Symbiosis

www.symbiosisonline.org
www.symbiosisonlinepublishing.com
ISSN Online: 2374-8362

Research Article | Journal of Advanced Research in Biotechnology | Open Access

Structural And Functional Annotation Of Rv1514c Gene Of *Mycobacterium Tuberculosis* H_{37}Rv As Glycosyl Transferases

Amjad Beg[1, 2], Shivangi[2], Fareeda Athar[1] and Laxman S Meena[2*]

[1]*Centre for Interdisciplinary Research in Basic Science, Jamia Millia Islamia, Jamia Nagar, New Delhi-110025, INDIA*

[2]*CSIR-Institute of Genomics and Integrative Biology, Mall Road, Delhi-110007, INDIA*

Received: October 16, 2018; Accepted: October 31, 2018; Published: December 20, 2018

***Corresponding author:** Laxman S Meena, CSIR-Institute of Genomics and Integrative Biology, Mall Road, Delhi-110007, INDIA; Telephone no: 0091-11-27002200; Fax No: 0091-11-27667471; E-mail id: meena@igib.res.in*

Abstract

We need an emergent cure against tuberculosis after seeing its capability of scattering all around the world. The causative pathogen *Mycobacterium tuberculosis* (*M. tuberculosis*) persistently gets accomplishment in spreading the repulsiveness of overall expanding death rate; thusly create an emergent requirement for curing people to interact with this pathogen. There might be numerous points of view of this bacterium to take a shot for focusing on the imperative components for its development and survival, one of those systems is recorded in the article underneath. Glycosylation is the universal and one of the most important phenomenon's for modification. This gene is a suitable one for studying exact functioning of glycosyl transferases enzymes as these are the non essential ones and require for growth of this bacterium. The target of the manuscript is to accentuate the *In silico* highlights of the Rv1514c gene of *M. tuberculosis* which prove being an essential role like transferases activity, transferring glycosyl group's action. For the study of Rv1514c gene, it comprises 789 bp long which codes for nearly approx 29kDa protein. In String database server for functional partner prediction demonstrates it interact with an fcl (EpiA), Rv1508A, gmdA, Rv1516c, Rv3264c, Rv1515c, Rv3032, Pima, Rv2188c and glgA. Rv1514c modelled by using I-TASSER server and for model evaluation we using ProSA, Rampage, Verify 3D and ERRAT which outcome score are satisfactory. In Rv1514c binding sites prediction done by COACH which confirmed this gene might be bounded with UDP (Uridine-Diphosphate) and it could be a Transferases activity protein which possibly, transferring glycosyl groups action. Assist exploratory investigations of this gene might be valuable in giving physiological and biological vitality as a remedial.

Keywords: Mycobacterium Tuberculosis; Rv1514c; Glycosylation; Macrophages; Transferases

Abbreviations

Tuberculosis (TB); Mycobacterium tuberculosis (M. tuberculosis); alveolar macrophages (AMs); Multidrug-Resistant TB (MDR-TB); Extremely Drug Resistant TB (XDR-TB); Protein Data Bank (PDB); Iterative Threading ASSEmbly Refinement (I-TASSER); Local Meta threading server (LOMETS); Protein Structure Analysis (ProSA); RAMPAGE (Ramachandran Plot Analysis); Protein Quality Analysis (ProQ); Gene Ontology Annotation (GOA)

Introduction

A standout amongst the most pulverization malady since the eighteenth century is 'Tuberculosis' which was first identified and analyzed by Robert Koch (1843-1910). This researcher and his associates were discovered an unambiguous staining method for the identification of the causative agent of this dangerous malady which is *Mycobacterium tuberculosis* (*M. tuberculosis*). This staining strategy depended on the mix of methylene blue color pursued by counterstaining with vesuvin (dark colored dye); this darker color dye distinguishes *M. tuberculosis* not only in culture but also in the tissues [1]. The strategy attained so far for identifying and diagnosing this pathogen are not working nowadays, one like BCG vaccination, which was one of the best drug of that time to cure this malady completely, but now its fails due to frequent mutations in the genome of this pathogen and this mechanism provoke easy escape from the immune system of host cell [2-4]. At the beginning form the eighteenth century, the researcher had been endeavoring to outline the medication that can give hundred percent treatment from this infection but yet unsuccessful due to non-consistent nature of this bacterium [5]. The unstable environment of this bacterium provides the capability to this bacterium to escape effectively from the immunological obstructions of the human body (innate immune system) [6-8]. The most distinctive antigen presenting cells known as dendritic cells renowned for their role in the killing of the parasitic intracellular or extracellular bacterium, however for the situation particularly of *M. tuberculosis*, these cells flop in their obligation. *M. tuberculosis* takes another approach to get away from the killing system of invulnerability yet precisely how [9-10]. M. tuberculosis bacterium after the entering into an individual person through the nasal track it fixes itself in the host alveolar macrophages where it can continue for a more extended timeframe with no deterrent from host-specific insusceptible cells [11-12]. *M. tuberculosis* alters itself to remain in the macrophage by achieving some exceptional highlights like the adjustment in its cell wall component, communicating qualities that can oppose acidic condition, actuation of qualities that upgrade the resistance property of this bacterium and so forth [13].

This manuscript elaborates the significance of Rv1514c of *M. tuberculosis* which is predicted to be glycosyl transferases. Although Glycosylation is a universal, important phenomenon in prokaryotes and eukaryotes, most of its features are still unknown and poorly understood [14]. Mainly these proteins are basically membrane linked, unstable and present in very low concentration [15-17]. These enzymes help in synthesis of many cell wall components like peptidoglycan, lipopolysacharide, glycolipids etc. In case of *M. tuberculosis*, many drugs had been synthesized in order to target cell wall components but the exact mechanism behind this is still unknown [18]. This gene is a suitable one for studying exact functioning of glycosyl transferases enzymes as these are the non essential ones and require for growth of this bacterium. Earlier studies of *M. tuberculosis* found many glycosyl transferases genes that play significant roles in the cellular processes like mshA, gal transferases etc [19]. Therefore understanding this gene may provide us with better way to understand the mechanism of cell wall synthesis in M. tuberculosis and its effect on pathogenesis and thus help us to find a mode towards eradication of this disease [20-22].

Material and Methods

Prediction of the Interacting Functional Partner

Analysis of the protein-protein interaction is important for the protein significance, because the interacting partner may be function as like our interest protein. STRING database web server analyzes the interacting score for the prediction of functional partner. It predicts that cytoplasmic protein interacts with some other protein and it works in a web-like manner. In the STRING database server there are low, medium and high cutoff score where the score values is for <0.4; medium: 0.4 to 0.7; high: >0.7 [23, 24].

Ab initio Protein Modelling of Rv1514c

Ab initio which means (From the beginning) in protein modelling, for the protein modelling we required FASTA format sequence which is retrieved from the Mycobrowser database (https://mycobrowser.epfl.ch/gene/) which has the comprehensive data of *Mycobacterium tuberculosis* for genomics and proteomics studies [25]. The Protein Data Bank (PDB) structure of the Rv1514c was not found in PDB (http://www.rcsb.org/pdb/home/home.do), so we have started the analysis of structure modelling of Rv1514c. For the analysis of secondary structure of the protein and protein structure property prediction we are using RaptorX web server [26].

I-TASSER (Iterative Threading ASSEmbly Refinement) server is an online server for forecast of the structural modelling of the protein. For the protein demonstrating by I-TASSER (https://zhanglab.ccmb.med.umich.edu/I-TASSER/) we should have the FASTA arrangement of the protein sequence [27]. This server modelled the 3D structure of the protein by following the three phases though, Local Meta threading server (LOMETS) for enhancing the model of the optional structure of a protein by covertly presented it utilizes the alpha-helix, beta-sheet and loop individually. In I-TASSER server, structure examined by the C-score, z-score and inclusion of threading arrangement where the cutoff estimation of the C-score there are - 5 to 2, Z-score >1 and COV (Coverage of the amino acid in modelling) which implies inclusion of alignment greater than 70%. Evaluation of the, a TM-score ≤ 0.17 compares to a likeness between two arbitrarily chose structures from the PDB library; a TM-score > 0.5 relates roughly to two structures of the comparative topology [28-30].

Evaluation of Rv1514c Protein Model

The evaluation of the structural modeled protein of the protein is done by Protein Structure Analysis (ProSA) and SAVES metaserver. The validation of the protein was performed by the structure analysis and verification server ProSA (https://prosa.services.came.sbg.ac.at/prosa.php) analysis [31] and performed to the exactness and dependability of the displayed structure modeled protein. SAVES (http://nihserver.mbi.ucla.edu/SAVS/) was utilized to complete the confirmations of the model with ERRAT and Verify3D. The general characteristics of the demonstrated structures were assessed utilizing ERRAT [32]. Verify3D was utilized to approve the refined structure. The 3D structure of the protein was contrasted with its very own amino-corrosive grouping thinking about a 3D profile ascertained from the nuclear directions of the structures of right proteins [33]. The built model was assessed for its spine compliance utilizing a Ramachandran plot. The stereo chemical nature of the demonstrated proteins is surveyed from Ramachandran validate score for favored and unfavored regions [34, 35].

Intrinsic Dynamics Studies of the Rv1514c Model

For the understanding of structural dynamics of proteins is a primary need for increasing more prominent bits of knowledge into their biological functions [36]. The modeled protein structure dynamics analysis on the basic elements were performed by the WEBnm@ server (http://www.bioinfo.no/instruments/normalmodes) [37] to figure the slowest modes and related misshaping energies to compute typical mode investigation of the proteins adding to the comparing protein development. Simple mode investigation figures the likely developments of the proteins and is the strategy for determination for investigating the slowest movement of decision [38]. The Structural modeled protein representation visualization was performed by RasMol.

Rv1514c model quality analysis

For the forecast of quality of a protein by ProQ (http://www.sbc.su.se/~bjornw/ProQ/ProQ.html) online server had been used which depends on the neural system constructed apparatus which is based on the evaluation of the structural characters, there is the quality of a protein model and it is streamlined to discover revise models to discover local structures. The quality estimates the LG score and MaxSub. The cutoff scope of LG score> 1.5 is extremely great model, > 2.5 great model and > 4 to a great degree great model and there MaxSub score> 0.1 extremely great model, > 0.5 great model and > 0.8 amazingly great model [39, 40].

Protein Binding Site Prediction of Rv1514c protein

The binding sites forecast depends on the protein arrangement or likewise done by PDB file (3D structure) in binding sites prediction is done by the online server. COACH Metaserver (https://zhanglab.ccmb.med.umich.edu/COACH/) used for the ligand binding site forecast or ligand pocket prediction. For the investigation of the ligand binding prediction starts from the PDB structure or else pick document in PDB format of the target protein. After that this server will make correlative ligand restricting site expectation by using the two relative systems: TM-site and S-site which recognize the presence of the ligand-restricting site, it sorts out the specific restricting substructure and profile gathering structure the (BioLiP) database [41]. COACH server parameters are C-score, RMSD, Cov, BS-score Lig name and anticipated restricting buildup. In COACH server the C-score is confidence score, RMSD is deposits that are basically adjusted by TM-adjust, Cov is the coverage of global structural alignment, BS-score>1 proportion of neighborhood closeness (sequence & structure) between layout the template binding site and predicted binding site and the cutoff score BS-score >1 reflect a significant local match between the predicting and template binding site and Lig name are the ligand name [42].

Sequence-based functional analysis of Rv1514c protein

The useful investigation of protein is the key factor in its examination and it will be finished by the COFACTOR online server (https://zhanglab.ccmb.med.umich.edu/COFACTOR/) which predicts the function of the protein [43]. COFACTOR is model based server utilized for protein-protein interrelation and a natural clarification of protein molecules. Examination of this protein functioning for threading was done by BioLiP protein work database. By this database, investigation of the model structure figures out how to observe the homology and the functional state of the protein. COFACTOR model-based capacity expectation estimation was situated as the best methodology for protein determine protein structure configuration [44].

Results

Prediction of the Interacting Functional Partner

Interacting partner functional prediction analysis has done by the of STRING database server. The outcome demonstrates that Rv1514c protein interconnected with 10 number of protein like fcl (EpiA), Rv1508A, gmdA, Rv1516c, Rv3264c, Rv1515c, Rv3032, pimA, Rv2188c and glgA [45-51] are appeared in Figure 1. STRING database server produces the score based on associating accomplice which demonstrates the cutoff esteem in the middle of the base communication score (between 0.4 - 0.6) or more score indicated high interaction. Based on the connecting score, fcl (EpiA) (Nucleotide-sugar epimerase) protein demonstrates the high associating score i.e. 0.923 any other predicted functional partner score are shown in Table 1.

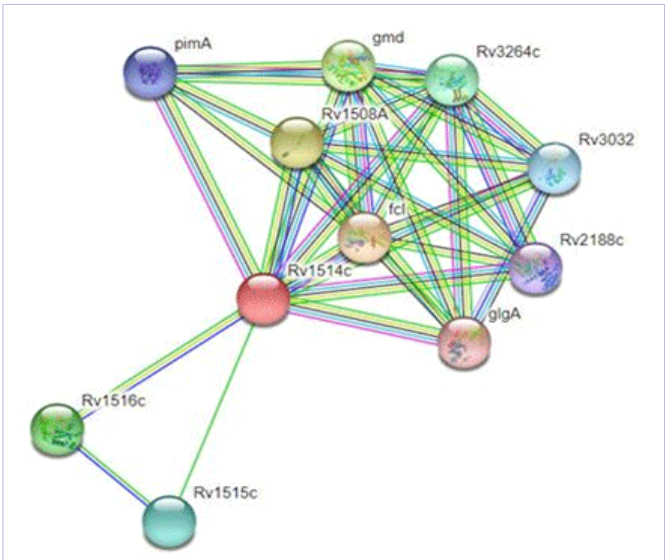

Figure 1:
Protein-interacting partners: In Rv1514c protein the interacting functional partner prediction done by STRING database server result shows the Rv1514c protein interact with Rv1508A, *gmdA*, Rv1516c, Rv3264c, Rv1515c, Rv3032, *pimA*, Rv2188c and glgA. In this figure the *fcl* (*EpiA*) protein is highly interact to the Rv1514c which score is 0.923.

Table 1:
Interacting functional partners: This table enlists the interacting functional partner of Rv1514c protein of *M. tuberculosis* interconnected with 10 number of proteins like *fcl (EpiA),* Rv1508A, *gmdA*, Rv1516c, Rv3264c, Rv1515c, Rv3032, *pimA*, Rv2188c and *glgA* with that the cutoff score.

No.	Functional Partner	Function	Score	Reference
1	*fcl (EpiA)*	Nucleotide-sugar epimerase	0.923	[45]
2	Rv1508A	Hypothetical protein	0.884	[46]
3	*gmdA*	GDP-D-mannose dehydratase;	0.819	[45]
4	Rv1516c	Sugar transferase	0.785	[46]
5	Rv3264c	D-alpha-D-mannose-1-phosphate guanylyltransferase	0.738	[47]
6	Rv1515c	Hypothetical protein	0.738	[48]
7	Rv3032	Transferase	0.695	[49]
8	*pimA*	Alpha-mannosyltransferase	0.695	[50]
9	Rv2188c	Hypothetical protein	0.695	[51]
10	*glgA*	Capsular glucan synthase	0.695	[49]

Ab initio Protein Modelling of Rv1514c

As we probably know for the protein modelling first we retrieved the Fasta format protein sequence of the Rv1514c gene by using the Mycobrowser database which has nucleotide and protein sequences of the M. tuberculosis species and physical-chemical properties of the genes. This quality has 789 base pair long quality and the protein sub-atomic mass ~ 29kDa protein. This protein appears as a speculative protein and not studies previously. For the analysis of protein structure property prediction using RaptorX (http://raptorx.uchicago.edu/) web server of Rv1514c, there are 33% α-Helix, 18% β-sheet, 48% Coil region and solvent accessibility 29%Exposed, 35% Medium and 35% Bury present. I-TASSER is an online server which modeled the 3D structure of the protein by utilizing amino corrosive arrangement and this demonstrating is known as ab initio modeling. It is figured by the threading layout formats arrangement. The structure of Rv1514c protein quality of structure modelled relies on the estimation of C-score and RMSD esteem. The level of the positive district lies over 90% and the C-score is the certainty score for each model which is from -5 to 2 where the higher conviction demonstrates is controlled by the higher estimation of C-score. By the I-TASSER demonstrated organized there is 5 protein models made by the C-score and situated by gathering measure. The demonstrated protein structure which chose on the 1st rank premise its PDB hit is 2Z86 (Crystal structure of chondroitin polymerase from Escherichia coli strain K4 (K4CP) complexed with UDP-GlcUA and UDP). The template alignment PBD hit 2Z86 is Transferases from organism *Escherichia coli.* The demonstrated protein PDB hit template identity is 17-20 percent and the Normalized z-score is 2.97 which indicate superior alignment and the quantity of adjusted buildups which appears by (Cov) is 0.85 which implies the 85% of the coverage of the threading alignment, the model protein of Rv1514c confidence score is -0.47 the estimated TM-score is 0.65±0.13 and Estimated RMSD = 6.9±4.1Å. Finally, the displayed structure of the Rv1514c protein is appeared in Figure 2.

Figure 2:
Protein modelling: Rv1514c structure modelled by using I-TASSER, the PDB hit of a template alignment is 2Z86, Normalized z-score is 2.97, Cov is 0.85.The C-score is -0.47 the estimated TM-score is 0.65±0.13 and the Estimated RMSD = 6.9±4.1Å. I-TASSER server model protein Rv1514c outcome score are satisfactory.

Evaluation of Rv1514c Protein Model

ProSA: Model evaluation essential for the demonstrated protein structure, if there should be an occurrence of ab initio displaying this, is provided with the measurable unwavering quality to the 3D structure of the demonstrated protein. For the model investigation Protein Structure Analysis (ProSA) server check the potential errors. The assessed z-score of the model protein is -6.31 which is appeared in Figure 3(a). ProSA server demonstrates the second plot, which indicates plotting energies as a function of amino acid arrangement position i. As a rule, positive qualities compare to problematic or incorrect parts of the input structure. A plot of single buildup energies, as a rule, contains large variances and is of limited an incentive for model assessment. Thus the plot is smoothed by ascertaining the normal vitality over every 40-buildup part s (i, i+39), which is then allocated to the 'focal' deposit of the piece at position i+19. The second line with a little window size of 10 deposits is appeared out of sight of the plotted Figure 3(b).

Rampage: The protein model evaluation by SAVES metaserver some server like (RAMPAGE, Verify3D and ERRAT). RAMPAGE (Ramachandran plot examination) which is an online server look at the displayed protein structure showed that the 92.3% residues are in the most favored region, 5.8% buildups in the additional allowed region and 1.9% residues in outlier region. These parameters of protein structure exhibiting that our showed protein was of good quality steady and satisfactory.

Verify3D: The assessment of Rv1514c protein three-dimensional profiles investigation by utilizing a Verify3D server. This program assesses the closeness of an atomic model (3D) with its own specific amino corrosive arrangement which is 1 dimensional. Each deposit is doled out an essential class in a brilliance of its zone and condition (alpha, beta, circle, polar, non-polar et cetera). The score ranges from-1 (poor score) to +1 (awesome score). 90.48% of the arrangement had found in the centre estimation of 3D-1D score >=0.2 that is discerning for our exhibited protein appeared.

Errat: ERRAT is an online server which favors the protein structure on the beginning of the atomic association between different sorts of molecules. The general quality factor is 98.8649 of our protein structure which is satisfactory.

Intrinsic Dynamics Studies of the Rv1514c Model

In typical mode examination (NMA) initial six modes coordinating with worldwide pivot and interpretation of the framework are for the most part disregarded and consequently, least recurrence mode of concern is the seventh one. Typical Mode Analysis of the Rv1514c protein showed that low twisting energies were related with generally unbending districts in the

protein. NMA demonstrated the vibration and thermal properties of a protein at the atomic level. The Rv1514c protein from M. tuberculosis had minimal deformation energies are 1541.04, 2992.26, 4095.30 and 5269.36 but the lowest deformation energies of 2992.26 in the seventh mode as shown in Table 2. It inferred that the seventh mode with extensive unbending regions had a superior probability of describing domain motions as shown in Figure 4.

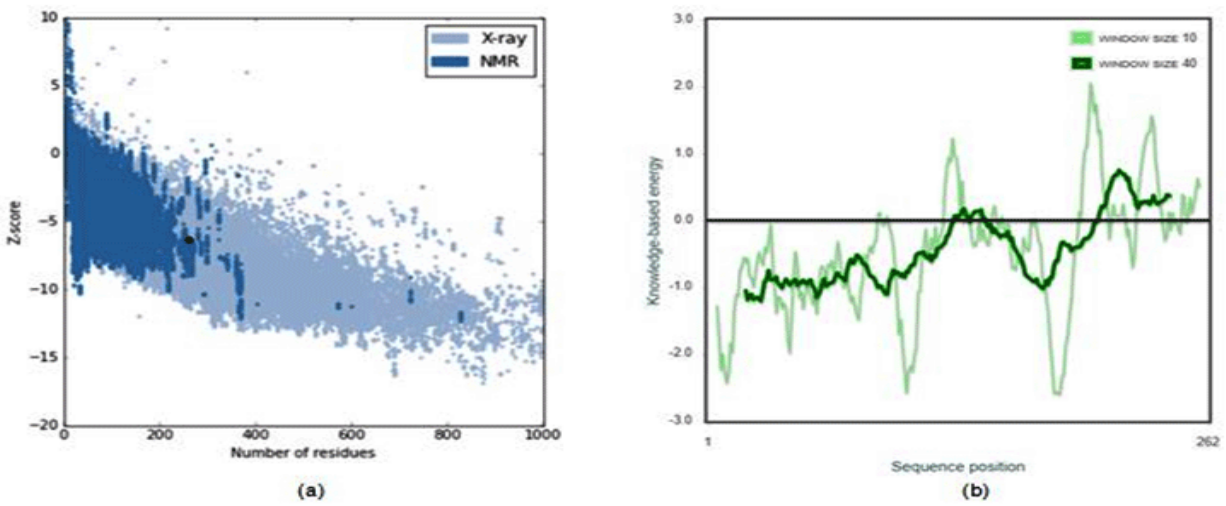

Figure 3:
Model Evaluation: For the model evaluation done by ProSA tool result showing the subfigures indicate that Rv1514c protein that's although homologous of (PDB code 2Z86, chain A). (a) ProSA-web z-scores of Rv1514c protein chains (A) in PDB determined by X-ray crystallography (light blue) or NMR spectroscopy (dark blue) with acclaim to their length (-6.31) shown in above studies and the z-scores of Rv1514c point as large dots. (b) The energy plot of Rv1514c shows the energies of the residue the window size is 40.

Table 2:

Deformation energies: This table shows the minimal deformation energies of Rv1514c protein from *M. tuberculosis* had deformation energies are 1541.04, 2992.26, 4095.30 and 5269.36 but the lowest deformation energies of 2992.26 in the seventh mode.

S. No.	Mode Index	Deformation Energies
1	7	616.28
2	8	1249.42
3	9	1591.43
4	10	1541.04
5	11	2055.35
6	12	2458.86
7	13	3046.28
8	14	2992.26
10	15	4127.75
11	16	4095.30
12	17	5176.75
13	18	5565.31
14	19	5269.36
15	20	6734.49

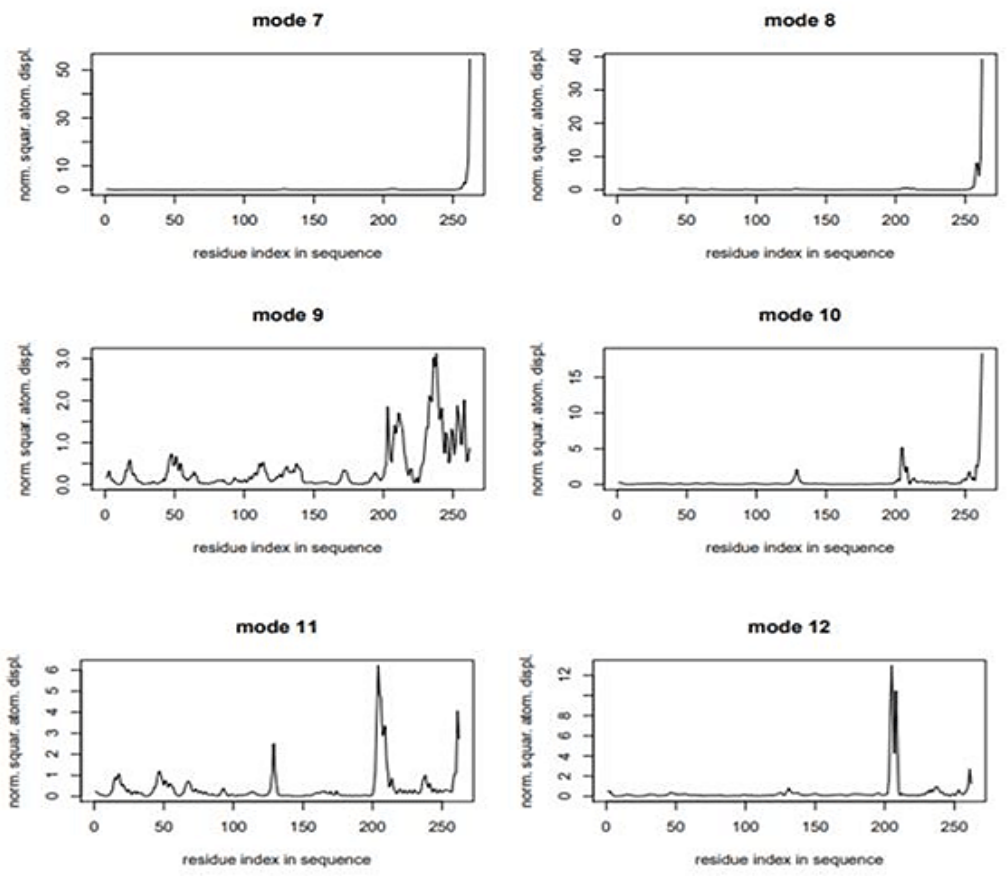

Figure 4:
Prediction of Intrinsic Dynamics analysis: The intrinsic dynamics prediction figure shows the Deformation energies and Normalized atomic displacement plot of the Rv1514c protein for *M. tuberculosis* $H_{37}Rv$.

Rv1514c model quality analysis

Protein quality analysis (ProQ) predicts the protein quality, it depends on the neural system which can figure the number of structural characters this demonstrates the quality of a protein model and its improved to assess the right models and discover native structures. ProQ assessment estimates LG score and MaxSub. In Rv1514c demonstrate structure protein, anticipated LG score was 4.060 and MaxSub 0.146 which implies Rv1514c protein quality is an extremely great model.

Ligand Binding Site Prediction of Rv1514c protein

For protein binding site prediction of Rv1514c protein done by COACH server they anticipated ligand binding (pocket) site which produce corresponding binding site forecasts utilizing two similar strategies, TM-SITE and S-SITE. By the examination of COACH, results demonstrate on rank 1 Protein Data Bank (PDB) hit 2D7I which have C-score 0.91 where the cluster size is 109 and it is a (Crystal structure of pp-GalNAc-T10 with UDP, GalNAc and Mn^{2+}) Transferase protein form the organism of *Homo sapiens* where ligand as UDP (Uridine-Diphosphate). In Rv1514c buildup consensus binding sites residues as 12, 13, 14, 16, 43, 68, 70, 71, 91, 92, 93, 173, 174, 200 and 204. In TM-SITE and S-SITE result rank 1 the PDB hit is also 2D7I but the C-score and cluster size is low with comparison earlier result but the other hand in rank 2 the PDB hit is 4FIY which C-score is 0.17 and cluster size is 3 ligand name are manganese(2+) binding sites are 16,93,94,203. At last, the server demonstrate that In Rv1514c protein binding sites predicted on the 12, 13, 14, 16, 43, 68, 70, 71, 91, 92, 93, 173, 174, 200 and 204 residues with Uridine-Diphosphate clearly shown in Figure 5.

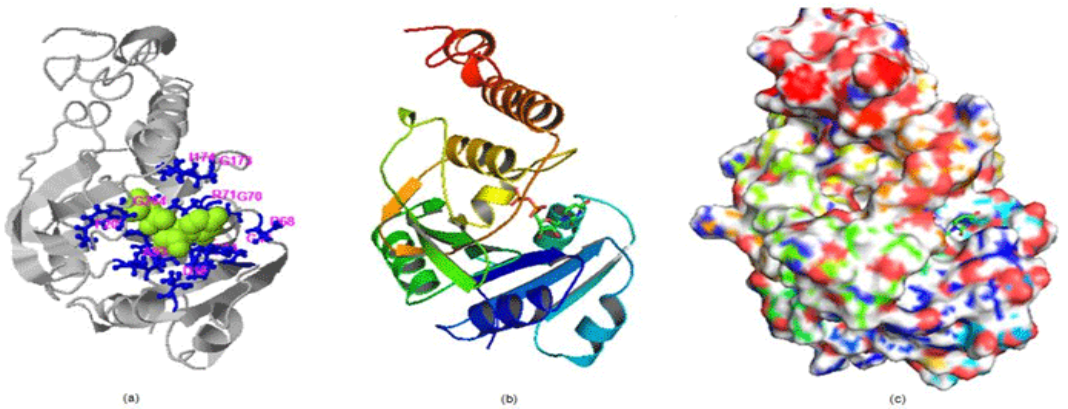

Figure 5:
Ligand Binding Site Prediction: For the ligand binding sites prediction of the Rv1514c protein has been done by Coach server demonstrates that Uridine-Diphosphate (UDP) as ligand with C-score 0.91 where the cluster size is 109 and the binding sites residues as 12, 13, 14, 16, 43, 68, 70, 71, 91, 92, 93, 173, 174, 200 and 204. **(a)** This figure shows the binding residues on the site of Rv1514c protein. **(b)** In this cartoon model of Rv1514c protein showing the ligand Uridine-Diphosphate (UDP). **(c)** At last, in this figure showing the Rv1514c surface model with covered the ligand in the protein structure

Functional analysis of Rv1514c protein

Rv1514c protein display based function forecast studies had been finished by COFACTOR servers. COFACTOR is a structured plan and protein-protein affiliation based strategy for common perception of protein particles. COFACTOR results anticipated analogue in PDB, Molecular limit, biological process, a cellular component, enzyme homolog in PDB, format protein with relative restricting locales appeared in Figure 6. COFACTOR online server results in the foreseen quality metaphysics Gene Ontology (GO) terms are organized by atomic capacity, organic process, and cell part with a distinct C-score. In a Rv1514c useful examination of atomic capacity by COFACTOR are with score 1.0 they anticipate synergist action, confidence score 1.00 is for transferases activity which means transferring glycosyl groups action, which means 100% confidence score for the function prediction of a Rv1514c as a Transferases protein which transferring glycosyl groups action. The outcome result of the COFACTOR results shown in Figure 6.

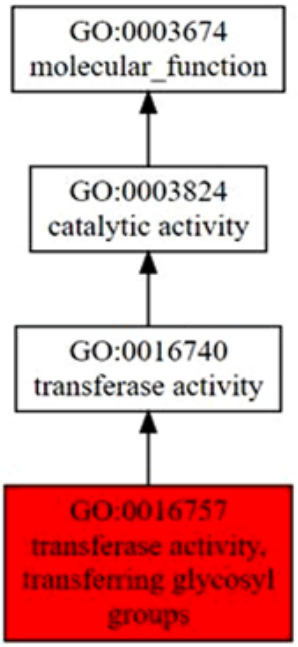

Figure 6:
Functional analysis: The figure demonstrates the gene ontology studies with the molecular Function and Cell segment. The entire procedure exhibits the molecular function studied C-score is 1.00 for Transferases action has appeared in figure 6 and the cellular component C-score is 0.95 for an intracellular part.

Discussion

In the wake of seeing the state of the world because of TB we presently require an eminent and regular treatment to secure or to fight against this disease. Advancement of different lines drugs additionally can't fix which influences the circumstances to compound and furthermore these medications make an opposing situation because of them or by comparative medications. After all the consistent endeavors by the researcher from numerous decades to make a lasting arrangement, there was an absence of the entire solution for this infection. In this manner in the current situation influences scientists to decide its total fix as a vital activity. The mentioned manuscript provides us with a new target which is glycosyl transferases to kill this pathogenic bacterium. Glycosyl transferases are the universal molecule help in transfer of the glycosyl residue on various moieties [14]. Various bioinformatics approaches have been applied on this gene to elaborate its various functions and essentiality in biological system. Glycosyl transferases are badly required for synthesizing glycosyl linkages in the cell wall components like arabinan LAM, Mycolic acids etc [19]. Rv1514c is predicted to be a glycosyl transferases gene and

various bioinformatics approaches have been used to understand the working. STRING database tool proves protein-interaction with various other proteins but the interaction capacity was higher for fcl (EpiA) which is nucleotide sugar epimerase [23, 24]. As fcl (EpiA) involved in the cell make up, its interacting partner Rv1514c might be an important key point for its functioning. The 3D structure for this protein has been modeled by I-TASSER [27] evaluated by ProSA, ERRAT, Verify3D and RAMPAGE [32-35] and their intrinsic dynamics studies by WEBnm@ which figure out the miss happening energies [37]. The model quality analysis predicted by ProQ which outcome result is Rv1514c model quality is satisfied [39, 40]. Binding sites on this protein have been derived by the help of COACH server which ensure the binding of UDP and manganese2+ on various sites [41, 42]. Functional analysis of this gene is forced by COFACTOR server which proves that this gene having transferases activity by a satisfied score [43, 44]. As discussed above that glycosyl transferases is one of the important class of enzymes, understanding and classification of these gene might be an important step in knowing the strategy of the survival of this bacterium inside host cell and therefore could be used a step in the eradication of this disease.

Acknowledgement

The authors thankful to Council of Scientific & Industrial Research-Institute of Genomics and Integrative Biology for fiscal support. Md. Amjad Beg also acknowledges University Grants Commission Maulana Azad National Fellowship for the financial support and Jamia Millia Islamia University.

References

1. Sakula A. Robert Koch: The story of his discoveries in tuberculosis. Irish Journal of Medical Science. 1985;154(1):3-9. doi: 10.1007/BF02938285
2. Martin C. The dream of a vaccine against tuberculosis; new vaccines improving or replacing BCG? European Respiratory Journal. 2005; 26(1):162-167. doi: 10.1183/09031936.05.00109904
3. Rodrigues LC, Mangtani P, Abubakar I. How does the level of BCG vaccine protection against tuberculosis fall over time? British Medical Journal. 2011;343-d5974. doi: 10.1136/bmj.d5974
4. Caws M, Duy PM, Tho DQ, Lan NT, Hoa DV, Farrar J. Mutations prevalent among rifampin and isoniazid-resistant Mycobacterium tuberculosis isolates from a hospital in Vietnam. Journal of Clinical Microbiology. 2006;44(7):2333-2337. doi: 10.1128/JCM.00330-06
5. Russell DG. *Mycobacterium tuberculosis*: here today, and here tomorrow. Nature Reviews Molecular Cell Biology. 2001;2(8): 569-577. doi: 10.1038/35085034
6. Barry CE, Boshoff HI, Dartois V, Dick T. The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. Nature Reviews Microbiology. 2009;7(12):845-855. doi: 10.1038/nrmicro2236
7. Pai M, Behr MA, Dowdy D, Dheda K, Divangahi M, Boehme CC. Tuberculosis. Nature Reviews Disease Primers. 2016;27(2): 16076. doi: 10.1038/nrdp.2016.76
8. Esmail H, Barry CE, Young DB, Wilkinson RJ. The ongoing challenge of latent tuberculosis. Philosophical Transactions of the Royal Society Biological science. 2014;369(1645):20130437. doi: 10.1098/rstb.2013.0437
9. Pieters J. *Mycobacterium tuberculosis* and the macrophage: maintaining a balance. Cell Host & Microbe. 2008;3(6):399-407. doi: 10.1016/j.chom.2008.05.006
10. Yerlikaya S, Broger T, MacLean E, Pai M, Denkinger CM. A Tuberculosis Biomarker Database: The Key to Novel TB Diagnostics. International Journal of Infectious Diseases. 2017;(56):253–257. doi: 10.1016/j.ijid.2017.01.025
11. Meena LS. An overview to understand the role of PE_PGRS family proteins in *Mycobacterium tuberculosis* H37Rv and their potential targets. Biotechnology Applied Biochemistry. 2015;62(2):145-153. doi: 10.1002/bab.1266
12. Kumari P, Meena LS. Factors affecting susceptibility to *Mycobacterium tuberculosis*: a close view of immunological defence mechanism. Applied Biochemistry and Biotechnology. 2014;174(8):2663-2673. doi: 10.1007/s12010-014-1217-3
13. Rajni, Meena LS. Survival mechanisms of pathogenic *Mycobacterium tuberculosis* H37Rv. FEBS Journal. 2010;277(11):2416-2427. doi: 10.1111/j.1742-4658.2010.07666.x
14. Unligil UM, Rini JM. Glycosyltransferase structure and mechanism. Current Opinion Structural Biology. 2000;10(5):510-517. doi: 10.1016/S0959-440X(00)00124-X
15. Liu J, Mushegian A. Three monophyletic superfamilies account for the majority of the known glycosyltransferases. Protein Science. 2003;12(7):1418-1431. doi: 10.1110/ps.0302103
16. VanderVen BC, Harder JD, CrickDC, Belisle JT. Export-mediated assembly of Mycobacterial lycoproteins parallels eukaryotic pathways. Science. 2005;309(5736):941-943. doi: 10.1126/science.1114347
17. Alderwick LJ, Seidel M, Sahm H, Besra GS, Eggeling L. Identification of a novel arabinofuranosyltransferase (AftA) involved in cell wall arabinan biosynthesis in *Mycobacterium tuberculosis*. Journal of Biological Chemistry. 2006;281(23):15653-15661. doi: 10.1074/jbc.M600045200
18. Kaur D, Berg S, Dinadayala P, Gicquel B, Chatterjee D, McNeil MR, Vissa VD, et al. Biosynthesis of mycobacterial lipoarabinomannan: role of a branching mannosyltransferase. Proceedings of the National Academy of Science of the United States of America. 2006;103(37):13664-13669. doi: 10.1073/pnas.0603049103
19. Buchmeiner N, Fahey RC. The mshA gene coding the glycosyltransferase of mycothiol biosynthesis is essential in *Mycobacterium tuberculosis* Erdman. FEMS Microbiology Letter. 2006;264(1):74-79. Doi 10.1111/j.1574-6968.2006.00441.x
20. WatkinsHA, BakerEN. Structural and functional analysis of Rv3214 from Mycobacterium tuberculosis,a protein with conflicting functionalannotations, leads toits characterization as a phosphatase. Journal of Bacteriology. 2006;188(10):3589-3599.
21. Monu, Meena LS. Biochemical characterization of PE_PGRS61 family protein of *Mycobacterium tuberculosis* H37Rv reveals the binding ability to fibronectin. Iranian Journal Basic Medical Science. 2016;19(10):1105-1113.

22. Beg MA, Shivangi, Thakur SC, Meena LS. Structural prediction and mutational analysis of Rv3906c gene of *Mycobacterium tuberculosis* H37Rv to determine its essentiality in survival. Advances in Bioinformatics. 2018;6152014:1-12. doi: 10.1155/2018/6152014
23. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Alberto Santos, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Research. 2017;45(D1):D362-D368. doi: 10.1093/nar/gkw937
24. Lewis AC, Saeed R, Deane CM. Predicting protein-protein interactions in the context of protein evolution. Molecular Biosystem. 2010;6(1):55-64. doi: 10.1039/b916371a
25. Kapopoulou A, Lew JM, Cole ST. The MycoBrowser portal: a comprehensive and manually annotated resource for mycobacterial genomes. Tuberculosis. 2011;91(1):8-13. doi: 10.1016/j.tube.2010.09.006
26. Gao Y, Wang S, Deng M, Xu J. RaptorX-Angle: real-value prediction of protein backbone dihedral angles through a hybrid method of clustering and deep learning. BMC Bioinformatics. 2018;19(Suppl 4):100. doi: 10.1186/s12859-018-2065-x
27. Kumar S, Nath O, Govil S, Pathak AN. Computational 3D prediction, evaluation and analysis of pyruvate dehtdrogenase an effective target for filarial infection by Brugia pahangi using homology modeling approach. International Journal of Pharmaceutical Science Drug Research. 2014;6(2):120-123.
28. Yang J, Yan R, Roy A, Xu D, Poissn J, Zhang Y. The I-TASSER Suite: protein structure and function prediction prediction. Nature Methods. 2015;12(1):7-8. doi: 10.1038/nmeth.3213
29. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein and function prediction. Nature Protocol. 2010;5(4):725-738. doi: 10.1038/nprot.2010.5
30. Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics. 2008;9:40. doi: 10.1186/1471-2105-9-40
31. Wallner B, Elofsson A. Can correct protein models be identified. Protein Science. 2003; 12(5):1073-86. doi: 10.1110/ps.0236803
32. Wiederstein M, Sippl M J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Research. 2007. doi: 10.1093/nar/gkm290
33. Colovos C, Yeates T O. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Science. 1993; 2(9): 1511-1519. doi: 10.1002/pro.5560020916
34. Bowie J U, Luthy R, Eisenberg D. A method to identify protein sequences that fold into a known three-dimensional structure. Science. 1991; 253(5016): 164-170. doi: 10.1126/science.1853201
35. Ho B K, Brasseur R. The Ramachandran plots of glycine and pre-proline. BMC Structural Biology. 2005. doi: 10.1186/1472-6807-5-14
36. Lovell S C, Davis I W, Arendall W B, Bakker P I, Word J M, Prisant M G, et al. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. Protein. 2003; 50(3): 437-450. doi: 10.1002/prot.10286
37. Hollup S M, Salensminde G, Reuter N. Webnm@ a web application for normal mode analyses of proteins. BMC Bioinformatics. 2005. doi: 10.1186/1471-2105-6-52
38. Sharon F B, Daniel R R. Homology Modeling of Nitrogenase Iron Protein of Nitrogen Fixing Actinomycete Arthrobacter sp. International Journal of Computer Applications. 2013; 61(1):13-19. doi: 10.5120/9891-4457
39. Kumar S, Nath O, Govil S, Pathak A N. Computational 3D Structure Prediction, Evaluation and Analysis of Pyruvate Dehydrogenase an Effective Target for Filarial Infection by Brugia pahangi Using Homology Modeling Approach. International Journal of Pharmaceutical Sciences and Drug Research. 2014; 6(2): 120-123.
40. Cristobal S, Zemla A, Fisher D, Rychlewski L, Elofsson A. A study of quality measures for protein threading models. BMC Bioinformatics. 2001. doi: 10.1186/1471-2105-2-5
41. Yang J, Roy A, Zhang Y. Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. Bioinformatics. 2013; 29(20): 2588-2595. doi: 10.1093/bioinformatics/btt447
42. Wang S, Sun S, Li Z, Zhang R, Xu J. Accurate De Novo Prediction of Protein Contact Map by Ultra-Deep Learning Model. PLoS Computational Biology. 2017; 13(1). doi.org/10.1371/journal.pcbi.1005324
43. Roy A, Yang J, Zhang Y. COFACTOR: an accurate comparative algorithm for structure-based protein function annotation. Nucleic Acids Research. 2012; 40 (Web Server issue): W471-477. doi: 10.1093/nar/gks372
44. Zhang C, Freddolino P L, Zhang Y. COFACTOR: improved protein function prediction by combining structure, sequence and protein-protein interaction information. Nucleic Acids Research. 2017; 45(W1): W291-W299. doi: 10.1093/nar/gkx366
45. Mattow J, Jungblut P R, Schaible U E, Mollenkopf H J, Lamer S, Zimny Arndt U, et al. Identification of proteins from Mycobacterium tuberculosis missing in attenuated Mycobacterium bovis BCG strains. Electrophoresis. 2001;22(14):2936-46. doi:10.1002/1522-683(200108)22:14<2936::AID-ELPS2936>3.0.CO;2-S
46. DeJesus M A, Gerrick E R, Xu W, Park S W, Long J E, Boutte C C, et al. Comprehensive Essentiality Analysis of the Mycobacterium tuberculosis Genome via Saturating Transposon Mutagenesis. MBio. 2017; 8(1): 2133-16. doi: 10.1128/mBio.02133-16
47. Ning B, Elbein A D. Purification and properties of mycobacterial GDP-mannose pyrophosphorylase. Archives of Biochemistry and Biophysics. 1999; 362(2): 339-345. doi: 10.1006/abbi.1998.1053
48. Mawuenyega K G, Forst C V, Dobos K M, Belisle J T, Chen J, Bradbury E M, et al. Mycobacterium tuberculosis functional network analysis by global subcellular protein profiling. Molecular Biology of Cell. 2005; 16(1): 396-404. doi:10.1091/mbc.e04-04-0329
49. Sambou T, Dinadayala P, Stadthagen G, Barilone N, Bordat Y, Constant P, et al. Capsular glucan and intracellular glycogen of Mycobacterium tuberculosis: biosynthesis and impact on the persistence in mice. Molecular Microbiology. 2008; 70(3): 762-774. doi: 10.1111/j.1365-2958.2008.06445.x
50. Jackson M, Crick D C, Brennan P J. Phosphatidylinositol is an essential phospholipid of mycobacteria. 2000; 275(39) 30092-9. doi: 10.1074/jbc.M004658200
51. Lea Smith DJ, Martin K L, Pyke J S, Tull D, McConville M J, Coppel RL, et al. Analysis of a new mannosyltransferase required for the synthesis of phosphatidylinositol mannosides and lipoarbinomannan reveals two lipomannan pools in corynebacterineae. Journal of biological Chemistry. 2008; 283(11): 6773-6782. doi: 10.1074/jbc.M707139200