© 2020 | International Journal of Scientific Research in Science and Technology
IJSRST | Volume 7 | Issue 1 | Print ISSN: 2395-6011 | Online ISSN: 2395-602X
DOI : https://doi.org/10.32628/IJSRST

Homology Modeling of Salicyl-AMP ligase involved in *Mycobacterium Tuberculosis*

Yogesh N. Joshi[1*], Shubham S. Gore [2], Akshay S. Rajeshwar[3]

[1 & 3]Department of Bioinformatics, Walchand Centre for Biotechnology, Walchand College of Arts & Science, Solapur, Maharashtra,India.

[2]Department of Biotechnology, Walchand Centre for Biotechnology, Walchand College of Arts & Science, Solapur, Maharashtra,India.

ABSTRACT

Tuberculosis (TB) is caused by *Mycobacterium tuberculosis* (strain ATCC 25618/H37Rv) that most often affect the lungs, but can also affect other parts of the body. A total of 1.5 million people died from TB in 2018 Worldwide. TB is spread from person to person through the air. The high prevalence of TB obligates the identification of new therapeutic targets and the development of anti TB vaccines that can control multidrug resistant and latent TB infection. Membrane protein has recently been suggested as key targets for bacterial viability. However, *M. tuberculosis* has a thick, waxy mycolic acid capsule that protects it from these toxic substances. *Mycobacterium tuberculosis* is able to reproduce inside the macrophage and will eventually kill the immune cell. Its treatments require the use of multiple antibiotics over long periods of time. Antibiotics resistance is a growing problem with increasing rates of multiple drug resistant tuberculosis (MDR-TB) and extensively drug resistant tuberculosis (XDR-TB). Subsequent work suggests that Salicyl-AMP ligase is involved in the initial steps of the mycobactin biosynthetic pathway. In the present study, we used different *In-Silico* tools and techniques which include retrieval of Salicyl-AMP ligase protein sequence from UniProt KB database and the sequence analysis was performed by using Protparam tool which concluded that the protein was unstable and basic in nature. The secondary structure was predicted by using PSI-PRED tool which indicated that the percentage of coils was higher than the percentage of alpha helix and extended strand. Then the 3D structure of Salicyl-AMP ligase was predicted by using SWISS MODEL server and the model was validated by using PROCHECK analysis tool. After validation of the model, the validation score was 90% indicating that the model was of good quality and the predicted 3D structure was deposited in protein model database (PMDB). Further study on Salicyl-AMP ligase protein can be carried out in molecular docking and in structure based drug designing (SBDD) to inhibit tuberculosis.

Keywords : Homology Modeling, *In-Silico*, Salicyl-AMP ligase, Tuberculosis, UniProtKB

I. INTRODUCTION

Tuberculosis (TB) is a contagious airborne disease caused by *mycobacterium tuberculosis* (strain ATCC 25618/H37Rv). The classic symptoms of active TB are a chronic cough with blood containing mucus, fever, night sweats and weight loss [1]. The association of *Mycobacterium tuberculosis* (strain ATCC 25618/H37Rv) infections with HIV is the main cause of death of these patients because of the weak immune response. Further, the rise of multidrug resistance (MDR) tuberculosis creates severe disease

condition, thus making current therapies in effective [2]. Chronic lung disease is another significant risk factor. Silicosis increases the risk about 30-fold [3]. Those who smoke cigarettes have nearly twice the risk of TB compared to nonsmokers[4]. Controlling TB is complicated due to several factors, of which the most important is the capacity of *Mycobacterium tuberculosis* (strain ATCC 25618/H37Rv) to persist for long periods of time in adverse conditions into macrophages [5]. The outcome of *mycobacterium tuberculosis* infection would be represented as a distribution between active TB and latent TB on the basis of the presence or absence of symptoms [6]. Thus, it is essential to identify specific therapeutic targets that can be used to control latent and active TB infections. In this regard, a new anti-TB drug should achieve the following criteria: (i) shorten the duration and dose of treatment; (ii) be active against both MDR and XDR *mycobacterium tuberculosis* strains; (iii) be able to eradicate latent TB; and (iv) be capable of being coadministered with drugs to treat HIV [7]. MbtA (Salicyl-AMP ligase) protein is involved in the pathway mycobactin biosynthesis, which is part of Siderophore biosynthesis. *Mycobacterium tuberculosis* synthesizes a suite of structurally related small molecule iron chelating agents (i.e. siderophores) collectively known as the mycobactins that vary by the appended lipid residue on the central lysine moiety [8]. *Mycobacterium tuberculosis* requires iron for its survival; however, the concentration of free iron is highly restricted in biological fluids due to the insolubility of iron under aerobic conditions and sequestration by iron-binding proteins such as transferrin [9]. *Mycobacterium tuberculosis* has iron acquisition systems based on siderophores, secreted iron-chelating compounds with extremely high $Fe3^{+}$ affinity. Siderophores have a critical role in bacterial iron acquisition inside the human host, where the free iron concentration is well below that required for bacterial growth and virulence [10]. Inhibition of siderophore biosynthesis has emerged as a novel strategy for the development of new antibacterial agents for *mycobacterium tuberculosis* and other pathogenic bacteria [11]. In the present study, we used different *In-Silico tools* and techniques for characterization, homology modeling and active site prediction of Salicyl-AMP ligase protein. The first step includes retrieval of Salicyl-AMP ligase protein sequence from UniProt KB database. The physiochemical properties were analyzed by using ProtParam tool and the secondary structure was predicted by using PSI-PRED secondary structure prediction tool. Later the 3D structure was predicted by using SWISS-MODEL server and the model was validated by using PROCHECK method. Then the active site of Salicyl-AMP ligase protein was predicted by using CASTp server. These active sites of the protein can be used for drug designing purpose.

II. METHODS AND MATERIAL

1. Retrieval of Sequence

The protein sequence of Salicyl-AMP ligase from *mycobacterium tuberculosis* (strain ATCC 25618/H37Rv) was retrieved from UniProtKB protein database (www.uniprotkb.org) and saved in FASTA file format along with its accession ID. UniProtKB is a protein database which is freely accessible to the public and it contains the amino acid sequences of proteins [12].

2. Physicochemical Analysis

The physicochemical properties was analysed by using ProtParam analysis tool. The ProtParam tool calculates parameters such as amino acid composition, molecular weight, theoretical pI, instability index, aliphatic index and grand average of hydropathicity (GRAVY) [13].

3. Secondary Structure Prediction

PSIPRED was a simple and accurate secondary structure prediction method, incorporating two feed-forward neural networks which perform an analysis on output obtained from PSI-BLAST (Position Specific Iterated - BLAST). Using a very stringent cross validation method to evaluate the method's performance, PSIPRED 3.2 achieves an average Q3 score of 81.6% [14].

4. Homology Modeling and Model Validation

The Salicyl-AMP ligase protein sequence was used for comparative homology modeling by using SWISS MODEL server. SWISS-MODEL has a fully automated protein structure homology modeling server to make the protein models accessible to all biotechnologist [15]. After modeling, the quality and validation of the model was carried out by PROCHECK method using PDBsum server [16].

5. Active Site Prediction

The active site was predicted by using CASTp (Computed Atlas of Surface Topography of proteins) server. CASTp provides an online resource for locating, delineating and measuring concave surface regions on three-dimensional structures of proteins [17].

III. RESULTS AND DISCUSSION

1. Sequence Retrieval

The protein sequence of Salicyl-AMP ligase was retrieved from UniProt KB database and the sequence was saved in FASTA file format in notepad.

The protein name, organism name, UniProt KB ID and sequence length were shown in following table 1.

Table 1. Sequence Information

Protein Name	Organism Name	UniProt KB ID	Sequence length
Salicyl-AMP ligase	Mycobacterium tuberculosis	P71716	565

2. Physicochemical analysis

The physicochemical properties was analysed by using ProtParam tool and the results were enlisted in table 2. As per the table, the Salicyl-AMP ligase protein is unstable and basic in nature. The total number of positively charged residues (Arg+Lys) was higher than the total number of negatively charged residues (Asp+Glu).

Table 2. Physicochemical properties

Parameter	Value
Number of amino acids	565
Molecular weight	59280.55
Theoretical pi	5.66
Total number of negatively charged residues (Asp+Glu)	62
Total number of positively charged residues (Arg+Lys)	51
Instability Index(∏)	36.06
Aliphatic Index	90.42
Grand average of hydropathicity	0.012

3. Secondary Structure Prediction

The secondary structure was predicated by using PSI-PRED method. The structure elements like alpha helix, extended strands and random coil were enlisted in table 3. As per table, the percentage of coil in Salicyl-AMP ligase protein was higher than the percentage of alpha helix and extended strand.

Table 3. Secondary Structure Prediction

Secondary structure elements	Number of residues	Percentage
Alpha helix	164	29.026
Extended strands	110	19.469
Coils	291	51.504

Fig.1. Secondary structure of Salicyl-AMP ligase protein

4. Homology Modeling and Model Validation

The 3D structure was predicted by using the SWISS Model server. The template name with its identity score, template PDBID and validation scores were enlisted in Table 4. From the table, PDB id 5kei.1. A was selected as template with its 71.99% sequence identity. The model was visualized in RasMol visualization software package as shown in figure 2. The quality and validation of the model was carried out by Ramachandran plot analysis using procheck analysis plot method at PDBsum server. In analysis it was showed that the percentage of favoured region or validation score is 90.8% which was higher than the percentage of additional allowed region 90% to concluding that the model were of reliable and good quality as shown in figure 4. The predicted three dimensional structure of Salicyl-AMP ligase was deposited in Protein Model Database with PMDB id: PM0082681.

Table 4. Homology modeling and model validation

Parameter	Value
% of identity	71.99%
Template PDB ID	5kei.1. A
Template name	2,3-dihydroxybenzoate-AMP ligase
Validation score	90.8%

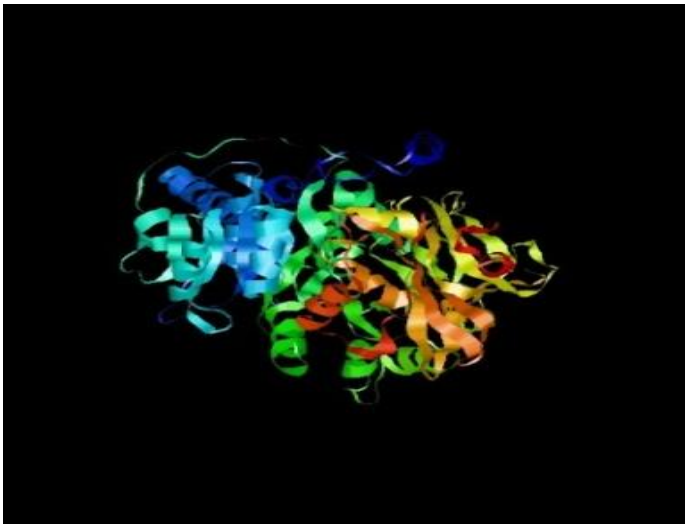

Fig.2. 3D model of Salicyl-AMP ligase protein

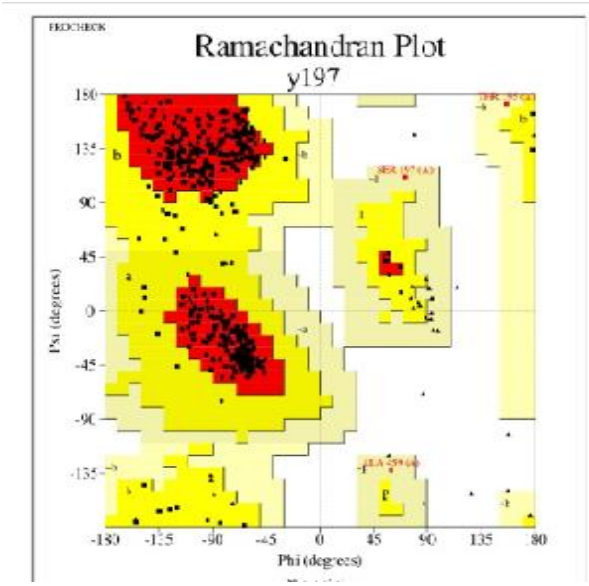

Fig.3. Ramachandran plot analysis

5. Active Site Prediction

The active site was predicted by using CASTp (Computed Atlas of Surface Topography of proteins) server. This server predicted that the Salicyl-AMP ligase protein structure contains 16 binding sites. The area (SA) of the pocket was 33335.666 and volume (SA) 21851.595.

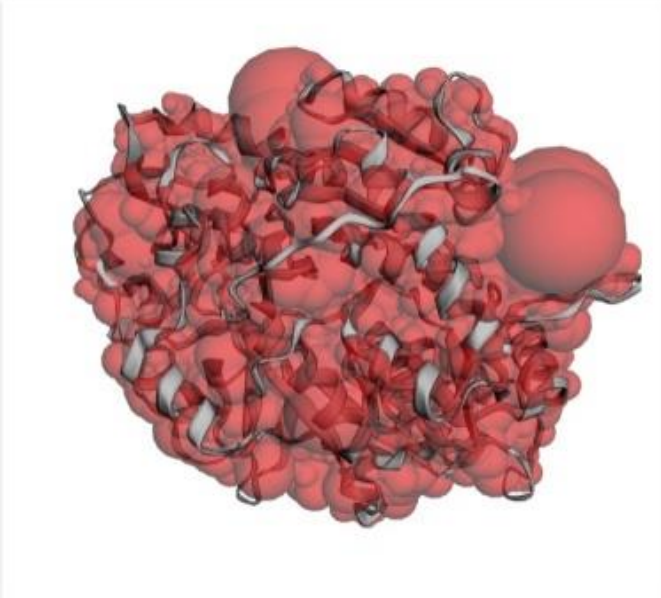

Fig.4. Active site prediction

IV. CONCLUSION

The present study mainly emphasize the prediction three dimensional structure with its of physiochemical and secondary structure of Salicyl-AMP ligase protein using various Bioinformatics tools and techniques. The primary structure indicates that protein consists of 565 amino acids while the physiochemical properties conclude that the protein is unstable and basic in nature. The secondary structure predicts that the protein consist of high percentage of coils as compared to alpha helix and extended strand. The 3D structure was modeled and the quality of the model was evaluated by PROCHECK and Ramachandran plot analysis. The active site of the protein was predicted by using CASTp server and this server predicts that it consists of 16 binding sites. From the present study it has been concluded that the Salicyl-AMP ligase protein can be used as a target for the inhibition of tuberculosis disease by using in- silico drug designing methods.

V. REFERENCES

[1]. WHO. Global tuberculosis report. 2014.

[2]. Global Tuberculosis Report (2014)http://www.who.int/tb/publications/global_report/en/. Last accessed 20 Mar 2015.

[3]. Mehta, P. K., C. H. King, E. H. White, J. J. Murtagh, Jr., and F. D. Quinn. (1996). “Comparison of in vitro models for the study of Mycobacterium tuberculosis invasion and intracellular replication.” Infect. Immun. 64:2673-2679.

[4]. van Zyl Smit RN, Pai M, Yew WW, et al. (2010). "Global lung health: the colliding epidemics of tuberculosis, tobacco smoking, HIV and COPD".

[5]. N. A. Knechel (2009), “Tuberculosis: pathophysiology, clinical features, and

diagnosis," Critical Care Nurse, vol. 29, no. 2, pp. 34-44.

[6]. C. E. Barry III, H. I. Boshoff, V. Dartois et al.,(2009) "The spectrum of latent tuberculosis: rethinking the biology and intervention strategies," Nature Reviews Microbiology, vol. 7, no. 12, pp. 845-855.

[7]. S. H. Cho, D. Goodlett, and S. Franzblau (2006), "ICAT-based comparative proteomic analysis of non-replicating persistent Mycobacterium tuberculosis," Tuberculosis, vol. 86, no. 6, pp. 445-460.

[8]. Vergne AF, Walz AJ, Miller MJ. Iron chelators from mycobacteria (1954-1999) and potential therapeutic applications. Nat. Prod. Rep. 2000; 17:99-116.

[9]. Crosa JH, Walsh CT. Genetics and assembly line enzymology of siderophore biosynthesis in bacteria. Microbiol. Mol. Biol. Rev. 2002; 66:223-249.

[10]. Ferreras JA, Ryu JS, Di Lello F, Tan DS, Quadri LE (2005). Small-molecule inhibition of siderophore biosynthesis in Mycobacterium tuberculosis and Yersinia pestis. Nat. Chem. Biol. 1:29-32.

[11]. Somu RV, Boshoff H, Qiao C, Bennett EM, Barry CE 3rd, Aldrich C.C (2006), Rationally designed nucleoside antibiotics that inhibit siderophore biosynthesis of Mycobacterium tuberculosis. J. Med. Chem.; 49:31-34.

[12]. Magrane M.(2009) And the UniProt consortium. UniProt Knowledgebase: a hub of integrated protein data Database, bar009.

[13]. Gasteiger E, Hooglan C, Gattiker A, Duvaud S, Wilkins MR, Appel RD et al. (2005), The Proteomics Protocols Handbook, Human Press, 571-607.

[14]. Marsden, R.L., McGuffin, L.J. & Jones, D.T. (2002) Rapid protein domain assignment from amino acid sequence using predicted secondary structure. Protein Science. 11, 2814-2824.

[15]. Arnold K, Bordoli L, Kopp J, Schwede T. (2006), The SWISSMODEL Workspace: A web-based environment for protein structure homology modelling. Bioinformatics; 22:195-201.

[16]. Laskowski RA, MacArthur MW, Moss DS, Thornton JM.(1993), PROCHECK - a program to check the stereochemical quality of protein structures. J App. Cryst.; 26:283-29.

[17]. Toh S., Holbrook-Smith D., Stogios P.J., Onopriyenko O., Lumba S., Tsuchiya Y., Savchenko A., McCourt P.(2015), Structure-function analysis identifies highly sensitive strigolactone receptors in Striga. Science.; 350:203-207.

Cite this article as :

Yogesh N. Joshi, Shubham S. Gore, Akshay S. Rajeshwar, "Homology Modeling of Salicyl-AMP ligase involved in Mycobacterium Tuberculosis", International Journal of Scientific Research in Science and Technology (IJSRST), Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 7 Issue 1, pp. 108-113, January-February 2020.
Journal URL : http://ijsrst.com/IJSRST207130