Comparative Modeling of Methionine S-adenosyltransferase A potential drug targeted protein of Mycobacterium

Anuradha Chevva Moremmagari

Assistant Professor,
Department of Biotechnology,
Sri Krishnadevaraya University, Anantapur-515 003. AP, India.

Abstract: Methionine S-adenosyltransferase plays a critical role in the active and dormant phases of mycobacterium life cycle which makes it an attractive antitubercular target. MAT is one of the 256 enzymes deemed necessary for the life. *Mycobacterium tuberculosis* (Mtb), causative organism of tuberculosis (TB), is a successful pathogen that overcomes the numerous challenges by the immune system of the host. In the last 40 years few 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 antiTB drug target, Mtb-MAT has been well validated.

IndexTerms **- TB drugs, homology modeling, Drug Design.**

Introduction

Over the past few decades, major advances in the field of molecular biology, coupled with advances in genomic technologies, have led to an explosive growth in the biological information generated by the scientific community. This deluge of genomic information has, in turn, led to an absolute requirement for computerized databases to store, organize, and index the data and for specialized tools to view and analyze the data.

The Function of a protein is directly linked to it's 3-D structure. If the tertiary structure is changed, the protein normally loses the ability to perform it's biological function, since this function depends on the geometrical shape of the active site in the interior of the molecule (lock-key principle).

Methionine S-adenosyltransferase plays a critical role in the active and dormant phases of mycobacterium life cycle which makes it an attractive antitubercular target. MAT is one of the 256 enzymes deemed necessary for the life. *Mycobacterium tuberculosis* (Mtb), causative organism of tuberculosis (TB), is a successful pathogen that overcomes the numerous challenges by the immune system of the host. In the last 40 years few 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 antiTB drug target, Mtb-MAT has been well validated.

Comparative or homology modeling or knowledge-base prediction exploits the fact that evolutionarily related proteins with similar sequences have similar structures. The degree of similarity is very high in the so called "core regions" comprising of secondary structural elements (" helices and sheets ") whereas the degree of similarity is usually low in loop regions connecting secondary structures.

While the high precision structures required for detailed studies of protein-ligand interaction can only be obtained experimentally, theoretically predicted models provide molecular biologists with “low resolution" models which hold enough information about preferred spatial arrangements of important residues to guide the design of experiments. Thus even though the current methods are still in their infancy, prediction of structures for all protein sequences of complete genomes in conjunction with experimental work is a realistic goal. Structural analyses of proteins for further mutagenesis, substrate and inhibitor

design, and enhanced function and stability are also possible. These methods can use structural data to probe organism function and evolution.

Functional characterization of a protein sequence is one of the most frequent problems in biology. This task is usually facilitated by accurate three dimensional (3D) structure of the studied protein.

These methods rely on detectable similarity spanning most of the modeled sequence and at least one known structure. When the structure of one protein in the family has been determined by experiment, the other members of the family can be modeled based on their alignment to the known structure. Comparative or homology protein structure modeling builds a three dimensional model for a protein of unknown structure (the target) based on one or more related proteins of known structure (the templates).

Methodology

Due to unavailability of X-ray/NMR crystallographic structure in PDB, a homology model of MAT has been constructed using the X-ray structures of *E.coli* MAT (PDB code: 1XRA) as template, by comparative protein modeling principles. Finding structural homolog to the target sequence. Comparative modeling usually starts by searching the PDB of known protein structures using the target sequence as the query. This search is generally done by comparing the target sequence with the sequence of each of the structures in the database. Once a list of potential templates is obtained using searching methods, it is necessary to select one or more templates that are appropriate for the particular modeling problems.

Tools used

BLAST (Basic Local Alignment Search Tool) is a set of similarity search programs designed to explore all of the available sequence databases regardless of whether the query is protein or DNA. The BLAST programs have been designed for speed; with a minimal sacrifice of sensitivity to distant sequence relationships .The scores assigned in a BLAST search have a well designed statistical interpretation. Making real matches easier to distinguish from random background hits. BLAST uses a heuristic algorithm, which seeks local as opposed to global alignments and is therefore able to detect relationships among sequences, which share only isolated regions of similarity.

PSI BLAST : Position Specific Iterative BLAST (PSI-BLAST) refers to a feature of BLAST 2.0 in which a profile (or Position Specific Scoring Matrix PSSM) is constructed (automatically) from a multiple alignment of the highest scoring hits in an initial BLAST search.The PSSM is generated by calculating position-specific scores of each position in the ali-gnment. Highly conserved positions recieve high scores and weakly conserved positions recieve scores near zero.The profile is used to perform a second BLAST search and the results of each 'iteration' used to refine the profile. This iterative searching strategy results in increased sensitivity.

MODELLER^{R} is a computer program that models 3D structure of proteins by satisfaction of spatial restraints. MODELLER^{R} is most frequently used for homology or comparative protein structure modeling. The user provides an alignment of a sequence to be modeled with known related structures and MODELLER^{R} will automatically calculate a model with all non-hydrogen atoms.
A 3D model is obtained by optimization of a molecular probability density function (pdf). The molecular pdf for comparative modeling is optimized with the variable target function procedure in Cartesian space that employs methods of conjugate gradients and molecular dynamics with simulated annealing. MODELLER^{R} can also perform multiple comparison of protein sequences and/or structures, clustering of proteins, and searching of sequence databases.

Getting the sequence file

The sequence of the enzyme, **Methionine S-adenosyltransferase** MAT of Mtb H37Rv, was obtained in the FASTA format using NCBI database.

Sequence analysis

The sequence analysis for the sequence obtained from NCBI, was done using the following servers.

Pfam: The protein sequence was submitted to the server for protein family identification. The Pfam results again confirms the presence of three domains of S-Ado-Met synt class.
Prosite: The sequence was submitted to Prosite database to find similar patterns. There were two recognizable patterns.

Results

The built 3-D model of Methionine S-adenosyltransferase has the correct stereochemistry as gauged from the Ramachandran plot (Fig.1) and good three dimensional (3D) structure compatibility as assessed by the *verify-3D* score and PROCHECK. The structurally and functionally important residues of Mtb-MAT have been identified using the *E.coli* crystal structure. The homology model conserves the topological and active features of the MAT family of proteins. Homology modeling is a multistep process which converts a linear amino acid sequence of a protein into a 3 dimensional structure. The amino acid sequence of Methionine S-adenosyltransferase (target) was taken from GenBank (NCBI) and subjected to similarity search through BLAST group of programs to identify the sequences with highest homology. The most homologous sequence is the one with lowest E value among all the hits. This sequence was selected as the template. The FASTA format of both the target and template was loaded on CLUSTALX software to obtain the alignment. This program gives an alignment file.

After a model is built, it is important to check it for possible errors. The quality of a model can be approximately predicted from the sequence similarity between the target and the template. Sequence identity above 30% is a relatively good predictor of the expected accuracy of a model. However, there are other factors too which can strongly influence the accuracy of a model.

Understanding the molecular function of proteins is greatly enhanced by insights gained from their three-dimensional structures. Since experimental structures are only available for a small fraction of proteins, computational methods for protein structure modeling play an increasingly important role. Comparative protein structure modeling is currently the most accurate method, yielding models suitable for a wide spectrum of applications, such as structure-guided drug development or virtual screening.

Combinatorial chemistry and molecular modeling are not only used in pharmaceutical companies but also have a broad range of applications in clinical diagnostic, immunology and molecular biology. Moreover, these technologies have potential applications in agriculture and polymer chemistry.

Molecular modeling is a rational approach requiring very low investments, thus it can be easily applied in R&D as well as in industries (even small and medium) in developing countries and countries in transition working on Parma, agro chemistry and natural product exploitation.

Energy Minimization

The Model structure could be refined in various ways using several tools available in SPDBV. Energy minimization is the first and best way of refining the modeled structures. It relaxes the clashing amino acids (Fig.1).
The energy of the model protein is completed with the following command:
Tools>Compute Energy (Force Field)
The energy minimization was carried out for the modeled MAT structure, using the following command (Select>All residues before performing Energy Minimization).
Tools>Energy Minimization
The energy of the modeled molecule before and after minimization is compared. Obviously, the energy has been reduced after minimization.

Energy Before minimization: 2408.190 KJ/mol
Energy After minimization: -11384.669 KJ/mol

Discussion

Tuberculosis represents one of the world's greatest sources of mortality and morbidity, with approximately 8 million new infections and 2 million deaths per year (Dye, 1999). The situation regarding

the control of tuberculosis has significantly worsened over the last decade, with the spread of strains resistant to multiple antimycobacterial agents. There is a profound need for the identification and development of novel chemotherapeutic compounds against tuberculosis. The characterization of mycobacterial biochemical pathways aids this process through the identification of enzymes amenable to therapeutic inhibition.

Mycobacterium tuberculosis is difficult to kill for a number of reasons. The organism is surrounded by a dense waxy coat consisting of unusual long-chain fatty acids (mycolipids) with hydroxyl, methyl, and cyclopropyl substitutions that prevent many common antibiotics from entering the cell (Bersa, 2015). In addition, the organism normally resides in the unfused lysosome of macrophages, which further complicates access by antibiotics. Finally, the bacterium is able to enter a very slow-growing, chronic phase, where many biochemical targets are down-regulated (Parrish, 1998). In this state, the bacteria shift their metabolic focus from sugars to β-oxidation of fatty acids, which entails a down-regulation of glycolysis and an up-regulation of the glyoxylate shunt (McKinney, 2000). Therefore, in order to cure tuberculosis, an active compound must penetrate the macrophage, the bacterial coat, and be active against both the acute and chronic growth phases. For these reasons, antimycobacterial therapy relies on the combination of several drugs.

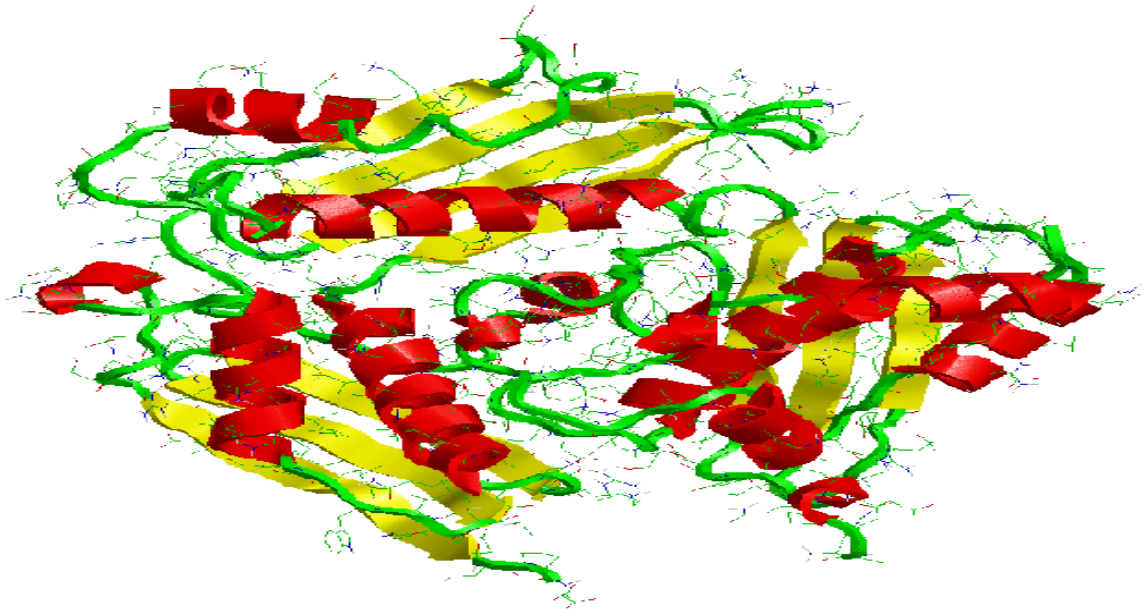
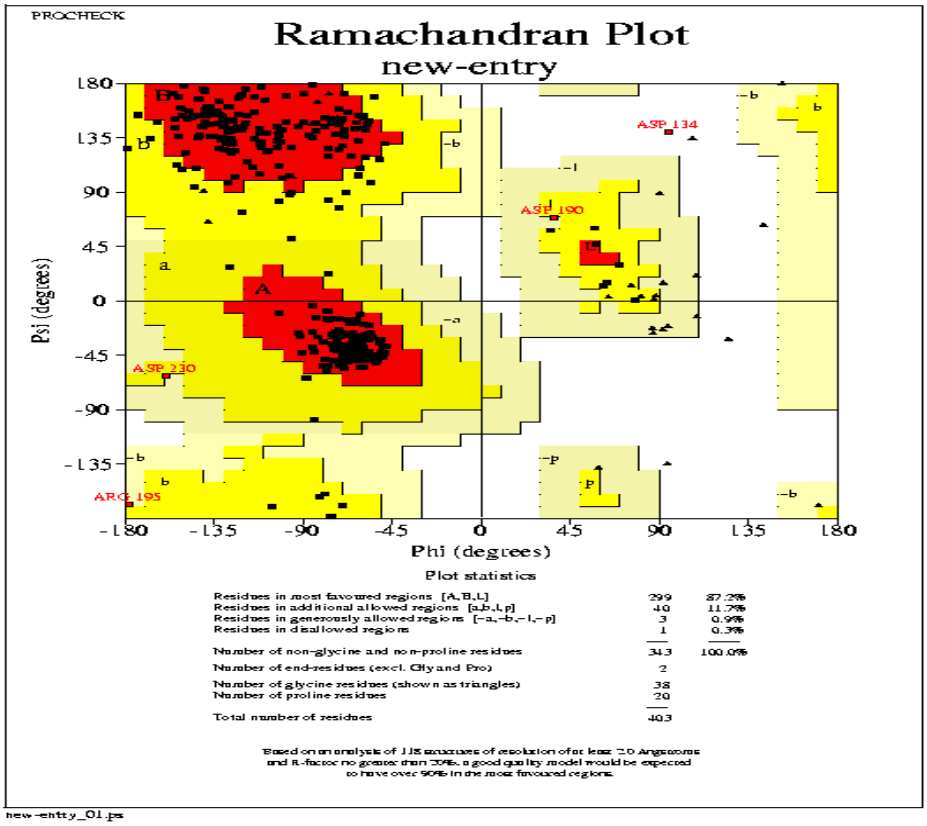


Fig.1 3D structure of Built Model of Methionine S-adenosyltransferase



In the examination of biochemical pathways in *Mycobacterium tuberculosis*, it would be ideal to identify processes where an enzyme plays a role in both active and chronic phase survival. In active, replicative growth cells require polyamines for cell division. While the exact function of these molecules is unknown, it is hypothesized that the positively charged spermidine and spermine act to stabilize DNA during unwinding and strand separation (Marton, 2014). In mycobacterium, polyamines may also play a role in transcriptional regulation (Sarkar, 2014) and have also been targeted for chemotherapeutic intervention (Paulin, 2015; Poso, 2013). In the biosynthesis of polyamines, decarboxylated S-adenosylmethionine acts as an aminopropyl donor for the formation of spermidine from putrescine, and of spermine from spermidine. These reactions give rise to methylthioadenosine, which can be recycled back to adenine and methionine for further synthesis of S-adenosylmethionine (SAM).

SAM is one of the most important cellular biochemical cofactors and plays a role in a large variety of essential metabolic pathways. The formation of SAM from methionine and ATP by Methionine S-adenosyltransferase. Therefore it represents a crucial checkpoint for numerous functions required for cell growth and division, such as polyamine biosynthesis and methylation reactions. Not surprisingly MAT is a very highly conserved enzyme and displaces a high sequence identity from bacteria through to humans. Even bacteria with known degenerate, minimal genomes such as *Mycoplasma Sps* (Fraser, 1995; Himmelreich, 1996), *B. aphidocola* (Shigenobu,2014), and *M. leprae* (Cole, 2016) contain a sequence with a high identity to MAT. The only exception to the ubiquity of MAT is the archaeobacteria which perform this enzymatic function with a highly divergent enzyme which shares only the active site residues with the *E. coli* MAT(Graham, 2014).

In mycobacterium, SAM plays an additional role beyond normal cellular methylation and aminopropylation reactions, as the organisms are reliant on the cofactor for the formation of methylated and cyclopropylated mycolic acids. These fatty acids are very long, and consist of 70–90 carbons (Bersa, 2015) and contain methyl, hydroxyl, and cyclopropyl substitutions that are diagnostic for individual mycobacterial species (Butler, 2017). In *M. tuberculosis*, there are as many as seven SAM-dependent methyltransferases involved in mycolic acid methylation and cyclopropylation (Cole, 2015). Interference with cyclopropyl formation in mycolic acid synthesis has been shown to impact virulence, persistence, and resistance of M. tuberculosis to oxidative stress (Glickman, 2016). When coupled with the role of SAM as an aminopropyl donor for polyamine biosynthesis during cell division (Marton, 2014), interference with SAM has the potential to impact both the active and persistent phases of tuberculosis. The built model

from the present study is making way for the interaction with proposed lead molecules which further results in developing more efficient drugs for control of Tuberculosis.

Acknowledgment

Author acknowledge the help rendered by HOD, Department of Biochemistry, Sri Krishnadevaraya university, Anantapur, AP, India, for providing facilities and structure prediction software.

References:

Altschul SF et al., J Mol Biol, 215: (2018); 403-410.
Altschul SF et al., Nucleic Acids Res, 25: (2019) ; 3389-3402.
Bersa G.S. et al., ASM *Press, Washington*, 2015, pp. 285-306. World Health Organization (WHO), Tuberculosis Fact Sheet (No. 104), 2015.
Butler WR, Guthertz LS, *Clin Microbiol Rev* 2017, **14:**704-26.
Cantoni.G.L, Annu. Rev. Biochem. 44 (2009) 435-451.
Chopra.P et al., *Indian J. Med. Res.* 117 (2018) 1-9.
Cole S.T. et al., *Nature* 2015, **393:**537-44.
Cole S.T et al., *Nature* 2016, **409:**1007-11.
Corbet.E.L et al., *Arch. Int. Med*. 163 (2013) 1009-1021.
Dye.C et al., *JAMA* 281 (2018) 677-686.
Diana L et al., Antimicrob. Agents Chemother. 2017. 38: 2380-2386.
Fraser C.M et al., *Science* 2013, **270:**397-403
Glickman MS, Cox JS, Jacobs WR Jr, *Mol Cell* 2016, **5:**717-27.
Graham D.E et al., *J Biol Chem* 2014, **275:**4055-9.
Himmelreich .R et al., *Nucl Acids Res* 1996, **24:**4420-49.
Jin, D. J., and C A Gross. 2018. J. Mol. Bio. 202:45-58.
Jin, D.J., and C. A Gross. 2017. J. Bacteriol.171: 5229-5231.
Jones G et al., J Mol Biol 2019; 267:727–748.
Kotb.M et al., *Pharmacol. Ther*. 59 (1993) 125-143).
Laskowski.R.A et al, J Appl Crystallogr 26: 283 (19193).
Lüthy R, Bowie JU, Eisenberg D, Nature, 356 (1992); 83-85.
Laskowski.R.A RA, MacArthur MW, Moss DS & Thorton JM. PROCHECK: a program to check the stereo chemical quality of protein structures. J. Appl. Cryst. 26. (1993); 283-291.
Mato.J.M et al., *Pharmacol. Ther.* 73 (2019) 265-280.
Marton.L.J et al., *Ann. Rev. Pharmacol. Toxicol*. 35 (2014) 55-91.
Needleman SB et al., J Mol Biol, 48: (1970); 443–453
Parrish NM, Dick JD, Bishai WR, *Tr Microbiol* 2012, **6:**107-12
Philip.W.J et al., Proc. Natl. Acad. Sci. U.S.A. 93 (2014) 3132-3137.
Sarkar.N.K et al., *Biochem.Mol.Bio.Int*.35 (2014) 1189-1198.
Sali & T. L. Blundell, J. Mol. Biol., 234: 779 (1993)
Shigenobu .S et al., *Nature* 2016, **407:**81-6.
SYBYL 6.7, Tripos Inc., 1699 South Hanley Rd., St. Louis, Missouri, 63144, USA
Tabor.C.W et al., *Adv. Enzymol*. 56 (2002) 251-282.
Takusagawa.F et al., Biochemistry 35 (2012) 2586-2596.
Takugasawa.F et al., J. Biol. Chem. 271 (2011) 136-147.
Thompson JD et al., Nucleic Acids Res, 25: (2009); 4673-4680.
WHO, Tuberculosis Fact Sheet (No. 104), 2020.