CODEN IEJMAT *Internet Electronic Journal of Molecular Design* **2007**, *6*, 151–166 ISSN 1538–6414
BioChem Press http://www.biochempress.com

Inter*net* Electronic Journal of Molecular Design

June 2007, Volume 6, Number 6, Pages 151–166

Editor: Ovidiu Ivanciuc

Inhibition of Methionine *S*–adenosyltransferase of *M. smegmatis and M. tuberculosis*: Homology Modeling, Docking and *De Novo* Inhibitor Design

Santosh A. Khedkar,[1] Alpeshkumar K. Malde,[1] and Evans C. Coutinho [1]

[1] Department of Pharmaceutical Chemistry, Bombay College of Pharmacy, Kalina, Santacruz (E), Mumbai 400 098, India

Received: December 26, 2006; Revised: February 8, 2007; Accepted: May 15, 2007; Published: June 30, 2007

Citation of the article:
S. A. Khedkar, A. K. Malde, and E. C. Coutinho, Inhibition of Methionine *S*–adenosyltransferase of *M. smegmatis and M. tuberculosis*: Homology Modeling, Docking and *De Novo* Inhibitor Design, *Internet Electron. J. Mol. Des.* **2007**, *6*, 151–166, http://www.biochempress.com.

Inter*net* Electronic Journal of Molecular Design

BioChem Press
http://www.biochempress.com

Inhibition of Methionine *S*–adenosyltransferase of *M. smegmatis and M. tuberculosis*: Homology Modeling, Docking and *De Novo* Inhibitor Design

Santosh A. Khedkar,[1] Alpeshkumar K. Malde,[1] and Evans C. Coutinho [1,*]

[1] Department of Pharmaceutical Chemistry, Bombay College of Pharmacy, Kalina, Santacruz (E), Mumbai 400 098, India

Received: December 26, 2006; Revised: February 8, 2007; Accepted: May 15, 2007; Published: June 30, 2007

Internet Electron. J. Mol. Des. **2007**, *6* (6), 151–166

Abstract

Mycobacterium tuberculosis (Mtb) is a successful pathogen that overcomes the numerous challenges presented by the immune system of the host. Increasing resistance to the currently available drugs has been a pressing force for the search for new anti–TB drugs which are active against drug–resistant strains. Also, any new inhibitor should be active against both the acute and chronic growth phases of the mycobacteria. Methionine *S*–adenosyltransferase (MAT) is an enzyme involved in the synthesis of *S*–adenosylmethionine, a methyl donor essential for mycolipid and polyamine biosynthesis (during active growth) and for methylation and cyclopropylation of mycolipids (necessary for survival during the dormant phase). Inhibitors of the MAT enzyme appear to be promising anti–TB agents. The homology models of MAT from *M. smegmatis* (non–virulent) and *M. tuberculosis* (pathogenic form) were built using comparative protein modeling principles. Docking studies (with the program GOLD) in conjunction with different scoring functions (*GoldScore*, *ChemScore* and *HINT* score) were employed to explore the binding modes of some known inhibitors of Msm–MAT and to decipher the crucial interactions between the inhibitors and the enzyme. Various terms of the scoring functions were correlated to the inhibition data. The docking and scoring protocol established and validated for Msm–MAT can be used to search small–molecule 3D databases to extract novel ligands for Mtb–MAT (pathogenic form). Further, new molecules were designed for Mtb–MAT using the *de novo* ligand design approach of the program Ludi.

Keywords. Methionine S–adenosyltransferase (MAT); docking; HINT; GoldScore; ChemScore; *de novo* inhibitor design.

1 INTRODUCTION

Mycobacterium tuberculosis is a pathogen that successfully overcomes the numerous challenges presented by the immune system of the host [1]. In 1993, the World Health Organization (WHO) declared tuberculosis (TB) as a "global emergency" on account of its pathogenic synergy with

The work has been presented at the *10th ISCB International Conference on Drug Discovery: Perspectives and Challenges* (*Including Symposium on Infectious Diseases*) jointly organized by the Indian Society of Chemists and Biologists (ISCB), India and the University of Toronto, Canada during 24th–26th February, 2006 at the Central Drug Research Institute (CDRI), Lucknow, India.

* Correspondence author; phone: +91–22–26670871; fax: +91–22–26670816; E–mail: evans@bcpindia.org.

human immunodeficiency virus (HIV) infection. An HIV infection significantly increases the risk that new or latent TB infection will progress to the active stage [2,3]. The problem of increasing drug–resistance has been a pressing force to develop new anti–TB drugs. There is thus a vital need for the identification and development of novel chemotherapeutic compounds active against both the acute and the chronic growth phases of the mycobacterium. Focusing on the existing antitubercular targets for drug development [4] may be of limited value because chances of resistance by mutation in the protein target may render the drugs ineffective. Precisely, because of this observed drug–resistance by the bacterium, it is imperative to develop smart new drugs that inhibit novel targets that are structurally and functionally different from those currently known.

Methionine *S*–adenosyltransferase (MAT) is a cytosolic enzyme that utilizes L–methionine and ATP as substrates to synthesize *S*–adenosylmethionine (SAM), a methyl donor essential for mycolipid biosynthesis [5–9]. It has a crucial role in polyamine biosynthesis [10], which occurs during active growth, and in the methylation [11] and cyclopropylation of mycolipids that are necessary for survival during the chronic (dormant) phase. Because of this, MAT has been categorized as an essential enzyme, and is a druggable target for development of new antitubercular drugs. This has been corroborated by the study of Berger *et al*. [12], where they characterized the MAT enzyme from *M. smegmatis* (Msm) and *M. tuberculosis* (Mtb) and simultaneously validated MAT as an anti–mycobacterial target, by inhibiting its activity with some nonselective methionine and purine analogs. The MAT enzymes from Msm and Mtb are highly homologous (88 % sequence identity with identity matrix, Insight II) and conserve 100% sequence similarity in their active sites.

We report here a study involving docking and ranking of known inhibitors of Msm–MAT in the active site of a homology model of Msm–MAT, to establish a protocol that could be applied in virtual screening of databases to uncover novel inhibitors for Mtb–MAT. Since the sequences of Msm– and Mtb–MAT [13] are highly conserved, we expect the protocol to be successful with Mtb–MAT. Further, we have used the homology model to design some novel molecules with the structure–based drug design program Ludi.

2 COMPUTATIONAL METHODS

All computations and molecular modeling of the MAT enzyme and its inhibitors were carried out on a Pentium IV PC using the programs Insight II (v 2005; Accelrys Inc., USA) and Sybyl (v 7.1, Tripos Inc., USA) working under the Linux Red Hat Enterprise WS (v 2.1) OS, and GOLD (v 3.0.1; Cambridge Crystallographic Data Center, UK) [14–16] running under the Windows environment.

2.1 Protein and Ligand Preparation

The Msm–MAT enzyme was modeled by comparative protein modeling principles following the same protocol as described for Mtb–MAT [13]. Mtb– and Msm–MAT share ~60% sequence identity with MAT of *E. coli*. The sequence alignment that was used to define structurally conserved regions (SCRs) is shown in Figure 1.

```
Ecol  01:    AKHLFTSESVSEGHPDKIADQISDAVLDAILEQDPKARVACETYVKT
Mtb   01: MSEKGRLFTSESVTEGHPDKICDAISDSVLDALLAADPRSRVAVETLVTT
Msm   01: MS-KGRLFTSESVTEGHPDKICDAISDSVLDALLEQDPKSRVAVETLVTT
          **    ******* ******* * *** ******  **   ** **   *
Ecol  48: GMVLVGGEITTSAWVDIEEITRNTVREIGYVHS---DMGFDANSCAVLSA
Mtb   51: GQVHVVGEVTTSAKEAFADITNTVRARILEIGYDSSDKGFDGATCGVNIG
Msm   50: GQVHVAGEVTTTAY---ADIPKIVRDRILDIGYDSSTKGFDGASCGVNVA
          * * * ** *  *      *       *         ****   * *
Ecol 095: IGKQSPDI-------|------RADPLE-QGAGDQGLMFGYATNETDVLM
Mtb  101: IGAQSPDIAQGVDTAHEARVEGAADPLDSQGAGDQGLMFGYAINATPELM
Msm  097: IGAQSPDIAQGVDTAHETRVEGKADPLDLQGAGDQGLMFGYAIGDTPELM
          ** *****              ****  *************   *  **
Ecol 135: PAPITYAHRLVQRQAEVRKNGTLPWLRPDAKSQVTFQY--DDGKIV--GI
Mtb  151: PLPIALAHRLSRRLTEVRKNGVLPYLRPDGKTQVTIAY--EDNVPV--RL
Msm  147: PLPIALAHRLARRLTEVRKNGVLDYLRPDGKTQVTIQY--DGTTPV--RL
        * **  ****  *  ********  **** * ***  *       *
Ecol 181: DAVVLSTQHSEEIDQKSLQEAVMEEIIKPILPA-----EWLTSATKFFIN
Mtb  197: DTVVISTQHAADIDLEKTLDPDIREKVLNTVLDDLAHETLDASTVRVLVN
Msm  193: DTVVLSTQHADGIDLEGTLTPDIREKVVNTVLADLGHETLDTSDYRLLVN
        * ** ****   **         *                 *      *
Ecol 226: PTGRFVIGGPMGDCGLTGRKIIVDTYGGMARHGGGAFSGKDPSKVDRSAA
Mtb  247: PTGKFVLGGPMGDAGLTGRKIIVDTYGGWARHGGGAFSGKDPSKVDRSAA
Msm  243: PTGKFVLGGPMGDAGLTGRKIIVDTYGGWARHGGGAFSGKDPSKVDRSAA
          *** ********* ************** ********************
Ecol 276: YAARYVAKNIVAAGLADRCEIQVSYAIGVAEPTSIMVETFGTEKVPSEQL
Mtb  297: YAMRWVAKNVVAAGLAERVEVQVAYAIGKAAPVGLFVETFGTETEDPVKI
Msm  293: YAMRWVAKNVVAAGLAERVEVQVAYAIGKAAPVGLFVETFGSETVDPAKI
          ** * **** ****** * * ** **** * *    ***** *      *
Ecol 326: TLLVREFFDLRPYGLIQMLDLLHPIYKETAAYGHFGREHF--PWEKTDKA
Mtb  347: EKAIGEVFDLRPGAIIRDLNLLRPIYAPTAAYGHFGRTDVELPWEQLDKV
Msm  343: EKAIGEVFDLRPAAIVRDLDLLRPIYAPTAAYGHFGRTDIELPWEQTNKV
               * *****      * ** ***  *********     ***   *
```

Figure 1. Sequence alignment of MAT enzymes from *E. coli* (Ecol), *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) and the corresponding residues of Ecol (101–108) which are missing in the crystal structures are indicated by "– | –".

The loop regions were modeled using *de novo* loop generate algorithm or by searching the PDB database (by loop search method). Side chain conformations were searched using the Auto Rotamer Search and the splice regions were optimized by an energy minimization protocol using the consistent valence forcefield (CVFF) [17] as implemented in the *Discover* program (v 2005L, Accelrys Inc., USA) with steepest descents and conjugate gradients methods to a convergence criterion of 0.001 kcal/mole/Å as the maximum derivative, followed by simulated annealing. The

three major regions, N– and C– terminal ends and "active site mobile loop" were further optimized using 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 picoseconds 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/mole/Å^2. The minimization was carried out with 10,000 steps each of steepest descents and conjugate gradients. Finally, the optimized model was checked for the bond lengths, bond angles, torsions and chirality of the Cα atoms using the *ProStat* module. The accuracy and validity of the model was tested with *Profiles–3D* (Accelrys Inc., USA) [18], which calculates a 3D to 1D compatibility score, and graphically portrays the properly folded and misfolded region(s) in the protein structure by performing an Eisenberg analysis [19,20] of the model.

For compatibility with the GOLD program, Sybyl atom and bond types were assigned to the protein atoms. For meaningful results with the GOLD program, it was found necessary to consider the Mg^{2+} ion and this was placed in the homology models by superimposing Asp175, Phe247 and Asp255 (the residue numbering refers to Msm–MAT sequence throughout the article, unless otherwise specified) over corresponding residues in the X–ray structure of *E. coli* (PDB code 1MXB) [5,6], such that the distances of the carboxylate oxygens from the Mg^{2+} atom were set between 2.0 and 2.2 Å. Compatibility in the Sybyl *HINT* program necessitated modification of the N– and C–terminal residues as described in the Sybyl *HINT* manual. The MAT inhibitors (Table 1) reported in the literature [12] were sketched in Sybyl, the atom and bond types were assigned, and all possible tautomeric states were considered. All ligands were geometry optimized to the gradient of 0.001 kcal/mol/Å using the MMFF94 force field.

2.2 Docking Protocol

GOLD (Genetic Optimization for Ligand Docking) is a genetic algorithm (GA) for docking flexible ligands into protein binding sites while considering partial protein flexibility. It uses a genetic algorithm to generate and dock the best conformations of a ligand. Residues in a 10 Å radius from the Mg^{2+} ion (which is coordinated to the carboxylate group of Asp179 and also with one of the carboxylate oxygens of Asp259) were defined as the active site for the docking study. The program was set to terminate when ten solutions within an rmsd of 1.5 Å were identified. The options such as *flip ring corners*, *flip amide bonds*, *flip pyramidal nitrogen* were activated, while unionized carboxylates were allowed to rotate so as to form favorable H–bonds. Docking was

carried out with the default GA settings and the number of GA runs was set to 30. The ten best poses for every ligand were stored for post–docking analysis. The two fitness functions – *GoldScore* [21] and *ChemScore* [21–23] – accessible through the GOLD program were selected for scoring the docking poses.

Table 1. Inhibitors of Msm–MAT with Their % Inhibition Data

Mol ID	Chemical Name	% Inhibition[a]
Methionine analogs		
M1	Cycloleucine	25.8±5.1
M2	Methioninamide	23.9±4.8
M3	Ethionine	20.4±4.1
M4	α–Methyl methionine	18.8±3.7
M5	Methylthiopropanaldehyde	18.4±3.7
M6	Methionine, methyl ester	17.7±3.5
M7	Penicillamine	15.0±3.0
M8	L–Methionine sulfoximine	12.6±2.5
M9	Methionine sulfone	9.2±1.8
M10	Methionine sulfoxide	4.0±0.8
Purine analogs		
P1	6–Methylpurine	0.0±0.0
P2	2–Aminopurine	11.0±2.1
P3	6–Benzyloxy–purine	17.7±3.4
P4	2–Amino–6–chloropurine–9–acetic acid	23.5±8.4
P5	6–Cyanopurine	24.2±4.6
P6	6–Propoxypurine	27.9±4.0
P7	6–Dimethyl aminopurine	28.0±5.3
P8	2,6–Diaminopurine	29.3±7.9
P9	6–Chloropurine	31.4±7.3
P10	6–Bromopurine	31.4±1.5
P11	2–Amino–6–carboxyethyl–mercaptopurine	31.9±9.8
P12	2–Hydroxypurine	33.8±0.6
P13	2,6–Dichloropurine	35.5±6.9
P14	8–Aza–2,6–diaminopurine	40.0±2.2
P15	6–Mercaptopurine	40.1±2.3
P16	O–Methylguanine	60.3±5.3
P17	Azathioprine	75.5±5.4
P18	8–Azaguanine	81.7±2.8
Uric acid and xanthine analogs		
U1	1,3,7–Trimethyluric acid	0.0±0.0
U2	7–Methyluric acid	11.4±2.2
U3	3,7–Dimethyluric acid	27.9±1.6
U4	1–Methyluric acid	43.4±2.9
U5	Uric acid	45.6±11.9
X1	8–Chlorotheophylline	7.0±4.1
X2	Xanthine	35.4±2.5
X3	1–Methylxanthine	35.9±5.8
X4	7–Methylxanthine	36.3±4.2

[a] per cent (%) inhibition of MAT enzyme at 10 mM inhibitor concentration [Ref. 12]

2.3 Rescoring using *HINT*

The best poses ranked by *ChemScore* for each ligand were taken for *H*ydropathic *INT*eraction (*HINT*) [24,25] analysis. The protein was corrected to be recognized during *HINT* calculations as described in the Sybyl *HINT* user manual. The '*calculate*' and '*dictionary*' methods as implemented in Sybyl *HINT* were employed for partitioning ligand molecules and protein molecule respectively. The receptor area covered by the respective ligand pose was considered for calculating the intermolecular *HINT* interaction score. The *HINT* intermolecular scores were calculated using the polar and hydrophobic functions separately as well as in combination.

2.4 Rescoring using *HINT*

The active site of Mtb–MAT comprising residues in a 10 Å radius from the carboxylate carbon atom of Asp179 as the centre was considered in *Ludi* [26] (accessible through Insight II). *Ludi* was run in the 'standard mode' using the *Ludi* CAP2002 library (presently comprising approximately 78,000 fragments). Conformational search of the fragments was done considering two rotatable bonds simultaneously. The remaining parameters were considered at their default settings. The *Ludi* scoring functions [27–33] (*Energy Estimate* 1, 2 and 3) as implemented in *InsightII Ludi* were considered to prioritize the hits from the *Ludi CAP* database.

3 RESULTS AND DISCUSSION

In the post–genomic phase, the exploitation of new biochemical pathways in pathogenic microorganisms for development of prophylactic and therapeutic interventions is an attractive alternative to antibiotics. The 3D structural information of targets is useful in the drug discovery program, as it provides an understanding of the binding interactions at the molecular level. Determination of the structures of all the therapeutically important proteins by experimental methods puts a strain on present day techniques (high–throughput X–ray crystallography or nD–NMR) and the use of theoretical methods of structure generation (such as homology or comparative protein modeling) has proven useful in the structure based drug design studies including docking and scoring during early phase of drug discovery [34–38].

The ligand–receptor binding is a complicated process that poses a challenge to the correct theoretical prediction of the binding of a ligand to its receptor. The prediction of the correct binding pose is possible with most docking programs. However, the success of a docking method is critically dependent on the accuracy of the scoring function that is employed to score these poses (*ranking*). Further, among the available docking programs some do well with certain protein families and fail with others. Therefore, it is necessary in docking studies to use a variety of docking programs with multiple scoring functions to find the one most suitable for the target under investigation. The development of fast and accurate scoring functions is an intensive area of

research these days. As an alternative, to overcome the limitations of the individual scoring function, a Consensus Scoring strategy in which the docked poses are evaluated using different and independently developed scoring functions is often used. The use of multiple scoring functions is advantageous; since each scoring function is developed independently, its application in ranking the ligand poses is similar to repeated sampling. Further, different methods will often agree more on the ranking of the actives than on the inactives, and thus together they strengthen the confidence of the evaluator that the results are not by chance. In this study, we have used the scoring functions accessible through GOLD (*GoldScore* and *ChemScore*) that have been tested on a large number of complexes extracted from the Protein Data Bank (PDB) and found to be accurate in 70–80% of the cases. In addition, we have used the *HINT* scoring function that is reported to be a good indicator of biological activity, to re–rank the poses produced with the program GOLD.

Msm–MAT has a Mg^{2+} ion which is coordinateed to the two carboxylate oxygens of Asp175 (Asp179 in Mtb) and plays a crucial role in the activity of this enzyme. In its recent version, GOLD has been programmed to handle metal ions with a defined coordination geometry. In GOLD, the interaction of a ligand with a metal atom is treated as "*pseudo hydrogen bonding*" *i.e*. GOLD does not consider the interaction between the ligand and the metal atom as a coordinate bond. Dummy atoms are added to the metal atom, which are 'fitting points' used to guide the docking process. Any interaction of the ligand with these 'fitting points' will be scored accordingly within the S(hb_ext) part of the GOLD scoring function. Further, GOLD handles protein flexibility in a limited way – only the conformations of OH groups of Ser and Thr are adjusted so as to form strong H–bonds with the ligand.

3.1 MAT Active Site

Even though the active site of MAT is located in a large cavity at the dimer interface, the methionine (and also the inhibitor) binding site is essentially formed by the residues of one monomer; the residues of the other monomer make up the ATP and phosphate binding sites. X–ray crystallographic studies have shown that methionine analogs bind at the methionine (*i.e*. substrate) binding site [5–9]. Therefore, in the present study, only the substrate binding site (residues within a 10Å radius from the Mg^{+2} ion) was considered for docking studies. The active sites of MAT enzyme of Msm and Mtb are almost identical, except for the two mutations: Asp203 and Gly204 in Msm to Ala207 and Asp208 in Mtb, respectively. However, these residues are ~10 Å away from the center (Mg^{+2} ion) of the active site. The residues His16, Lys177, Asp175, Ser198, Thr244, Phe247, Leu249 and Asp255 (numbering refers to Msm–MAT sequence) are conserved in both the mycobacterial species and form the core of the MAT active site. The site–directed mutagenesis studies, carried out for MAT enzyme from rat, have led to the conclusion that Phe247 (Phe251 in Mtb), Asp175 (Asp179 in Mtb), and Lys177 (Lys181 in Mtb) are critical for SAM synthesis, either by binding to methionine or possibly by being involved in the subsequent steps of the mechanism

[8,9]. These conclusions were based on the fact that their mutations resulted in an absolute loss of MAT activity when compared to the native enzyme.

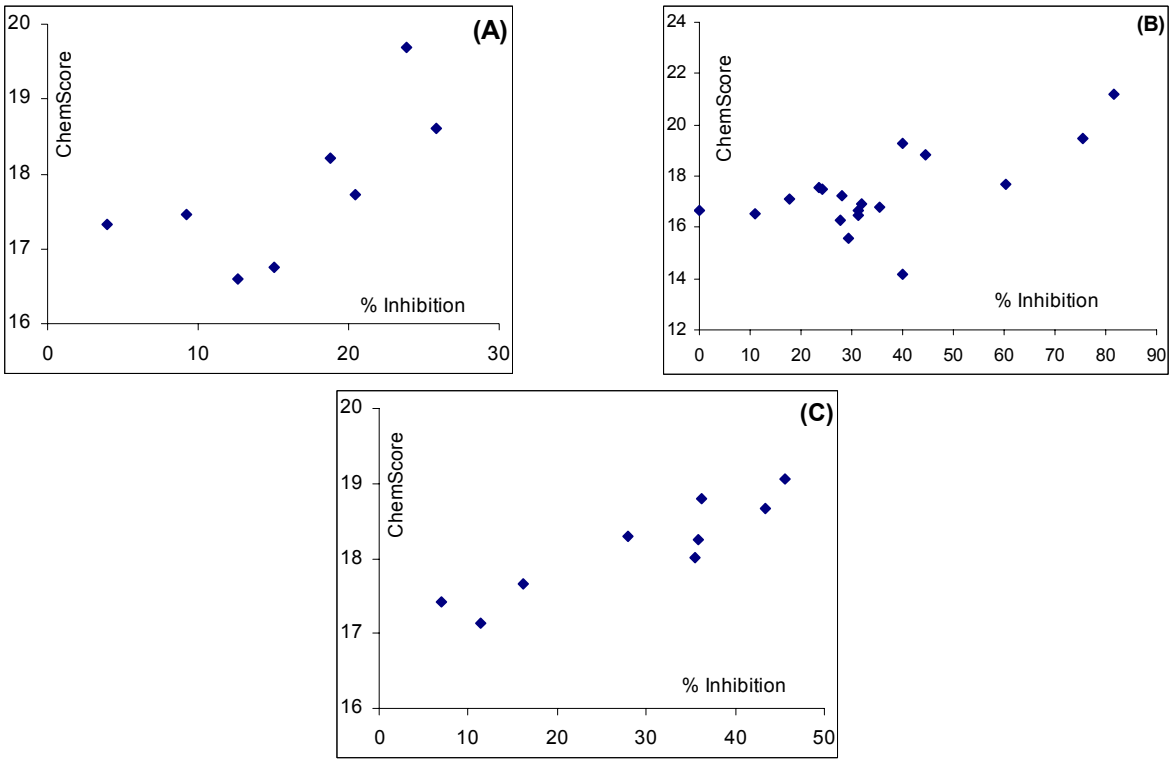


Figure 2. The correlation of *ChemScore* with the % inhibition of Msm MAT by: (A) methionine analogs (correlation statistics: *Pearson R* = 0.69, *Spearman rho* = 0.76, *P* = 0.028, *Kendall's tau* = 0.57, P= 0.048, correlation is significant at the 0.05 level), (B) purine analogs (correlation statistics: *Pearson R* = 0.64; *Spearman rho* = 0.47, *P* = 0.49; *Kendall's tau* = 0.33, P = 0.05, correlation is significant at the 0.05 level), and (C) uric acid and xanthine analogs (correlation statistics: *Pearson R* = 0.92; *Spearman rho* = 0.92, *P* = 0.001; *Kendall's tau* = 0.78, P = 0.05; correlation is significant at the 0.05 level). Refer text for details.

3.2 Post–Docking Analysis

As mentioned earlier, three scoring functions, *GoldScore*, *ChemScore* and *HINT* score, were employed for docking and/or re–scoring of the docked poses. The statistical parameters such as *Pearson R* (referred henceforth to as *R*), *Spearman rho*, *Kendall's tau* were calculated to relate the scores with the experimental activity data. In the present case, the activity of MAT inhibitors has been reported as per cent (%) inhibition of MAT enzyme at 10 mM inhibitor concentration. It is seen that *ChemScore* correlates well with the experimental activity of the Msm–MAT inhibitors. The rescoring of these poses with *HINT* reassures the correlation; an improved correlation statistics for *HINT* over *ChemScore* is achieved. The plots of *ChemScore* and *HINT* score of docked compounds against experimental data are shown in Figures 2 and 3 respectively. The biological activity of the methionine series of inhibitors shows a correlation (*R*) of 0.69 and 0.85 with *ChemScore* and *HINT* score, respectively. The 'metal interaction' and 'lipophilic' terms have major accounts in the *ChemScore* (with *R* of 0.86 and 0.67, respectively) of the inhibitors. In *HINT* scores (which is the sum of polar and hydrophobic interactions) the affinity correlates strongly with polar

interactions (R = 0.84).

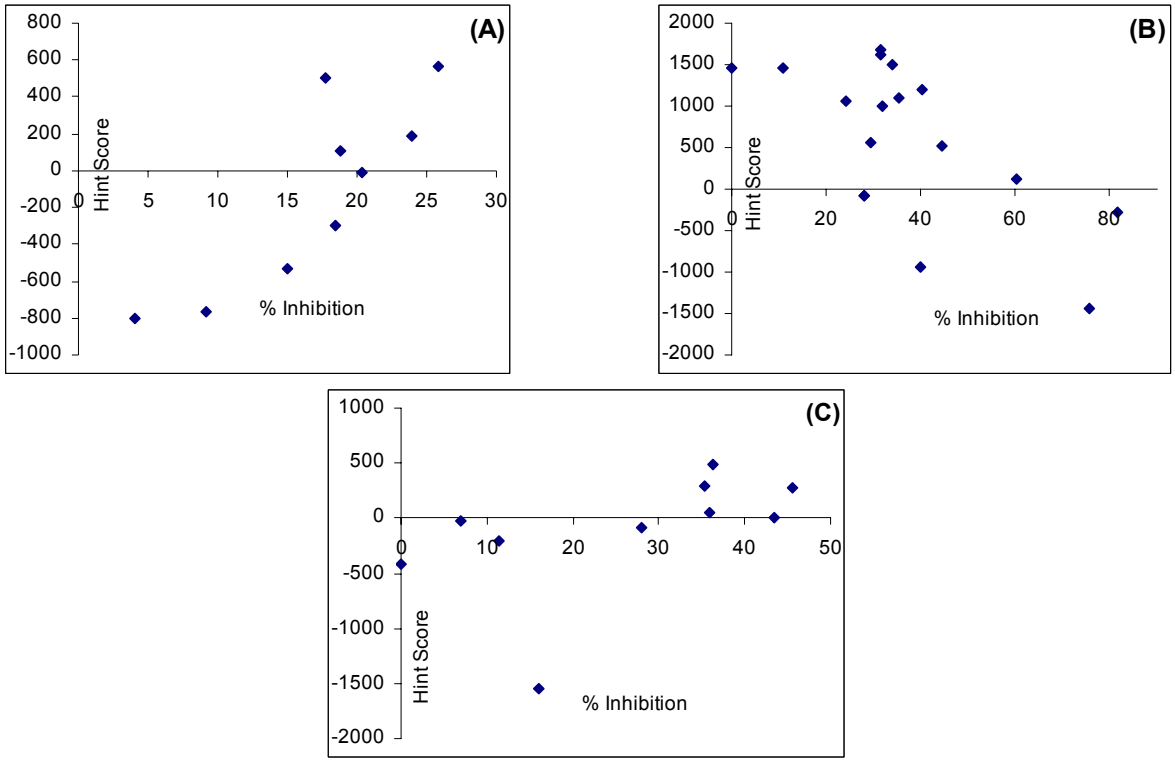


Figure 3. The correlation of *HINT* Score with the % inhibition of Msm MAT by: (A) methionine analogs (correlation statistics: *Pearson R* = 0.85, *Spearman rho* = 0.85, *P* = 0.004, *Kendall's tau* = 0.72, P= 0.007, correlation is significant at the 0.05 level), (B) purine analogs (correlation statistics: *Pearson R* = 0.65; *Spearman rho* = 0.45, *P* = 0.07; *Kendall's tau* = 0.33, P = 0.064, correlation is significant at the 0.05 level), and (C) uric acid and xanthine analogs (correlation statistics: *Pearson R* = 0.53; *Spearman rho* = 0.72, *P* = 0.019; *Kendall's tau* = 0.51, P = 0.04; correlation is significant at the 0.05 level). Refer text for details.

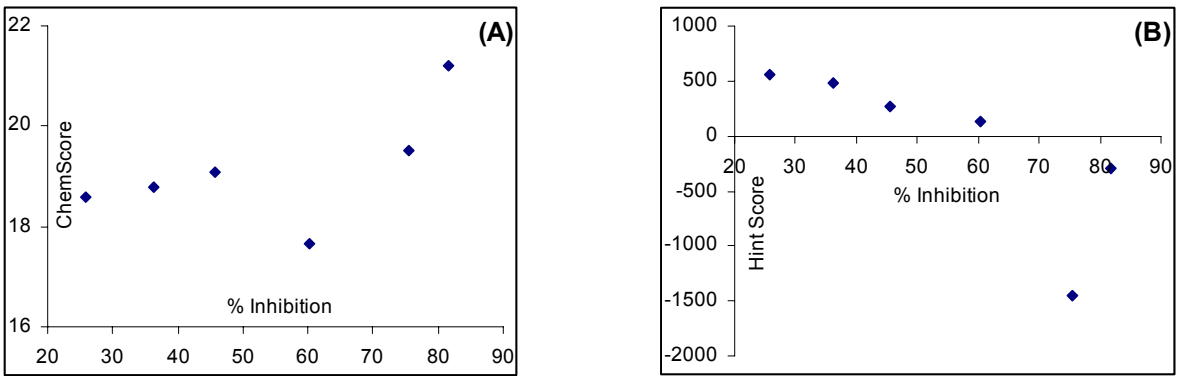


Figure 4. Correlation of *ChemScore* (A) and *HINT* score (B) with % inhibition by representative molecules (molecules **M1**, **X4**, **U5**, **P16**, **P17**, **P18** in Table 1) from each inhibitor class.

The correlation of percent Msm–MAT inhibition by purine analogs with *ChemScore* and *HINT* score exhibit *Pearson R* of 0.64 and 0.65 respectively. In *HINT* score, the polar component has a major contribution (R = 0.64). The correlation between the score and activity (% inhibition) for this class is sufficiently good to discriminate between the active and inactive compounds. The correlation statistics is seen to improve (R = 0.74) when low active compounds (with MAT

inhibition less than 25%) were neglected in the analysis.

The activity of xanthine and uric acid analogs shows a better correlation with *ChemScore* (R = 0.92) than with HINT score (R = 0.75). Figure 4 shows that both *ChemScore* and *HINT* scores are able to discriminate molecules of different chemical classes as active, moderately active and poorly active.

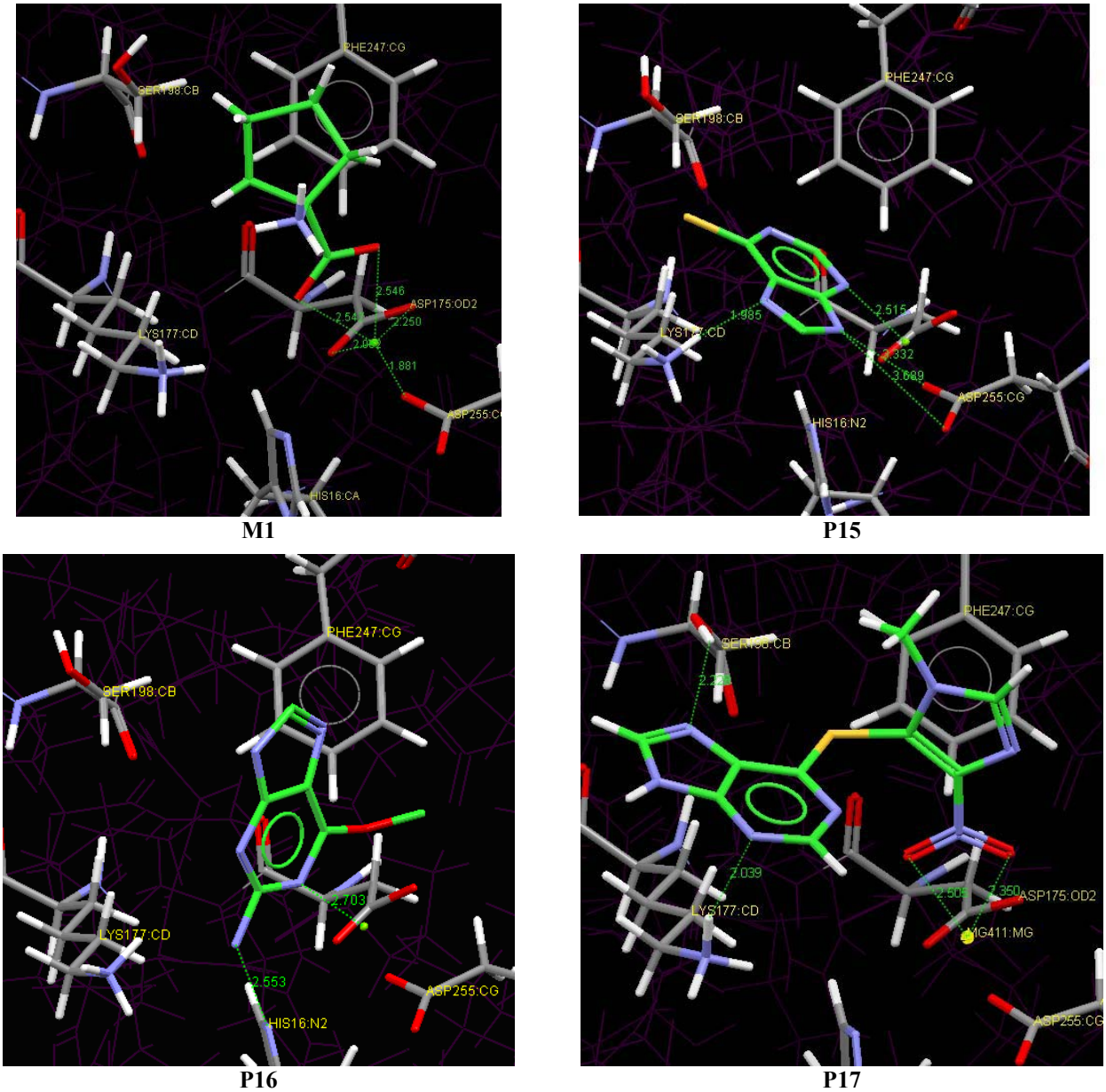
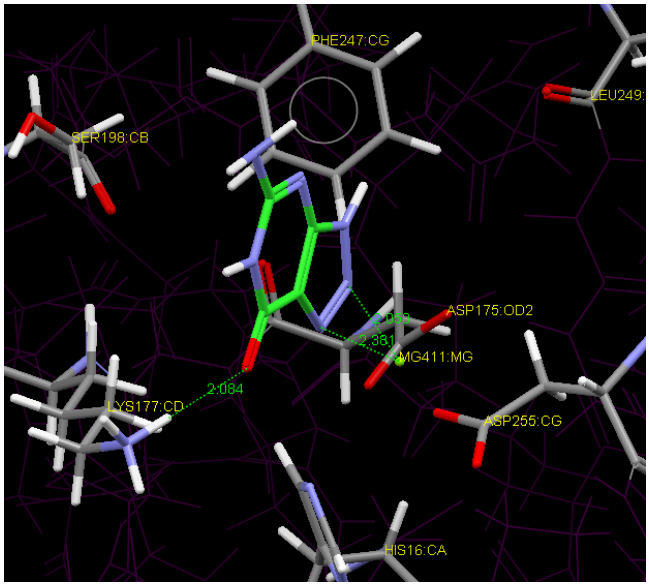


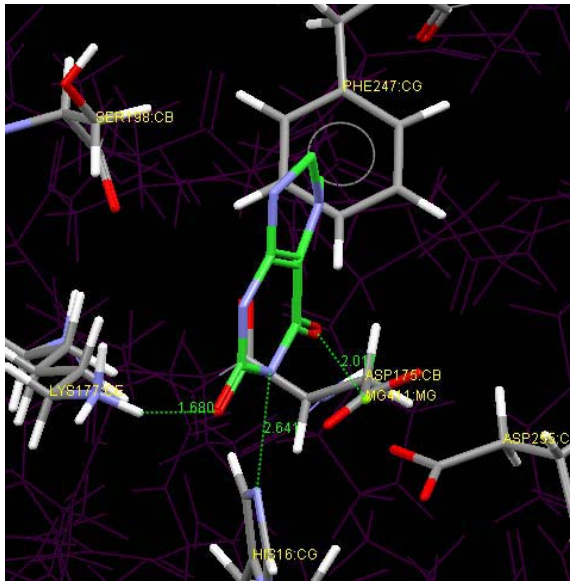
Figure 5. Pictorial views of the most potent inhibitors, cycloleucine (**M1**), 6–Mercaptopurine (**P15**), O–Methylguanine (**P16**), azathioprine (**P17**), 8–azaguanine (**P18**), Uric acid (**U5**) and 7–Methylxanthine (**X4**) in the Msm–MAT active site obtained through docking.

A docking study of the Msm–MAT inhibitors in the active site of the Mtb–MAT model reveals the same pattern of binding configurations and predicted affinities (*ChemScore*s); this is not surprising as the active site of the MAT enzyme from Msm and Mtb are highly conserved. From

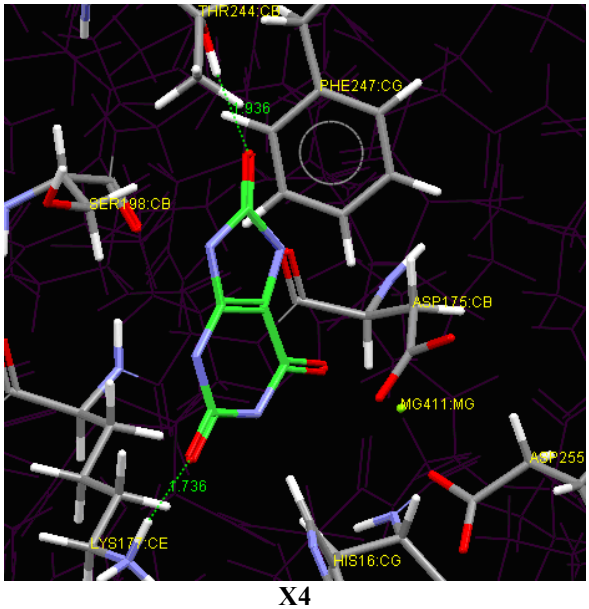
this analysis, it is clear that *ChemScore* can be confidently employed here as a scoring function for docking small organic molecules in 3D–databases. Even though the *HINT* score performs better than *ChemScore*, in that it is well able to discriminate actives from inactives, its value as a scoring function in docking is limited. However, it can be used to more accurately rank a small set of MAT inhibitors that have been filtered based on the *ChemScore*.



P18



U5



X4

Figure 5. (Continued)

The enzyme–inhibitor complexes (Figure 5) were visually analyzed with the Silver interface of the program GOLD. The interactions between the inhibitors and the active site residues in Msm–MAT are described below.

It is observed that the oxygen atoms of the carboxylate group of the methionine analogs interact with Asp175 (Asp179 in Mtb) through coordination with the Mg^{2+} ion (distance ranges from 1.9 to

2.4 Å; see M1 in Figure 5). The aliphatic side chains of the inhibitors are stabilized through hydrophobic interactions with the phenyl ring of Phe247 (Phe251 in Mtb). The aromatic rings of purine, uric acid and xanthine analogs participate in a π–stacking interaction with the phenyl ring of Phe247 (Phe251 in Mtb). Phe247 has been reported to exert a crucial role in the activity of the MAT enzyme [8,9], and it is not surprising that this residue enters into several crucial interactions with the inhibitors. In addition, several important interactions between the Mg^{2+} ion and the electronegative atoms in the inhibitors are seen. In majority of the purine analogs, this occurs through the sp^2 hybridized N3 atom (for example, molecules **P1**, **P2**, **P5**, **P9**, **P10**, **P13**, **P15**, **P16** in Table 1) or the N7 atom (molecules **P3**, **P6**, **P8**, **P14**, **P18**, **X01**, **X2**, **X3** in Table 1). In addition, the N7 atom of some of the inhibitors is also involved in a hydrogen bond with Lys177 (Lys181 in Mtb; molecules **P1**, **P4**, **P9**, **P10**, **P12**, **P15** in Table 1) and with Ser198 (Ser202 in Mtb; molecules **P11**, **P17** in Table 1). In cases where the N7 atom is substitued, such as the uric acid analogs **U2** and **U3** (Table 1), the carbonyl oxygen at C8 coordinates with the Mg^{+2} ion while the N1 atom forms a hydrogen bond with Thr244. In molecules with no substitution at N7 and N9 atoms (molecules **U4**, **U5** in Table 1), the C8 carbonyl oxygen forms a hydrogen bond with Thr244. The molecule **P16** does not coordinate with the Mg^{+2} ion. Azathioprine (molecule **P17** in Table 1) is one of the molecules with high activity and forms hydrogen bonds with Lys177 through the N3 atom and with Ser198 through the N7 atom; the Mg^{2+} ion is stabilized by coordination to the oxygen of the nitro group. The most active inhibitor known for Msm–MAT, 8–azaguanine (molecule **P18** in Table 1) forms a hydrogen bond with Lys177 through the 6–carbonyl oxygen and the Mg^{2+} ion is coordinated with its N7 and N8 atoms.

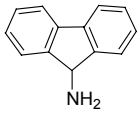
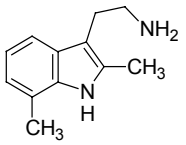
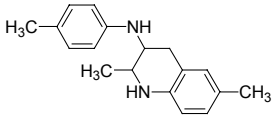
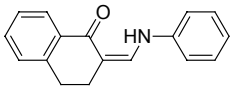
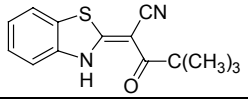
In conclusion, it can be said that while hydrogen bonds with Lys177, Ser198 and Asp255, and hydrophobic/π–stacking interactions with Phe247 are important for strong binding, coordination with Mg^{2+} ion is a minimum requirement for inhibition of the MAT enzyme. A combination of these *features* can be used as a query in a database search.

3.3 Design of Novel MAT Inhibitors

The design of new and selective inhibitors of an enzyme is one of the most important applications in contemporary rational drug design. When information on the binding site of the target protein is available, novel inhibitors can be designed from scratch using *de novo* approaches. As a *de novo* design tool, *Ludi* first generates the interaction sites using a set of rules that are intended to cover the complete range of energetically favorable orientations for hydrogen bonds and hydrophobic contacts. Then using the fragment approach, *Ludi* positions small fragments in the clefts of the protein structure (*i.e.*, active site) such that hydrogen bonds can be formed with acceptor/donor groups, and hydrophobic pockets are filled with hydrophobic groups. A fragment is chosen for its potential to have good hydrophobic and hydrogen bonding interactions with the target receptor. The molecular fragments placed in the active site are then *linked* together to form a

molecule that can fit into the cavity of the protein under study.

Table 2. Some inhibitors of Mtb–MAT designed using *de novo* approach of *Ludi*.

Sr. No.	Structure	Ludi Score	Predicted K_i (micro molar)
1		579	1.58
2		551	3.16
3		415	63.1
4		391	100
5		378	158.5

The structures of novel inhibitors suggested with this approach are shown in Table 2, along with their *Ludi* scores and predicted binding affinities (K_i). Among the scoring functions accessible through *Ludi* for prioritizing fragment hits, Energy Estimate 2 was used for *de novo* design of MAT inhibitors. This scoring function estimates the change in free energy upon binding of the fragment to the receptor, and the score correlates with the dissociation constant, K_i. Each fragment is evaluated as a function of the potential number of hydrogen bonds, hydrophobic and ionic contacts, and a penalty towards the freezing of the internal degrees of freedom of the ligand. A few inhibitors (Table 2) with high scores are predicted with dissociation constant (K_i) values in the mM to μM range (corresponding to *Ludi* 'Energy Estimate 2' scores in the range 300 to 600). The designed molecules bear the properties of lead–like molecules (rule–of–three) [39] with MW below 300, number of H–bond donors/acceptors less than 3 and logP less than 3, so that after lead optimization, the drug–like molecules will fulfill Lipinski's rule–of–five [40]. These moieties offer a basic lead structure for recognition with MAT active site and could be further optimized for potency and selectivity.

4 CONCLUSIONS

The homology models have been used in the absence of any experimental structures of Msm and Mtb–MAT to understand the binding modes of known inhibitors *via* a docking and scoring protocol. A consensus scoring strategy seems best suited for the problem at hand. Though the correlation between activities and scores are statistically significant, there is some room for

improvement. One reason for the lack of a very good correlation of docking scores with the inhibition data could be due to the variation (up to ±12 %) in the reported percentage (%) inhibition data for these inhibitors, which was not considered during the correlation analysis. Our previous study [13] on the molecular electrostatic potential of MAT enzyme from Mtb and humans revealed that the electrostatic potential covering the active site residues (Asp179, Gly16 and His17) have sharply differing values, with the active site of Mtb–MAT being more deeply negative [13]. The pattern of the MEP in the two MATs (mycobacterium and human) indicates that the two proteins would bind inhibitors in the same relative sense, 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. Since the MAT proteins from Msm and Mtb are highly conserved, many of the inferences derived for Msm–MAT can be confidently extrapolated for Mtb–MAT. This study gives an insight into the molecular interactions that play a significant role in the binding affinity of the MAT inhibitors. Further, new lead molecules designed using the *de novo* approach could be developed into useful drugs for treatment of tuberculosis.

Acknowledgment

This work was possible by a grant from the All India Council of Technical Education, New Delhi (F. No. 8022/RID/NPROJ/RPS–5/2003–04), and Department of Science and Technology, New Delhi (Order No. SR/FST/LSI–163/2003). SAK and AKM thank the Council of Scientific and Industrial Research (CSIR), New Delhi for financial assistance.

5 REFERENCES

[1] S. H. Gillespie, MiniReview: Evolution of drug resistance in *Mycobacterium tuberculosis*: Clinical and molecular perspective, *Antimicrob Agents Chemother*. **2002**, *46*, 267–274.

[2] G. S. Besra and D. Chatterjee, Lipids and carbohydrates of *Mycobacterium tuberculosis*, In Tuberculosis: Pathogenesis, protection, and control, (Ed.: B. R. Bloom), ASM Press, Washington, **1994**, pp. 285–306.

[3] The World Health Organization Global Tuberculosis Program. http://www.who.int/tb/en/.

[4] I. Chopra, L. Hesse and A. J. O'Neill, Exploiting current understanding of antibiotic action for discovery of new drugs, *J. Appl. Microbiol*. **2002**, *92* (Suppl.), 4S–15S.

[5] F. Takusagawa, S. Kamitori and 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* **1996**, *35*, 2586–2596.

[6] F. Takusagawa, S. Kamitori, S. Misaki and G. D. Markham, Crystal structure of S–adenosylmethionine synthetase, *J. Biol. Chem*. **1996**, *271*, 136–147.

[7] Z. Fu, Y. Hu, G. D. Markham and F. Takusagawa, Flexible loop in the structure of S–adenosylmethionine synthetase crystallized in the tetragonal modification, *J. Biomol. Struct. Dyn*. **1996**, *13*, 727–739.

[8] B. Gonzales, M. A. Pajares, J. A. Hermoso, L. Alvarez, F. Garrido, J. R. Sufrin and J. Sanz–Aparicio, The crystal structure of tetrameric methionine adenosyltransferase from rat liver reveals the methionine–binding site, *J. Mol. Biol*. **2000**, *300*, 363–375.

[9] B. Gonzales, M. A. Pajares, J. A. Hermoso, D. Guillerm, G. Guillerm and 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*. **2003**, *331*, 407–416.

[10] N. K. Sarkar, S. Shankar and A. K. Tyagi, Polyamines exert regulatory control on mycobacterial transcription: a study using RNA polymerase from *Mycobacterium phlei*, *Biochem. Mol. Biol. Int*. **1995**, *35*, 1189–1198.

[11] G. L. Cantoni, Biological methylation: Selected aspects, *Annu. Rev. Biochem*. **1975**, *44*, 435–451.

[12] B. J. Berger, M. H. Knodel, Characterization of methionine adenosyltransferase from *Mycobacterium smegmatis*

and *M. tuberculosis*, *B.M.C. Microbiol*. **2003**, *3*, 12–24.

[13] S. A. Khedkar, A. K. Malde and E. C. Coutinho, Comparative protein modeling of methionine S–adenosyltransferase (MAT) enzyme from Mycobacterium tuberculosis: a potential target for antituberculosis drug discovery, *J. Mol. Graph. Model*. **2005**, *23*, 355–366.

[14] G. Jones, P. Willett, R. C. Glen, A. R. Leach and R. Taylor, Development and validation of a genetic algorithm for flexible docking, *J. Mol. Biol*. **1997**, *267*, 727–748.

[15] G. Jones, P. Willett and R. C. Glen, Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation, *J. Mol. Biol*. **1995**, *245*, 43–53.

[16] G. Jones, P. Willett and R. C. Glen, A genetic algorithm for flexible molecular overlay and pharmacophore elucidation, *J. Comput.–Aided Mol. Des*.**1995**, *9*, 532–549.

[17] P. Dauber–Osguthorpe, V.A. Roberts, D. J. Osguthorpe, J. Wolff, M. Genest, and A. T. Hagler, Structure and energetics of ligand binding to proteins: *Escherichia coli* dihydrofolate reductase–trimethoprim, a drug–receptor system. *Proteins* **1988**, *4*, 31–47.

[18] R. Luthy, J. U. Bowie, and D. Eisenberg, Assessment of protein models with three–dimensional profiles. *Nature* **1992**, *356*, 83–85.

[19] M. Gribskov, A. D. Mclachlan, and D. Eisenberg, Profile analysis: detection of distantly related proteins. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 4355–4358.

[20] M. Gribskov, R. Luthy, and D. Eisenberg, Profile analysis. *Meth. Enzymol*. **1990**, *183*, 146–159.

[21] GOLD (v3.1) User Guide, CCDC, UK.

[22] M. D. Eldridge, C. W. Murray, T. R. Auton, G. V. Paolini and R. P. Mee, Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes, *J. Comput.–Aided Mol. Des*., **1997**, *11*, 425–445.

[23] C. A. Baxter, C. W. Murray, D. E. Clark, D. R. Westhead and M. D. Eldridge, Flexible docking using TABU search and an empirical estimate of binding affinity, *Proteins* **1998**, *33*, 367–382.

[24] G. E. Kellogg, S. F. Semus and D. J. Abraham, *HINT*: a new method of empirical hydrophobic field calculation for CoMFA, *J. Comput.–Aid. Mol. Des*. **1991**, *5*, 545–552.

[25] G. E. Kellogg, G. S. Joshi and D. J. Abraham, *Med. Chem. Res*. **1992**, *36*, 444–453.

[26] LUDI, Release 2005, USER GUIDE, Accelrys Inc., San Diego, CA, USA.

[27] H. J. Bohm, A novel computational tool for automated structure–based drug design, *J. Mol. Recognit*. **1993**, *6*, 131–137.

[28] H. J. Böhm, The computer program LUDI: A new method for the de novo design of enzyme inhibitors, *J. Comput.–Aided Mol. Des*. **1992**, *6*, 61–78.

[29] Böhm, H. J. LUDI: rule based automatic design of new substituents for enzyme inhibitor leads, *J. Comput.–Aided Mol. Des*. **1992**, *6*, 593–606.

[30] H. J. Böhm, The development of a simple empirical scoring function to estimate the binding constant for a protein–ligand complex of known three–dimensional structure, *J. Comput.–Aided Mol. Des*. **1994**, *8*, 243–256.

[31] H. J. Böhm, On the use of LUDI to search the Fine Chemicals Directory for ligands of proteins of known three–dimensional structure, *J. Comput.–Aided Mol. Des*., **1994**, *8*, 623–632.

[32] M. Stahl and H. J. Bohm, Development of filter functions for protein–ligand docking, *J. Mol. Graph. Model*. **1998**, *16*, 121–132.

[33] H. J. Bohm, Prediction of binding constants of protein ligands: a fast method for the prioritization of hits obtained from de novo design or 3D database search programs, *J Comput.–Aided Mol. Des*. **1998**, *12*, 309–323.

[34] A. Hillisch, L. F. Pineda and R. Hilgenfeld, Utility of homology models in the drug discovery process, *Drug Discov. Today*. **2004**, *9*, 659–669.

[35] L. Miguet, Z. Zhang, M. Barbier, M. G. Grigorov, Comparison of a homology model and the crystallographic structure of human 11beta–hydroxysteroid dehydrogenase type 1 (11betaHSD1) in a structure–based identification of inhibitors. *J Comput Aided Mol Des*. **2006**, *20*, 67–81.

[36] M. S. Gomaa, S. W. Yee, C. E. Milbourne, M. C. Barbera, C. Simons, A Brancale, Homology model of human retinoic acid metabolising enzyme cytochrome P450 26A1 (CYP26A1): active site architecture and ligand binding. *J. Enzyme Inhib. Med. Chem*. **2006**, *21*, 361–369.

[37] W. M. Rockey, A. H. Elcock, Structure selection for protein kinase docking and virtual screening: homology models or crystal structures? *Curr. Protein. Pept. Sci*. **2006**, *7*, 437–457.

[38] A. Pedretti, G. Vistoli, C. Marconi, B. Testa, Muscarinic receptors: a comparative analysis of structural features and binding modes through homology modelling and molecular docking. *Chem. Biodivers*. **2006**, *3*, 481–501.

[39] M. Congreve, R. Carr, C. Murray and H. Jhoti, A 'rule of three' for fragment–based lead discovery?, *Drug Discov. Today* **2003**, *8*, 876–877.

[40] C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Adv. Drug Del. Rev.* **2001**, *46*, 3–26.

Biographies

Santosh A Khedkar has completed Ph.D. (Tech.) from the Bombay College of Pharmacy, University of Mumbai, under the guidance of Prof. Evans Coutinho. He obtained an M. Pharm. Sci. degree in Pharmaceutical Chemistry at the University Institute of Chemical Technology (UICT), Mumbai. His research interests include homology modeling, docking studies, development of new QSAR methods and *ab initio* studies in peptide design. He has presented papers in several national and international conferences and published in reputed international journals.

Alpeshkumar K Malde has completed Ph.D. (Tech.) from the Bombay College of Pharmacy, University of Mumbai, under the guidance of Prof. Evans Coutinho. His research interests include applications of *ab initio* calculations, protein modeling, new QSAR methods, solution structures of peptides by 2D–NMR and design of novel boron peptides. He has presented / published papers in reputed conferences and journals. He is a recipient of two Gold Medals awarded by the Indian Pharmaceutical Association (IPA) and the National Institute of Pharmaceutical Education and Research (NIPER).

Evans C Coutinho is Professor of Pharmaceutical Chemistry at the Bombay College of Pharmacy. He received his Ph.D. degree from the University of Mumbai and Masters and Bachelors degree in Pharmaceutical Technology from the University Institute of Chemical Technology (UICT). His research interests are in the areas of QSAR, docking, structure–based drug design, quantum mechanical calculations, and multinuclear and multidimensional NMR of proteins and peptides. He has received the Cipla Distinguished Fellowship Award, the UGC Research Award, and the Career Award for Young Teachers from the All India Council of Technical Education.