric methionine adenosyltransferase from rat liver reveals the methionine-binding site, J. Mol. Biol. 300 (2000) 363–375.
[22] B. Gonzales, M.A. Pajares, J.A. Hermoso, D. Guillerm, G. Guillerm, J. Sanz-Aparicio, Crystal structures of methionine adenosyltransferase complexed with substrates and products reveal the methionine-ATP recognition and give insights into the catalytic mechanism, J. Mol. Biol. 331 (2003) 407–416.
[23] InsightII, Version 98, Accelrys Inc., San Diego, CA, USA, (1998).
[24] in: http://www.ncbi.nlm.nih.gov/protein.
[25] S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucl. Acids Res. 25 (1997) 3302–3389.
[26] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The protein data bank, Nucl. Acids Res. 28 (2000) 235–242.
[27] P.S. Shenkin, D.L. Yarmush, R.M. Fine, H. Wang, C. Levinthal, Predicting antibody hypervariable loop conformation. I. Ensembles of random conformations for ringlike structures, Biopolymers 26 (1987) 2053–2085.
[28] U. Hobohm, M. Scharf, R. Schneider, C. Sander, Selection of representative protein data sets, Protein Sci. 1 (1992) 409–417.
[29] P. Dauber-Osguthorpe, V.A. Roberts, D.J. Osguthorpe, J. Wolff, M. Genest, A.T. Hagler, Structure and energetics of ligand binding to proteins: *Escherichia coli* dihydrofolate reductase-trimethoprim, a drug-receptor system, Proteins 4 (1988) 31–47.
[30] Collaborative Computational Project, Number 4. The CCP4 Suite: Programs for Protein Crystallography, Acta Cryst. D50 (1994) 760–763.
[31] R. Luthy, J.U. Bowie, D. Eisenberg, Assessment of protein models with three-dimensional profiles, Nature 356 (1992) 83–85.
[32] M. Gribskov, A.D. Mclachlan, D. Eisenberg, Profile analysis: detection of distantly related proteins, Proc. Natl. Acad. Sci. U.S.A. 84 (1987) 4355–4358.
[33] M. Gribskov, R. Luthy, D. Eisenberg, Profile analysis, Meth. Enzymol. 183 (1990) 146–159.
[34] S.A. Khedkar, A.K. Malde, E.C. Coutinho, Topology of human methionine *S*-adenosyltransferase, Ind. J. Chem. Sec. A, communicated.
[35] Sybyl, Version 6.7, Tripos Inc., St. Louis, MO, USA, (2002).
[36] D.J. Diller, R. Li, Kinases, homology models, and high throughput docking, J. Med. Chem. 46 (2003) 4638–4647.
[37] InsightII: Homology, User Guide, Accelrys Inc., USA, (1994) pp. 5–22.
[38] J.C. Taylor, G.D. Markham, Conformational dynamics of the active site loop of *S*-adenosylmethionine synthetase illuminated by site-directed spin labeling, Arch. Biochem. Biophys. 415 (2003) 164–171.
[39] A.K. Bhattacharrjee, J.M. Karle, Functional correlation of molecular electronic properties with potency of synthetic carbinolamines antimalarial agents, Bioorg. Med. Chem. 6 (1998) 1927–1933.
[40] A.K. Bhattacharrjee, J.M. Karle, Stereoelectronic properties of antimalarial artemisinin analogues in relation to neurotoxicity, Chem. Res. Toxicol. 12 (1999) 422–428.
[41] J.M. Mato, L. Alvarz, P. Oritz, M.A. Pajares, *S*-adenosylmethionine synthesis: molecular mechanisms and clinical implications, Pharmacol. Ther. 73 (1997) 265–280.