ISSN: 2574-1241
DOI: 10.26717/BJSTR.2018.12.00231
Laxman S Meena. Biomed J Sci & Tech Res

Research Article Open Access

Insights of Rv2921c (Ftsy) Gene of *Mycobacterium tuberculosis* $H_{37}Rv$ To Prove Its Significance by Computational Approach

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

[1]*CSIR-Institute of Genomics and Integrative Biology, India*

[2]*Centre for Interdisciplinary Research in Basic Science, India*

Received: December 14, 2018; **Published:** December 17, 2018

***Corresponding author:** Laxman S Meena, CSIR-Institute of Genomics and Integrative Biology, India

Abstract

After seeing the epidemic of tuberculosis in our country and across world according to WHO report, we right now present with an emergence of treatment/research against this disease. In this study, we therefore exaggerate some important aspects of ftsY (Rv2921c) gene of *Mycobacterium tuberculosis* (*M tuberculosis*) which is a GTP binding and hydrolyzing protein. This gene is 1269bp long and contains four GTP binding motif. The above said protein is involved in Signal Recognition Particle (SRP) pathway. The *M tuberculosis* SRP pathway comprises of two proteins ffh (Rv2916c), ftsY and a RNA subunit 4.5s RNA. This protein interacts with ffh (Rv2916c) gene and another predicted GTP binding protein known as Rv3362c. The said protein is an important part of protein export system which is an essential process for importing and exporting protein that are synthesized in cytoplasm to plasma membrane and other organelles. Thus, this protein might also be important for pathogenesis. This study enlists the effect of disruption of ftsY gene on its interaction with ffh and secretion system. Hence this study might be an important step in the way of eradication of this disease.

Abbreviations

TB: Tuberculosis; M Tuberculosis: Mycobacterium Tuberculosis; AMs: Alveolar Macrophages; RRTB: Resistance to Rifampicin Tuberculosis; MDR-TB: Multidrug-Resistant TB; SDG: Sustainable Development Goals; XDR-TB: Extremely Drug Resistant TB; GTPases: Guanosine Triphosphatases; G-Proteins: GTP-Binding proteins; SRP: Signal Recognition Particle; PDB: Protein Data Bank; NCBI: National Center for Biotechnology Information; SAPS: SAP Application Performance Standard; I-TASSER: Iterative Threading ASSEmbly Refinement; LOMETS: Local Meta Threading Server; RAMPAGE: Ramachandran Plot Analysis; UniProt-GOA: Gene Ontology Annotation.

Introduction

Tuberculosis is a fatal disease which is broadly conveyed and open. It is caused by *Mycobacterium tuberculosis* $H_{37}Rv$ (*M. tuberculosis* $H_{37}Rv$) which is a gram-positive and aerobic bacterium [1]. Pathogenic strain H37Rv of this bacterium is differing from other non-virulent strains like *Mycobacterium smegmatis* (*M. smegmatis*) in various means [2]. After crossing the respiratory tract, this bacterium permanently resides in the alveolar macrophages where they reside for long time without any hindrance by host immune system [3]. It is very important to decrease the level of this disaster which appears to be at its peak and to comprehend this situation, struggle is going on. As it is already known that, BCG is the only candidate vaccine available till now for the curative therapy of tuberculosis (TB). There are many more drugs used in the treatment of this disease but not as effective as BCG vaccine, although in some of the cases BCG also fails to protect an individual from the drastic nature of this disease [4]. In the year 2016 only, there were 6 million novel cases had been reported which develop resistance to Rifampicin (RRTB)–the most effective and first line drug to battle TB [5]. Although there were some strategies made by WHO to end this disease and the management also succeed to a minor level, we right now deal with the situation that need a quick treatment with respect to the total population and the environment of our nation. *M. tuberculosis* $H_{37}Rv$ and HIV co infection are prone for diminishing the CD4+ T cell population which is known as cell mediated arm of immunity and this co-infection is now widespread in all over

the world [6]. Persons with this co-infection are lessening host cell survival due to multifaceted nature of both these bacterium [7]. Thus, as to figure out problem of the pathogenesis of this bacterium, we have to figure out important aspects of this bacterium that may play significant role in its survival inside host cell and avoid many host immunological barrier [8]. To grasp trouble pathogenesis, it is imperative to portray significant highlights of *M. tuberculosis* H_{37}Rv that sanction it to evade the host barrier framework and add to its destructiveness [9]. Many previous studies show the importance of GTP binding and hydrolyzing genes in the survival of the many prokaryotes as well as eukaryotes.

GTP related genes are most significant molecules in various signaling mechanism. As Guanosine Triphosphatases (GTPases) are known to accept a basic part in the survival and bind of various pathogens so accordingly the qualities which tie to GTP likewise have an imperative part in its survival inside the host macrophages [10,11]. GTPases are generally called sub-nuclear switch proteins [12-14]. The key piece of these proteins incorporates restriction in phagosomes improvement, enabling pathogens to get shielded from making tracks in an opposite direction from lysosomes and unsafe free radicals prompted as inborn invulnerable reactions of the host after disease by this bacterium. This discernment gives another path to the advance of threatening to TB drugs [15,16]. In a past couple of years, an expansive work has been done to understand the piece of GTPases in the improvement and advancement of organisms to make them not vulnerable to immune system of the host [17,18]. GTP-restricting proteins (G-proteins) are much monitored signaling substance that takes an interest in cell signaling and bacterial pathogenesis by controlling the movement of related GTPases [19,21]. These proteins especially attach and hydrolyze GTP, which in this way orders or inactivates the GTPases consistently GTPases are particularly checked and work through RNA or ribosome legitimate. G1, G2, G3 and G4 subjects are responsible for specific participation with the guanine nucleotide and effectors proteins.

The underlying two segments are related with interchanges with the phosphate part of the GTP iota and the last segment is locked in with nucleotide specificity [22-24]. The accord grouping contains three concord sequence, GXXXXGK, DXXG and NKXD. According to previously reported literatures this protein acts as Signal Recognition Particle (SRP) in combination with Rv2916c (ffh) [25]. The both genes act as a unit to introduce some cytoplasmic protein across plasma membrane or some other cell organelle. There is an assortment of major pathways have been involved in the protein exporting system of *M. tuberculosis* H_{37}Rv like comprehensive Secretion (Sec) pathway, Twin-Arginine Translocation (Tat) pathway and some inconsequential pathways involved in ESAT-6 secretion system (Esx) and SRP pathway [26]. General secretion pathway system cooperates with post translational secretion of proteins whereas SRP pathways involved in co translational exporting of proteins [27]. SRP is a cytoplasmic ribonucleoprotein and is well conserved in eukaryotes and prokaryotes with somewhat varying composition. The all-around contemplated *Escherichia coli (E. coli)* SRP framework involves 4.5S RNA, ffh (SRP54) and ftsY (SRα), while the greater part of the bacterial SRP pathways comprise of just a single protein, SRP54 and a 4.5S RNA particle. The *M. tuberculosis* H_{37}Rv SRP pathway comprises of two proteins ffh, ftsY and an RNA subunit 4.5S RNA. The said protein is an important part of protein export system which is an essential process for importing and exporting protein that is synthesized in the numerous organelles. Thus, this protein might also be important for pathogenesis.

Hence in this literature, author wants to describe some Insilico aspects about a GTP binding protein known as Rv2921c (ftsY) of *M. tuberculosis* H_{37}Rv. The study comprises of comparison of sequences of ftsY gene of *M. tuberculosis* H_{37}Rv with the other species of the bacterium, search of various interactive partners, phosphorylation capacity and mutation study [28]. As we noted above that ftsY contains GTP binding and hydrolyzing properties which is also important for its activity therefore by mutating specific amino acid of the motif should change the ability of the GTP binding and hydrolyzing properties which thus affects secretion. Also, mutation should also affect its interaction with ffh gene. Therefore, this study gives us the initial steps for proofing the actual mechanism of secretion system of SRP and thus might help in treatment procedure as shown in Table 1.

Table 1: Different computational servers used in this study.

S.No.	Tools	URL	Function	Reference
1	MycoBrowser	https://www.ebi.ac.uk/Tools/msa/clustalo/	For the retrieval of the sequence.	[29]
2	MUSCLE	https://www.ebi.ac.uk/Tools/msa/muscle/	Multiple sequence alignment	[30,31]
3	STRING	https://string-db.org/cgi/network.pl?taskId=BUe3enVFzh8M	For the study of protein-protein interaction.	[32,33]
4	TBpred	http://crdd.osdd.net/raghava/tbpred/	For the prediction of subcellular localization.	[34]
5	DEPP	http://www.pondr.com/cgi-bin/depp.cgi	For phosphorylation sites prediction	[35,36]
6	BCPREDs	http://ailab.ist.psu.edu/bcpred/	For the B-cell epitopes prediction.	[38,39]
7	ProPred	http://crdd.osdd.net/raghava/propred/	For the MHC Class-II binding peptide prediction server	[40]
8	I-TASSER	https://zhanglab.ccmb.med.umich.edu/I-TASSER/	For the protein modeling	[43,44]
9	SAVES	http://servicesn.mbi.ucla.edu/SAVES/	For the model validation	[51,52]

10	I-Mutant 3.0	http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi	For prediction of stability change upon single point mutation.	[53,54]
11	HDOCK	http://hdock.phys.hust.edu.cn/	Protein-protein docking based on the hybrid algorithm of template-based modeling and ab initio free docking.	[56]

Methods and Material

Retrieval of Protein Sequence Database

Mycobrowser database has been used for retrieving sequence (gene and protein) of *ftsY* (Rv2921c) gene [29]. This *ftsY* (Rv2921c) protein sequence has been disentangled in FASTA format *ftsY* is predicted to be involved in insertion and reception of various proteins in the outer side of cell membrane and contains ATP/GTP binding motif.

Multiple Sequence Alignment

For the analysis of multiple sequences alignment of *ftsY* gene (Rv2921c) of *M. tuberculosis* had been done with *ftsY* gene of *Mycobacterium bovis (M. bovis)*, *Mycobacterium marinum (M. marinum)*, *M. smegmatis* and *Mycobacterium leprae (M. leprae)*. MUSCLE online server for multiple sequence alignment approach for ftsY protein analysis. MUSCLE remains for Multiple Sequence Comparison by Log-Expectation it is professed to accomplish both better normal exactness and preferred speed over ClustalW2 or T-Coffee, contingent upon the picked choices [30,31].

Interaction Study by String

STRING database server is utilized for demonstrating the protein-protein association between two or more entities. In the cell cytoplasm, a protein may collaborate with different proteins and work in the web-like manner. The associations incorporate direct (physical) and aberrant (functional) interactions; They branch from computational expectation, from arrangement pass on among life forms and from connections collected from other primary databases [32]. The total number of interactions for each dataset had been kept in STRING and measure somewhere in the range of 0 and 1. The principle of its working is as the score is <0.4 means low interaction, score is between 0.4 to 0.7 means medium interaction and score is >0.7 is high interaction [33].

Prediction of Sub Cellular Localization

Protein Subcellular Localization and prediction confined protein destinations. TBpred is a subcellular localization prediction method for mycobacterial proteins depend on support vector machine learning (profile portion SVM) to foresee the local subcellular sites. Several parameters may be tuned for their appropriate values to get optimum results. The nth SVM model learns from nth class samples with positive labels and rest other samples with negative labels. Prediction of an unknown sample is based upon the maximum score out of four scores, generated by four models specific to four different subcellular compartments [34].

Prediction of Phosphorylation site

DEPP (Disorder Enhanced Phosphorylation Predictor) server evaluates the number of phosphorylated serine, threonine and tyrosine site. This server is depending on the Support Vector Machines prepared on succession profiles improved by data from the spatial setting of tentatively distinguished P-locales. DEPP server predicted the phosphorylated sites of serine, threonine and tyrosine with accurate score. DEPP server is able for predicting phosphorylation by the serine kinases PKA, PKC, MAPK, CKII and by the tyrosine kinases SRC. The nature of expectations is incredibly reliant on the nature of submitted protein structures. Incorrect or inadequate protein structures may prompt wrong forecasts [35,36].

Prediction of B-cell and T-cell epitopes

The prediction of (B & T-cell) epitopes found in ftsY (Rv2921c) protein was done by various online bioinformatics tools. B-cells epitopes prediction was done by using BCPREDS server, T-cell epitopes prediction by the ABCpred server and (MHC-Class II Binding Peptide Prediction) is done by ProPred tool [37-40].

Model Building

Structure modeling of ftsY protein was done by using Iterative Threading Assembly Refinement (I-TASSER). I-TASSER is utilized for the protein structure and function expectation [41-44]. I-TASSER needs the FASTA prearranged sequence of protein and assembles the 3D model of protein by Ab Initio display approach. I-TASSER server is an online stage for protein structure and functions forecasts [45]. I-TASSER pursues three phases to envision the 3D model of the protein. For advance illustration of the secondary structure of the protein, this tool secretly introduced Local Meta Threading Server (LOMETS) which uses H, E and C articulate for alpha-helix, beta-sheet and curl respectively. In I-TASSER server there are also predicted the desired dynamic binding sites of our selected protein were anticipated by COACH online server. Prior to dynamic site-specific docking, the affirmation of actual binding pocket is important. The binding pocket is the site of protein where the ligand interacts reversibly or irreversibly [46]. COACH server is a metaserver, it starts from given structure of target protein; At that point it will make correlative ligand restricting site forecast using the two relative systems, TM-site and S-site, which perceive ligand restricting plan from the database (BioLiP) protein work database by restricting specific substructure and collection profile examination. In the COACH server, yield has positioned top 10 displays by the bunch measure, given C-score, PDB hit, ligand name, complex structure download and agreement restricting buildup. Range estimations of C-score prediction lie somewhere in the range of 0 and 1, where the most noteworthy score demonstrate greater unwavering quality [47].

Model Validation

The evaluation of approval of the created protein structure has been completed by online server RAMPAGE (Ramachandran Plot Analysis). The RAMPAGE server endorses the protein structure on the hypothesis of φ, ψ purpose of individual stores [48-50]. The approval of protein was performed by the structure Analysis and

confirmation server rendition for (SAVES) or, in other words server that depend on checking the stereo compound nature of a protein structure by breaking down residue geometry and by and large structure geometry [51,52].

Mutational Analysis of The Protein

In the Mutational assessment of *ftsY* (Rv2921c) had been finished by via I-MUTANT 3.0 suite server. For studying of all the energy proteins, thinking about analysis of protein quality, free vitality change (ΔΔG) upon single point transformations may in engage the clearing up of process. The vast majority of the ΔΔG esteem is about zero (around 32% of the ΔΔG edifying record ranges -0.5 kcal/ mole) and both the consideration and suggestion of ΔΔG might be either positive or negative for a similar change blurring the relationship among imprecise and expected ΔΔG esteem. Remembering the last goal to vanquish this issue, we describe another pointer that confines among the three change classes: destabilizing changes (ΔΔG<−1.0 kcal/mol), counter balancing changes (ΔΔG>1.0 kcal/mole), and impartial changes (−1.0≤ΔΔG≤1.0 kcal/mole). For the I-MUTANT 3.0 suite score, DDG<-0.5 means (extensive decline of stability), DDG>0.5 infers (increment of stability) and - 0.5<=DDG<=0.5 means (impartial stability) [53,54]. For the figure of protein, consistency change upon a singular point alteration was foreseen by I-MUTANT 3.0 Suite server.

Molecular Docking

Docking has been studied for confirming protein-protein and protein-DNA interaction analysis. Complexes formed after docking were visualized by the PyMOL. Molecular docking of ffh (Rv2916c) with wild type and mutant *ftsY* protein revealed the variation of binding energies, formation of hydrogen bonds and their distances. Docking of the two proteins we have been used HDOCK (http://hdock.phys.hust.edu.cn/) docking server [55]. This server takes input in form of either FASTA sequence or PDB structure. This feature of taken up FASTA sequence makes this server easily available for new comers. The output of the server contains docking score and RMSD value. The lower the RMSD value stronger is the association. HDOCK predicted the protein–protein and protein–DNA/RNA docking. The interactions of the protein-protein and protein–DNA/RNA play an essential role in the assortment of biological process. For the docking, HDOCK is the novel tool for hybrid docking algorithm of template-based modeling and free docking, in which cases with the misleading templates it can be rescued by the free docking protocol. The docking process is fast and consumes about 10-20 min for a docking run [56,57]. HDOCK server performance will become enhanced when more predictions were considered.

Result

Retrieval of Target Protein Sequence

The genomic and proteomic sequences for the gene *ftsY* (Rv2921c) have been retrieved in FASTA format from Mycobrowser database. The gene is 1269bp long and encodes protein of 43kDa. The predicted function of this protein is involving in reception and insertion of the protein at membrane site. Probably it functions as Signal Recognition Particle (SRP). The protein sequence contains four GTP binding motif (DXXG).

Protein-Protein Interaction

For the prediction of protein-protein interaction of ftsY protein is involved with the secretion system, involved the insertion of promising membrane proteins and it is in cytoplasmic membrane. The *ftsY* protein acts as a receptor for the complex formed by the signal recognition particle and the ribosome-nascent chain. The prediction of ftsY functional partner by STRING are the ffh, rplJ, rplQ, Rv3362c, rplK, rplR, rplS, rpmA, rplW, rpmG1

proteins due to the decreasing order of the String server version 10.5 score. The interactions score of ftsY and ffh interaction are very high 0.995 and both gene having same function and act as SRP protein Figure 1.

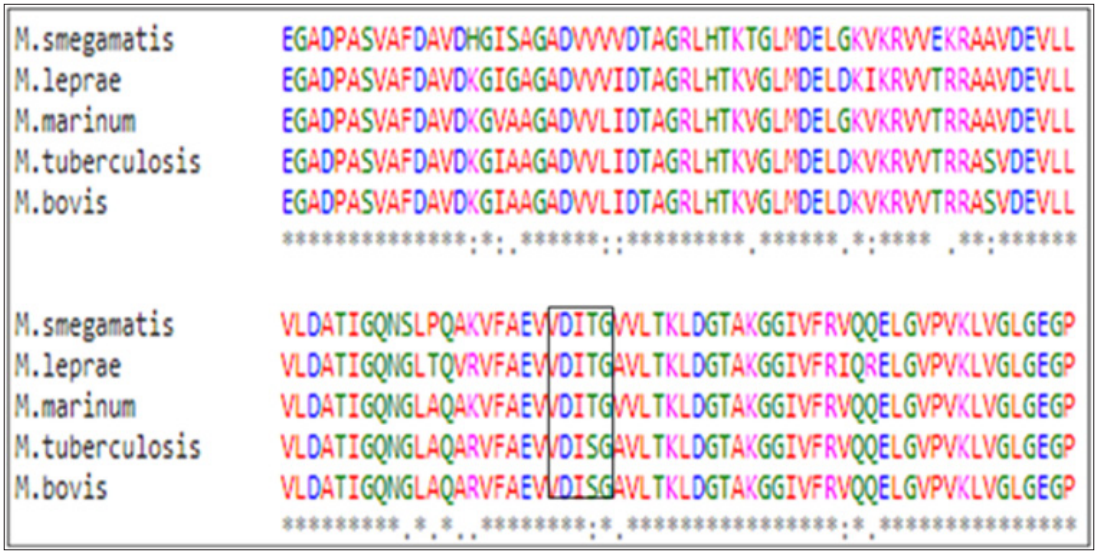
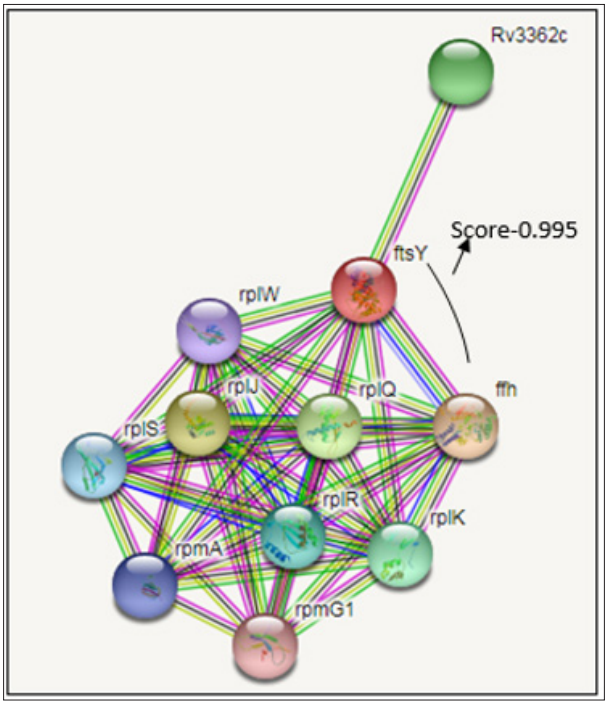


Figure 1: Multiple sequence alignment: Multiple sequence alignment of *ftsY* gene of *Mycobacterium tuberculosis* with other species of this bacterium had been done by MUSCLE server which shows that this is a common protein present in all species and GTP binding motif at D312 position of DXXG motif is common in all variants.

Multiple Sequence Alignment

Multiple sequence alignment of the protein ftsY (Rv2921c) of *M. tuberculosis* with other species of the bacterium like *Mycocaterium bovis, Mycocaterium marinum, Mycocaterium smegmatis, Mycocaterium leprae* has been done by using MUSCLE online server. After putting sequence in FASTA format for all five proteins the server aligns sequence and gives the perfect matches. The percent identity matrix proves its identity among all species. The result also shows the presence of consensus sequence DXXG in all sequences (Figure 2). The presence of the motif in all species is might be an indicator for its importance of the process in the survival of these prokaryotes.



.Figure 2: Protein-protein interaction study by the String server version 10.5: Interaction study of *ftsY* gene has been check by STRING database server which found that this gene interacts with several other genes including ffh gene which is also involved with co translational secretion system and Rv3362c which is probable GTP binding proteins.

Prediction of Sub Cellular Localization

Prediction of the subcellular localization of a protein by using the TBpred server the length of this gene is 422 amino acid residues and the selected approach are Dipeptide composition based SVM. There are different class wise scores from SVM models like cytoplasmic protein, integral membrane protein, secretory protein and protein attached to membrane by lipid anchor. At last, this TBpred server finally predicts that ftsY protein is localized in membrane as an integral membrane protein.

Phosphorylation Site Prediction

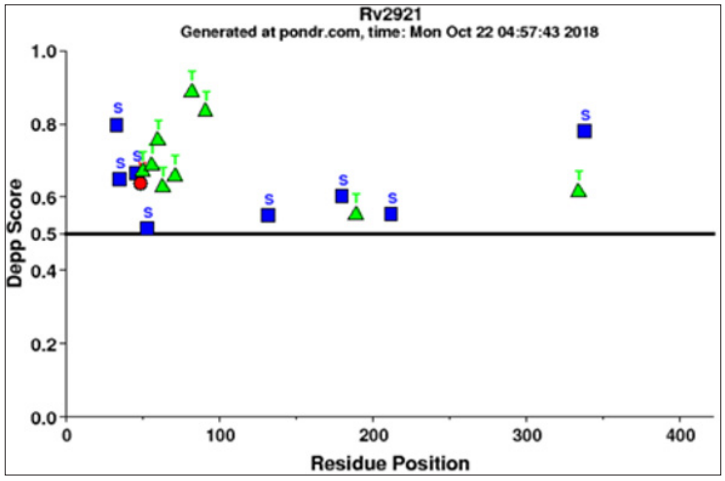


Figure 3: Phosphorylation site prediction: Phosphorylation prediction of this gene has been done by DEPP server which shows that this protein phosphorylated at serine, threonine and tyrosine residues.

DEPP (Disorder Enhanced Phosphorylation Predictor) server evaluates the number of phosphorylated serine, threonine and tyrosine site. Notwithstanding serine, threonine and tyrosine result shows that in ftsY protein there were 8 phosphorylated serine residues out of 18 residues, 9 phosphorylated threonine residues out of 25 residue and 1 phosphorylated tyrosine residue out of 2 residues. In DEPP statistic score are 44.44% are phosphorylated serine, 36% are phosphorylated threonine and 50% tyrosine, phosphorylation prediction also shown in Figure 3. The nature of expectations is incredibly reliant on the nature of submitted protein structures. Incorrect or inadequate protein structures may prompt wrong predictions.

Prediction of B-Cell and T-Cell Epitopes

The prediction of B-cell epitopes by ABCpred server which has predicted the epitopes on this gene taking the overlapping window of 14 amino acids that consequences in the preeminent possibility of the score is 0.90 from the residues region "SVLLVVGVNGTGKT" that starts at 224th position as shown in supplementary Figure 4a. The BCPREDS server are B-cell epitopes prediction which have shows two types prediction as Fixed length epitopes prediction by BCPred and flexible length epitopes prediction by FBCpred which have been shown in (Figure 4b) with the set specificity at 90% and epitopes length set on 14. B cell epitopes prediction fixed length method are predicted on residues 278th -291th, 165th-178th, and 114th-127th and for Flexible length epitopes prediction residues committed are as 282th -295th, 109th -122nd, 165th-178th and 330th-343th in the ftsY protein. For the T-Cell epitopes prediction there is multiple DR-β1 (DRB) alleles were used like HLA-DRB1*0101, HLA-DRB1*0102, HLA-DRB1*0301 which is the T-cell epitopes prediction for the prediction with the MHC Class-II binding region in the antigenic protein sequence of ftsY protein. The predicted binder was visualizing in graphical peak interface as well as in color residue in an HTML interface. Two consensus epitopes were LVVIAALTL in (DRB1_0101) at position 15th-23rd in (DRB1_0102) epitopes are present were – LWIATAVIA and LVVIAALTL, at the 5th -13th and 15th-23rd and in (DRB1_0301) VVIAALTLG residue are present 16th-24th position. In MHC Class-II binding peptide prediction result there are individual alleles at 1% threshold as shown in T-cell epitopes prediction. While at 3% threshold it gives another residue,(not mentioned in the article). For T-cell epitopes prediction result are shown in Figure 4c.

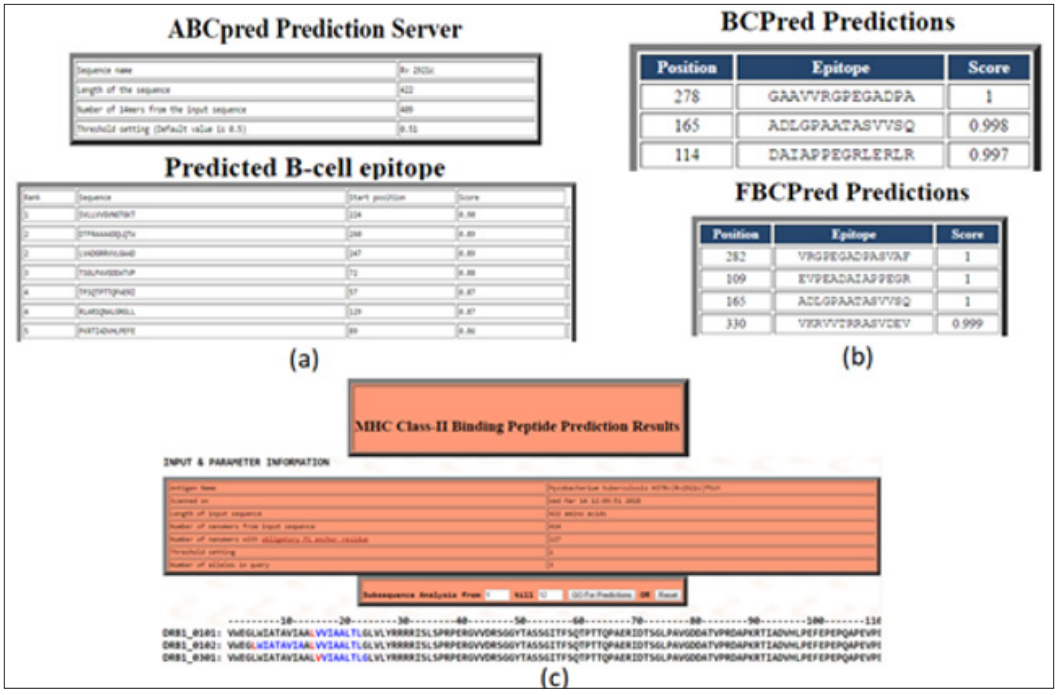


Figure 4: B-cell and T-cell epitopes prediction: Predictions of B-cell and T-cell epitopes had been done by using ABCpred, BCpred, Propred and FBCpred server which results in different prediction of epitopes on this gene.

Model Building

The structural modeling of the wild type and mutant ftsY (Rv2921c) protein and its interactive partner ffh (Rv2916c) protein were prepared from I-TASSER. The quality of modelled protein depends upon the percentage of the favorable region lies above 90% of the value of C-score and RMSD value. The C-score is the confidence score for each model. It is computed by threading layouts arrangement. The C-score changes inside the range from -5 to 2 and higher certainty show is controlled by the higher estimation of C-score. At long last, I-TASSER creates top 5 models according to C score and positioned by group measure among which the figure with higher C score. For protein modeling we modeled the ffh (Rv2916c), ftsY (Rv2921c), mutant ftsY (D45) in this protein modeling we mutate 45 number residue aspartates into alanine, mutant ftsY (D71) residue, mutant ftsY (D312) residue and mutant ftsY (D367) residues as shown in Figure 5.

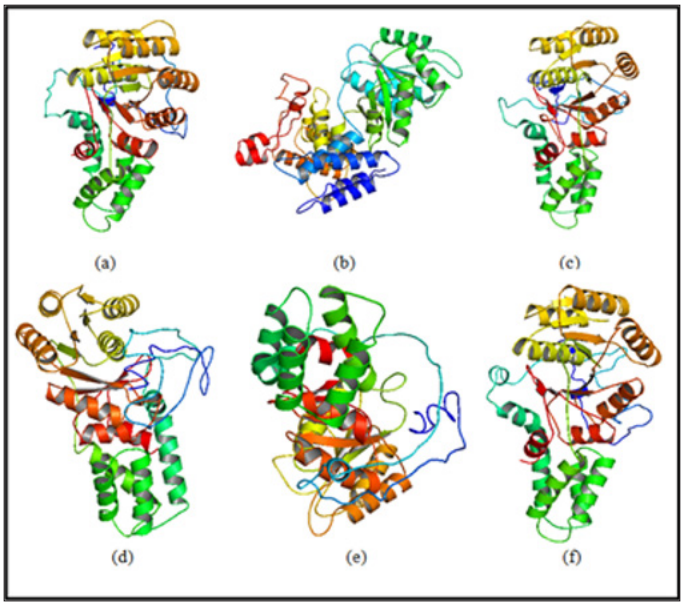


Figure 5: Model building by I-TASSER server: (a) 3D model of Ffh (Rv2916c) (b) 3D model of FtsY (Rv2921c), (c) 3D model of Mutant FtsY (D45), (d) 3D model of Mutant FtsY (D71) (e) 3D model of Mutant FtsY (D312) (f) 3D model of Mutant FtsY (D367).

Model Validation

After modelling of structure, the protein structure was validated through and SAVES server (RAMPAGE, ERRAT and Verify3D). The demonstrated protein was validated by RAMPAGE (Ramachandran plot investigation) which is an online server. After examination of Ramachandran plot of our proteins, the structure demonstrated that have been present in a favored region. Although, other residues were laid in the allowed region and number of residues were laid in outlier region. These parameters of protein structure demonstrating that our displayed protein was of good quality stable and adequate. ERRAT is an online server which approves the protein structure on the premise of the nuclear connection between various sorts of atoms. The ERRAT analysis shows overall quality factor of our model protein is good and satisfactory. The Verify3D strategy evaluates protein structure by utilizing three-dimensional profiles. This program examines the similarity of a nuclear model (3D) with its own amino acid sequence which is 1 dimensional. Every deposit is doled out a basic class in radiance of its area and condition (alpha, beta, circle, polar, non-polar and so on). The score ranges from -1 (poor score) to +1 (great score). 82.29 - 95.25% of the buildup had a found the middle value of 3D-1D score >=0.2 that is perceptive for our demonstrated protein result. is pass as shown in Table 2 Verify 3D result as shown in The protein model ftsY (Rv2921c) validation which is done by SAVES metaserver results convoluted that all of the protein model build are good and satisfactory which is shown in Table 2.

Table 2: Model evaluation by SAVES (Statistical Analysis and Verification Server).

S.No.	Model	RAMPAGE		Verify 3D	ERRAT
1	*ftsY*	Residues in favoured region Residues in allowed region Residues in outlier region	70.7% 13.6% 15.7%	93.13% (PASS)	81.8841
2	*ffh*	Residues in favoured region Residues in allowed region Residues in outlier region	78.8% 14.5% 6.7%	82.29% (PASS)	87.6953
3	*ftsY* D45A	Residues in favoured region Residues in allowed region Residues in outlier region	68.7% 15.8% 15.5%	92.40% (PASS)	76.5133
4	*ftsY* D71A	Residues in favoured region Residues in allowed region1 Residues in outlier region	71.8% 5.0% 13.1%	95.25% (PASS)	74.8184
6	*ftsY* D312A	Residues in favoured region Residues in allowed region Residues in outlier region	70.2% 15.3% 14.6%	89.55% (PASS)	77.1845
7	*ftsY* D367A	Residues in favoured region Residues in allowed region Residues in outlier region	69.9% 17.2% 12.9%	90.02% (PASS)	78.4504

Mutational Analysis

In ftsY (Rv2921c) protein had been characterized as a GTP binding protein which has consensus binding sequence DXXG which promote the link sub sites for binding in Mg^{2+} and -phosphate of GTP. In the consensus sequence DXXG, aspartate is crucial for GTP binding and hydrolyzing activity; Therefore we study the prediction of stability change upon the single point protein mutation by I-Mutant suite. Earlier studies have been shown that for GTP binding protein, when aspartate mutate with the alanine then the functional protein may loss the function. So, we mutate the DXXG aspartate into the Alanine for predicting the stability of the protein after change the single amino acid. In I-Mutant suite studies, we have seen that there are four consensus sequences in this protein on residue 45, 71, 312 and 367. After the analysis of predicted score by I-Mutant we have found that large stability decreased on the 45-position residue because the parameter of stability decreasing value is when -0.5 or below value towards negative and our outcome result is -1.39 which clearly shows at the position on 45 there will largely decreasing value are shown in Table 3.

Table 3: I-Mutant 3.0 server.

S.No.	Position	New Residue	Temp.	pH	DDG value
1	ftsY D45	A	25	7	-1.39
2	ftsY D71	A	25	7	-0.13
3	ftsY D312	A	25	7	-0.36
4	ftsY D367	A	25	7	-0.76

Molecular Docking

In the docking content, we here doing protein-protein docking for knowing the interaction changes upon mutation of important residue aspartate involved in GTP binding and hydrolyzing motif in the protein. Aspartate would be mutating with alanine. The wild type ffh protein dock with wild type ftsY as a control value for our study for the experimental study like wild type ffh protein dock with mutant ftsY (D45A), ftsY (D71A) , ftsY (D312A) , ftsY (D367A) to identify change in binding patterns using HDOCK docking server [14,15] as shown in Figure 6. The binding energy and formation of hydrogen bonds to each molecule by the drug were calculated and the RMSD value was 0.38 the highest for D312 are seen in Table 4.

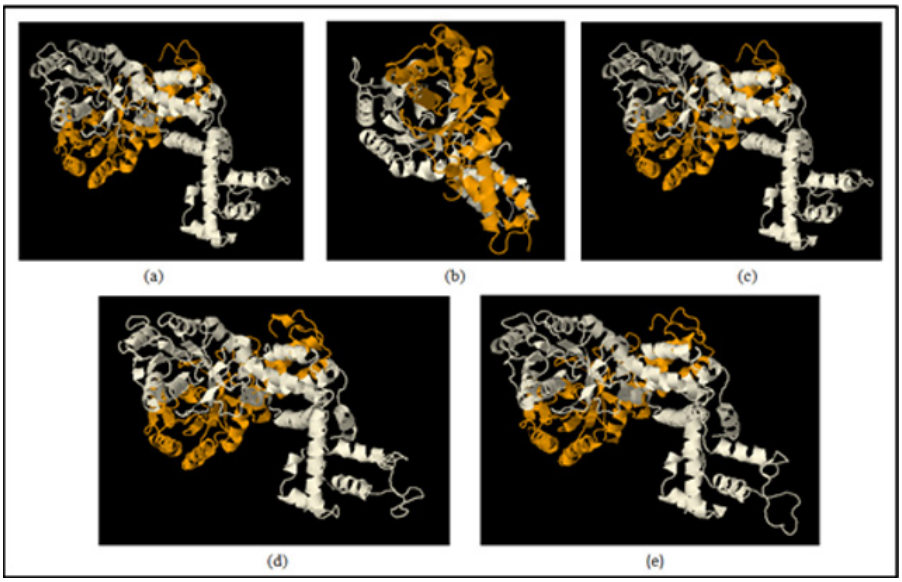


Figure 6: Protein-Protein docking by HDOCK: (a) 3D model of wild type *ffh* and *ftsY* (b) 3D model of docking of *ffh* with mutated *FtsY* (D45A) (c) 3D model of docking of *ffh* with mutated FtsY (D71A) (d) 3D model of docking of *ffh* with mutated FtsY (D312A), (e) 3D model of docking of *ffh* with mutated FtsY (D367A).

Table 4: Molecular docking by HDock server.

S.No.	Receptor	Ligand	PDB ID	Docking Score	Ligand RMSD (Å)
1	ffh	ftsY	2XXA	-306.79	0.21
2	ffh	ftsY D45A	5L3Q	-362.89	0.21
3	ffh	ftsY D71A	2XXA	-362.89	0.21
4	ffh	ftsY D312A	2XXA	-308.08	0.38
5	ffh	ftsY D367A	2XXA	-362.72	0.19

Discussion

In the current circumstance, we can see that there no defensive and healing treatment to destroy tuberculosis totally aside from BCG. Past many years of research as of now demonstrates that BCG gives constrained insurance against tuberculosis yet Fails in securing MDR, TDR and XDR instances of tuberculosis. There is a persistent exertion has been put by researchers with the end goal to build the adequacy of the antibody and in looking for new medication targets. GTP binding genes come out to be as novel targets for treatment of this disease [58-59]. GTP binding and hydrolyzing protein ftsY (Rv2921c) has been proved for possessing the same activity in many studies, this gene strongly interacts with ffh gene (Rv2916c) with score of 0.995 [32]. At the other hand, ftsY protein is also interacts with another predicted partner Rv3362 which also possess GTP binding activity [33]. Multiple sequence alignment result shows that this gene is universally present in all species of this bacterium with having some minute differences. All species contains same GTP binding motif DXXG at same location [30,31]. This protein is predicted to be present as an integral membrane protein [34]. Phosphorylation site prediction gives us result that this protein phosphorylated at serine, threonine and tyrosine [35,36]. The prediction of (B and T cell epitopes prediction) has been done by using ABCpred, BCpred, FBCPred and ProPred server [38-40]. 3D model of this protein has been made by I-TASSER server [43,44] and validated by SAVES server [51,52]. Mutational analysis has been done by I-Mutant 3.0 server and it shows that mutation on 45 residue aspartates with alanine at 25°C temperature and pH 7 noted the larger decrease in stability [53,54]. Molecular docking results of these proteins before and after mutation prove that interaction between these two proteins decrease maximal at position 312 residue with having value of 0.38 [56]. In summarizing our work we can narrate the essentiality of ftsY gene in secretion process and thus in pathogenesis of this bacterium. The two above mentioned proteins work in cooperative manner and produce a summative effect. The two proteins is key Regulator for the co-translational secretion system and thus might also play an important role in pathogenesis of the bacterium.

Conclusion

We need an emergent step to stop TB epidemic all around the world. This study concludes with knowledge of some important aspects of ftsY gene and its interaction with ffh gene. This interaction is very much important for one of the secretory pathways assembled in prokaryotes. Although the experiments enlisted in this manuscript are not enough, but these experiments provide us with better knowledge and initial steps in further in vitro and in vivo experimental works.

Acknowledgement

The authors acknowledge financial support from the Department of Science and Technology-SERB, Council of Scientific and Industrial Research-Institute of Genomics and Integrative Biology under the research project GAP0145.

References

1. Brosch R, Gordon SV, Marmiesse M, Brodin P, Buchrieser C, et al. (2002) A new evolutionary scenario for the Mycobacterium tuberculosis complex. Proc Natl Acad Sci USA 99(6): 3684-3689.
2. Tyagi JS, Sharma D (2002) Mycobacterium smegmatis and tuberculosis. Trends in Microbiol 10(2): 68-69.
3. Gagneux S (2018) Ecology and evolution of *Mycobacterium tuberculosis*. Nature Reviews Microbiology 16(4): 202-213.
4. Teo SSS, Shingadia DV (2006) Does BCG have a role in tuberculosis control and prevention in the United Kingdom? Arch Dis Child 91(6): 529-531.
5. (2018) World Health Organization, Global Tuberculosis Report.
6. SD Lawn (2004) AIDS in Africa: The impact of coinfections on the pathogenesis of HIV-1 infection. J Infect 48(1): 1-12.
7. Cunha R, Maruza M, Montarroyos (2017) Survival of people living with HIV who defaulted from tuberculosis treatment in a cohort, Recife, Brazil. BMC Infectious Diseases 17: 137.
8. Barry CE, Boshoff HI, Dartois V (2009) The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. Nature Reviews Microbiology 7(12): 845-855.
9. Nagu TJ, Aboud S, Mwiru R, Matee MI, Rao M, et al. (2016) Tuberculosis associated mortality in a prospective cohort in Sub Saharan Africa: association with HIV and antiretroviral therapy. Int J Infect Dis 56: 39-44.
10. Rajni, Meena LS (2010) Survival mechanisms of pathogenic Mycobacterium tuberculosis H_{37}Rv. FEBS Journal 277(11): 2416-2427.
11. Rajni, Meena LS (2010) Guanosine triphosphatases as novel therapeutic targets in tuberculosis. International Journal of Infectious Diseases 14(8): e682-e687.
12. Meena LS, Chopra P, Bedwal RS, Singh Y (2008) Cloning and characterization of GTP-binding proteins of Mycobacterium tuberculosis H_{37}Rv. Enzyme and Microbial Technology 42(2): 138-144.
13. Brennan PJ (2003) Structure, function and biogenesis of the cell wall of Mycobacterium tuberculosis. Tuberculosis 83(1-3): 91-97.
14. Brumell JH, Scidmore MA (2007) Manipulation of rab GTPase function by intracellular bacterial pathogens. Microbiology and Molecular Biology Reviews 71(4): 636-652.
15. Dever TE, Glynias MJ, Merrick WC (1987) GTP-binding domain: Three consensus sequence elements with distinct spacing. Proc Natl Acad Sci USA 84(7): 1814-1818.
16. Seabra MC, Mules EH, Hume AN (2002) Rab GTPases, intracellular traffic and disease. Trends in Molecular Medicine 8(1): 23-30.
17. Freedman NJ, Lefkowitz RJ (1996) Desensitization of G protein-coupled receptors. Recent Progress in Hormone Research 51: 319-351.
18. Watkins HA, Baker EN (2006) Structural and Functional Analysis of Rv3214 from Mycobacterium tuberculosis, a Protein with Conflicting Functional Annotations, Leads to Its Characterization as a Phosphatase. J Bacteriol 188: 3589-3599.
19. Carvalho ATP, Szeler K, Vavitsas K, Shina JA, Kamerlin CL, et al. (2015) Modeling the mechanisms of biological GTP hydrolysis Author links open overlay panel. Archives of Biochemistry and Biophysics 582: 80-90.
20. DuX, FreiH, KimSH (2000) The mechanism of GTP hydrolysis by Ras probed by Fourier transform infrared spectroscopy. Journal of Biological Chemistry 275(12): 8492-8500.
21. Capriotti P, Fariselli I, Rossi, Casadio R (2008) A three state prediction of single point mutations on protein stability changes. BMC Bioinformatics 9: S2-S6.

22. Kjeldgaard M, Nyborg J, Clark BF (1996) The GTP binding motif: variations on a theme. FASEB Journal 10(12): 1347-1368.
23. Dines M, Sendersky E, David L, Schwarz R, Adir N, et al. (2008) Structural, functional, and mutational analysis of the NblA protein provide insight into possible modes of interaction with the phycobilisome. The Journal of Biological Chemistry 283(44): 30330-30340.
24. Bereswill S, Waidner U, Odenbreit S, Lichte F, Fassbinder F, et al. (1998) Structural, functional and mutational analysis of the pfr gene encoding a ferritin from Helicobacter pylori. Microbiology 144(6): 2505-2516.
25. Kjeldgaard M, Nyborg J, Clark BFC (1996) The GTP binding motif: variations on a theme. FASEB J 10(12): 1347-1368.
26. Feltcher ME, Sullivan JT, Braunstein M (2010) Protein export systems of Mycobacterium tuberculosis: novel targets for drug development. Future Microbiology 5(10): 1581-1597.
27. Natale P, Bruser T, Driessen AJ (2008) Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane--distinct translocases and mechanisms. Biochim Biophys Acta 1778(9): 1735-1756.
28. Watkins HA, Baker EN (2006) Structural and functional analysis of Rv3214 from Mycobacterium tuberculosis, a protein with conflicting functional annotations, leads to its characterization as a phosphatase. J Bacteriol 188(10): 3589-3599.
29. Kapopoulou A, Lew JM, Cole ST (2011) The MycoBrowser portal: a comprehensive and manually annotated resource for mycobacterial genomes. Tuberculosis 91(1): 8-13.
30. Li W, Cowley A, Uludag M, Gur T, McWilliam H, et al. (2015) The EMBL-EBI bioinformatics web and programmatic tools framework. Nucleic Acids Research 43(1): 580-584.
31. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution 30(4): 772-780.
32. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, et al. (2017) The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Research 45(Database issue): D362-D368.
33. Lewis AC, Saeed R, Deane CM (2010) Predicting protein-protein interactions in the context of protein evolution. Molecular Biosystem 6(1): 55-64.
34. Rashid M, Saha S, Raghava GP (2007) Support Vector Machine based method for predicting subcellular localization of mycobacterialproteins using evolutionary information and motifs. BMC Bioinformatics 8: 337.
35. Blom N, Gammeltoft S, Brunak S (1999) Sequence- and structure-based prediction of eukarotic protein phosphorylation sites. Journal of Molecular Biology 294(5): 1351-1362.
36. Blom N, Sicheritz Ponten T, Gupta R, Gammeltoft S, Brunak S, et al. (2004) Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. Proteomics 4(6): 1633-1649.
37. Monu, Meena LS (2016) Biochemical characterization of PE_PGRS61 family protein of *Mycobacterium tuberculosis* H_{37}Rv reveals the binding ability to fibronectin. Iran J Basic Med Sci 19(10): 1105-1113.
38. EL Manzalawy Y, Dobbs D, Honavar V (2008) Predicting linear B-cell epitopes using string kernels. Journal of Molecular Recognition 21(4): 243-255.
39. Y EL Manzalawy, Dobbs D, Honavar V (2008) Predicting flexible length linear B-cell epitopes. Comput Syst Bioinformatics Conf 7: 121-132.
40. Singh H, Raghava GP (2001) ProPred: Prediction of HLA-DR binding sites. Bioinformatics 17(12): 1236-1237.
41. Beg MA, Shivangi, Thakur SC, Meena LS (2018) Structural prediction and mutational analysis of Rv3906c gene of *Mycobacterium tuberculosis* H_{37}Rv to determine its essentiality in survival. Adv Bioinformatics 1-12.
42. Kumar S, Nath O, Govil S, Pathak AN, Sumit Govil (2014) 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 6(2):120-123.
43. Yang J, Yan R, Roy A, Xu D, Poissn J, et al. (2015) The I-TASSER Suite: protein structure and function prediction prediction. Nature Methods 12(1): 7-8.
44. Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein and function prediction. Nat Protoc 5(4): 725-738.
45. Zhang Y (2008) I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 9: 40.
46. Yang J, Roy A, Zhang Y (2013) Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. Bioinformatics 29(20): 2588-2595.
47. Wang S, Sun S, Li Z, Zhang R, Xu J, et al. (2017) Accurate De Novo Prediction of Protein Contact Map by Ultra-Deep Learning Model. PLoS Computational Biology 13: e1005324.
48. Wallner B, Elofsson A (2003) Can correct protein models be identified. Protein Science 12(5): 1073-1086.
49. Ho BK, Brasseur R (2005) The Ramachandran plots of glycine and pre-proline. BMC Struct Biol 5:14.
50. Lovell SC, Davis IW, Arendall WB, Bakker PI, Word JM, et al. (2003) Structure validation by Calpha geometry: phi,psi and Cbeta deviation. Protein 50(3): 437-450.
51. Colovos C, Yeates TO (1993) Verification of protein structures: patterns of non bonded atomic interactions. Protein Science 2(9): 1511-1519.
52. Bowie JU, Luthy R, Eisenberg D (1991) A method to identify protein sequences that fold into a known three-dimensional structure. Science 253(5016): 164-170.
53. Capriotti E, Fariselli P, Calabrese R, Casadio R (2005) Predicting protein stability changes from sequences using support vector machines. Bioinformatics 21: 5-8.
54. Capriotti E, Fariselli P, Calabrese R, Casadio R (2005) I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Research 33: 306-310.
55. Remmert M, Biegert A, Hauser A, Soing J (2011) HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nature Methods 9(2): 173-175.
56. Yan Y, Zhang D, Zhou P, Li B, Huang SY, et al. (2017) HDOCK: a web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy. Nucleic Acids Res 45: 365-373.
57. Yan Y, Wen Z, Wang X, Huang SY (2017) Addressing recent docking challenges: A hybrid strategy to integrate template-based and free protein-protein docking. Proteins 85(3): 497-512.
58. Shaw S, Meena LS (2016) GTPases: Prerequisite Molecular Target in Virulence and Survival of Mycobacterium tuberculosis. International Journal of Molecular Biology 1(1): 1-2.
59. Meena LS (2018) GTPases: a significant signalling molecule in TB infection. International Journal of Biotechnology & Bioengineering 4(6): 124-129.

ISSN: 2574-1241

DOI: 10.26717/BJSTR.2018.12.002231

Laxman S Meena. Biomed J Sci & Tech Res