Volume : 5 | Issue : 1 | JANUARY 2016 • ISSN No 2277 - 8179

Research Paper

In Silico Analysis of Mmpl Gene Family of Mycobacterium Tuberculosis: a Novel Target for Anti-Tb Drugs

Medical Science

KEYWORDS : In silico analysis, Multi drug resistant Mycobacterium tuberculosis, ATP binding cassette, Resistance nodulation cell division, Multidrug and toxic efflux.

AFTAB ALAM	Center for Interdisciplinary Research in Basic Science, Jamia Millia Islamia, New Delhi, India- 110025
SHAHNAWAZ ALI	Center for Interdisciplinary Research in Basic Science, Jamia Millia Islamia, New Delhi, India- 110025
MD. ZUBBAIR MALIK	Center for Interdisciplinary Research in Basic Science, Jamia Millia Islamia, New Delhi, India- 110025
ROMANA ISHRAT	Center for Interdisciplinary Research in Basic Science, Jamia Millia Islamia, New Delhi, India- 110025

ABSTRACT

Mycobacterium tuberculosis is an infectious agent of human tuberculosis. The genome of Mycobacterium tuberculosis revealed that 13 membrane proteins of MmpL family play important role in drug efflux system. With the evolution of tuberculosis to Multi drug resistant tuberculosis and Extensive drug resistant, there is an urgent need of the new drugs for the less known MmpL family of transporters. With an initial assumption that MmpL is critically responsible for the host-pathogen contact, we analyzed MmpL protein family to uncover the important aspects like their structure, function, sub-cellular location, and domains. In this work we predicted function of each protein of MmpL family and two sets of important residues that are actively taking part in the transport mechanism and are presented as a promiscuous target. The present approach is an effective method in understanding the common mechanism of their function at molecular level; moreover, this study will help in developing more potent inhibitors that would worth in treating wide range of infections. Thus, this protein will serve a better target for devising new therapies. We hope that the information on MmpL protein family will be innovatory for further in-vitro analysis of this disease.

INTRODUCTION

Tuberculosis is an infectious disease, caused by the bacillus *Mycobacterium tuberculosis* (Ahmad, 2011). It typically affects the lungs (pulmonary TB) but can affect other sites as well (extra pulmonary TB). In 2013, it is estimated that 9.0 million people developed TB and about 16.7% (i.e. 1.5 million) died from the disease, while 4% (0.36 million) of whom were HIV-positive. According to the Global tuberculosis report 2014 by WHO about 37 million lives were saved in 13 years' time (from 2000 to 2013). India is second most prone area for tuberculosis harbouring 2.2 million patients, with 17.6% (i.e. 3.87 lakh) cases arriving per year including 0.66 lakh new cases annually.

Many *Mtb* strains are resistant to the commonly used antibiotics, due to the slow uptake of drugs as reported by Brennan and Nikaido, in 1995. Other factors contributing to the resistance may be target modification (Lambert et al., 2005), enzyme inactivation etc. (Schweizer, 2003). However, the most prevailing mechanism for the drug resistance is through drug system efflux (Adams et al., 2011). Efflux pumps include proteins that transport varied range of substrates (including anti-TB drugs). These efflux systems summarized and linked to MDR-TB are from various super-families like RND (resistance-nodulation-cell division), ABC (ATP binding cassette), MATE (multidrug and toxic efflux) and SMR (small multidrug resistance). ABC family of efflux pumps uses ATP as a source of energy, while RND, MATE and SMR require proton motive force (Paulsen et al., 1996). RND family of proteins mediates the transport of wide variety of substrates like cations, drugs, fatty acids etc. (Putman et. al., 2000). These are large membrane proteins with 12 transmembrane domains and two cytoplasmic loops between $1^{st}/2^{nd}$ and $7^{th}/8^{th}$ domains respectively. Long back Tseng and co-workers in 1999 reported the presence of RND family in all the organism of major sub kingdoms. RND family of transporters include our

gene of interest i.e. MmpL (Mycobacterial membrane proteins, Large).The genome of M. tuberculosis (*Mtb*) consists of 13 genes designated as MmpL (Cole et al 1998), the co-localization of MmpL gene with the *pks* genes and genes responsible for lipid metabolism suggests there similar function.

More than 50% of the drug targets are membrane proteins, the number is expected rise in the near future (Kwasi et. al., 2005). We still do not know the proper functions of MmpL's excluding 3, 4, and 7 for which expressed data are available. MmpL family serve as potential diagnostics targets, and also this family have role in virulence and studies need to be carried out on designing drug target against this (MmpL) family. So, we have tried to explore the maximum information of MmpL family by using of knowledge of updated bioinformatics tools and techniques including 3D-structure prediction.

In order to figure out the role, binding site, functions etc, we devised a simple workflow to explore this (MmpL) gene family as a drug target in MDR-TB, which is the main focus of the present work. This work will be a milestone in the emerging field of drug discovery development against TB.

MATERIALS & METHODS

Sequence Analysis and Phylogenetic Tree

The FASTA sequence along with their UniProt ID and primary accession number of 13 proteins of MmpL were retrieved. The BLASTp (Altchul et al., 1979) program screened homologous sequences form its database and select those sequences that are more similar or identical to our query sequence. We have derived the function of MmpL protein, which are uncharacterized till now through sequence similarity to a well characterized homologous protein in the gene bank. BLASTp (Altchul et al., 1979) is used for both, identifying a query of amino acid sequences and for finding similar sequence in a protein database. The database provides a picture of the functional properties of the related sequence. The Multiple Sequence Alignment (MSA) of MmpL gene pool shows the relationships among the family members, and it is the alignment of two or more sequences of protein or nucleic acid of similar length. Phylogeny analysis of MmpL protein was carried out using the UPGMA (Unweighted Pair Group

Method with Arithmetic Mean). This analysis gives us the clue about protein's ancestry image.

Analysis of Physiochemical Characteristics

The physicochemical properties of MmpL protein sequences were analyzed using property values of residues as averaged over the whole sequence. The number of amino acids and their frequencies were checked by CLC-free Workbench (version 6.8.2). The average molecular weight (AMW), number of negatively and positively charged residue, extinction coefficient (EC), isoelectric point (pI), instability index (II), Aliphatic index (AI),and grand average of hydropathy (GRAVY) value were predicted using the ExPASy bioinformatics resource tool (www.web.expasy.org/protparam/).

Structure and Function Prediction

The main goal of protein modelling is to predict a structure from its sequence within accuracy that is comparable to the best results achieved experimentally. The three-dimensional structure of a protein from its

amino acid sequences were predicted, that is the prediction of its folding and its secondary, tertiary, and quaternary structure from its primary structure. There are a number of tools available that can predict the 3D structure of MmpL family protein from its homology. The query protein sequences were used to for BLASTp search against the PDB database with default parameters to choose a homologous structure as templates. We focus completely on the template-based approaches, in which the function of a protein is allotted based mainly on its similarity to other proteins whose function is known (Khan et al., 2015). In our study, we have used five modeller tools and servers such as *Modeller* (Sali and Blundell, 1993)(Fiser et al., 2000), I-TASSER (Yang and Zhang, 2015), *RaptorX* (Källberg et al., 2014), *SWISS-MODEL* (Biasini et al., 2014), *and Phyre 2* (Kelley et al., 2015) for the generation of the 3D model of our proteins. The process of 3D modelling of MmpL proteins (Fig.

1) (Laskowski et al., 1993).

The query sequence of MmpL proteins were submitted to the various 3D structure prediction tools (Fig. 2). The various structure related parameters have been considered such as z-score, c-score, RMSD to generate the 3D structure model. The values of these parameters were analysed and selected the best built model. A protein structure predicted by tools and server will be subjected to many sources of errors. The validation of the obtained model was done using PROCHECK suit of programs, which provides a check on the stereo chemical quality of protein structure. It gives the overall quality of the structure as compared with well refined structure of the same resolution (Laskowski et al., 1993).

Automated function prediction is a dynamic research field, with a growing community of bioinformaticians as observed (Stephens et al., 2014). The accurate annotation of protein function is important in understanding life at the molecular level. MmpL gene family exists in the *Mycobacterium tuberculosis* and their biological functions are unknown. Three dimensional structures of these proteins may suggest biochemical and biological function. Here we have used various tools (Fig. 3) to predict biological function and their role in biological process on the basis of their structure.

Functional Residues and Ligands Prediction

Membrane proteins carry out a miscellaneous variety of functions and are used as mainly as drug targets for pharmaceutical agents. The set of information on potential amino acid residues for the function of membrane proteins is considerable for understanding the sequence–structure–function relationship of membrane proteins as well as predicting the functional residues from sequence/structure (Gromiha et al., 2009)(Ali et al., 2014). We used COACH server to predict the probable binding site and possible ligands for our MmpL gene family. COACH server generated possible ligand binding site predictions using two comparative methods, TM-SITE and S-SITE, which recognize ligand-binding templates from the BioLiP protein function database by binding- specific substructure and sequence profile comparisons. These predictions also combined with results from other methods (including COFACTOR, FINDSITE and ConCavity) to generate final ligand binding site predictions.

RESULT AND DISCUSSION

Sequence Analysis and Phylogenetic Tree

A sequence can be recognised as a homolog of a known protein if the pairwise sequence identity exceeds a statistically driven threshold (e.g. more than 30% sequence identify of an E-value less than 0.001 (Lubec et al.,

2005). Phylogenetic analysis show the evolutionary relationships among groups of MmpL family and give the indication about protein's ancestry background (Fig. 4-a). The MmpL family is split into three clusters has closely related to sequences, while MmpL1 just behave like an out group and ancient protein which divided into to give rise to cluster 1(MmpL4, 5 and 9), cluster 2 (MmpL6 and 2) and clusters 3 (MmpL3, 7, 8, 10, 11, 12, and 13). We analysed the sequence of MmpL gene family consisting of 13 genes. Coding equivalent but varied 13 proteins as obtained from the insight into the multiple sequence alignment for all these using ClustalW. They consist of various informative blocks possessing intentionally conserved and substantially conserved residues (shown in Green & Red) deriving same property along the time of gene evaluation. In MSA Glycine (G) at position 132 and 267, Leucine (L) at position 196, Prolene (P) at position 347 and 412, Alanine (A) at position 348 Threonine (T) at position are found conserved 344 (Fig. 4b). Since these residues are preferred to found in hydrophobic core at the protein. So, these conservations here convey us that these residues are part of transmembrane helical region. With conserved anatomical structure in all the MmpL's while other which are partially conserve. We further subjected the sequence for Inter ProScan and found similar results with domain conservation patterns which were then verified with Pfam for family that describes all MmpL's falls in transporter protein like familes.Various structure and sequence specific architecture of transporter like RND, ABC, MATE & SMR protein families. MmpL's was found more relevant to RND and ABC family of transporters thus giving us a view for being as a good target for MDR and XDR tuberculosis may function for drug-efflux pump thus, providing multi or extensive drug resistance for antibiotics.

Analysis of Physiochemical Properties of Selected Mmpl's Gene Family

By using Tuberculist database we could retrieved physiochemical properties like isoelectric point(PI), gene length (bp), protein length (kb), locations (kb) of MmpL family proteins in *Mycobacterium tuberculosis* genome, mass of each protein (Da), functional category and functions (Table. 1).

Structure and Function Prediction

Homology modelling and ab initio modelling (I-Tasser) techniques were used for 3D Structure prediction of all MmpL protein family. Our template-selection criterion includes (a) Large alignment (b) maximum identity (c) less gaps. ab initio modelling method uses at least 10 templates to build 3D structure of protein. The data obtained was comparatively analysed with the data of tools (Modellar, RaptoX, Phyre2 and Swiss Modellar) using homology modelling approach. On the basis of basic parameters such as z-score ($1 \leq z \leq 5$), C-score ($-5 \leq c \leq 2$), RMSD (≤ 4) and the percentage of allowed residues except Glycine (85 to 98% residues can be considered). The PROCHECK server validates each generated 3D structure by checking the stereo chemi-

cal properties of a protein (Fig. 5).

Once the 3D structure of protein is obtained, the protein function can be predicted using 3D structure, as 3D structure is considered evolutionarily conserved than to sequence (Kaczanowski and Zielenkiewicz, 2010). The dcGo server(Fang and Gough, 2013) used to predict the function on the basis of gene ontology, dcGO server gives reliable function predictions which are verified by other property explorer like ProFunc(Laskowski et al., 2005), CATH (Revigo), Prosite(Sigrist et al., 2010) and Ppopen (Rost et al., 2004). We selected some appropriate functions for each MmpL gene. Being of same family of proteins the predicted functions almost look similar (see Table. 2).

Possible Ligands and Binding Sites Predicted

Membrane proteins carry out a miscellaneous variety of functions and are used mainly as drug targets of pharmaceutical agents. The set of information on potential amino acid residues for the function of membrane proteins is considerable for understanding the sequence–structure–function relationship of membrane proteins as well as predicting the functional residues from sequence/structure (Gromiha et al., 2009). The COACH server used here to predict positions of all possible functional residues as well as their possible ligands. The LigPlot^{+} software (Laskowski and Swindells, 2011) was used to generate 2D ligand-protein interaction diagrams (Fig. 6).

MmpL1 (958 residue long protein) Set-A **(**Table 4**)** residues show the bonding relationship with human neutrophil peptides (HNPs) that belong to a family of antimicrobial and cytotoxic peptides known as 'defensins' (Fig. 6A). Set-B (Table 4) binding ZCT (Actinorhodin) that exhibits antibiotic activity (Fig. 6B).

MmpL2 (968 residue long protein) .Set-A (Table 4) residues are bind with the CLR (Cholesterol) which is vital for bacterial persistence (Fig. 6C) (Pasca et al., 2005). Set-B (Table 4) residues bind with the LMT (Dodecyl β-D-maltoside) that exert cleansing (oil-dissolving) and antimicrobial effects through a surface action possessing both hydrophilic and hydrophobic properties (Fig.6D).

MmpL3 (944 residue long protein) Set-A **(**Table 4**)** residues bind with the CLR (Cholesterol). (Fig. 6E) (Pasca et al., 2005). Set-B **(**Table 4**)** residues bind with the LMT (**Dodecyl** β-D-maltoside). (Fig. 6F).

MmpL4 (967 residue long protein) Set-A **(**Table 4**)** residues binds with the FES [2FE-2S]-CLUSTER. Usually,[2Fe-2S]-clusters are not directly involved in the catalytic transformation of substrate, but rather supply electrons to the active site (Fig. 6G) (Converse et al., 2003). Set-B **(**Table 4**)** residues bind with the MIY (MINOCYCLINE**).** A substance that is biostatic or biocidal at low concentrations towards bacteria, yeasts, moulds, or other form of life, especially pathogenic or noxious organisms (CHEBI: 50694) (Fig. 6H).

MmpL5 (964 residue long protein) Set-A **(**Table 4**)** residues bind with the RJA (Ratjadone A). It shows growth inhibitory effect against bacteria, yeast and human cancer cells (Fig. 6-I) (Heo et al., 2014). Set-B **(**Table

4**)** residues bind with the BCT (bicarbonate). The bicarbonate ion (hydrogen carbonate ion) is an anion with

empirical formula HCO_3

It has antimicrobial properties (Figure: 6-J).

MmpL6 (397 residue long protein) Set-A **(**Table 4**)** residues bind with the URE (UREA). The urea and thiourea derivatives (7a-7k, 8a-8f) were screened for their antibacterial activity against some of the pathogenic bacteria (Fig. 6-K) (Miner et al., 2009). Set-B **(**Table 4**)** residues show the bonding relationship with human neutrophil peptides (HNPs) has antimicrobial activity (Fig. 6-L).

MmpL7 (920 residues long protein) Set-A **(**Table 4**)** residues bind with the FES [2FE-2S]-CLUSTER (Fig. 6-M) (Cosper et al., 2009). Set-B **(**Table 4**)** residues bind with the MIY (MINOCYCLINE) which is biostatic towards bacteria, yeasts, moulds, or other form of life, especially pathogenic or noxious organisms (CHEBI: 50694) (Fig. 6-N).

MmpL8 (1089 residues long protein) Set-A **(**Table 4**)** residues bind with the LMT (Dodecyl β-D-maltoside) shows an antimicrobial effects through a surface action (Fig. 6-O). Set-B **(**Table 4**)** residues bind with the C8E

(tetraethylene glycol monooctyl ether). Involved in translocation of long-chain fatty acids across the outer membrane. It is a receptor for the bacteriophage T2. FadL may form a specific channel. Involved in the active translocation of vitamin B12 (cyanocobalamin) across the outer membrane to the periplasmic space (Figure: 6- P).

MmpL9 (962 residues long protein) Set-A **(**Table 4**)** residues bind with the ZLD (**Linezolid**). Linezolid is a synthetic antibiotic, against Gram-positive organisms tested in the pharmacodynamics model and reduce colonization density (Fig. 6-Q) (Burzlaff et al., 2003)(Reddy et al., 2011). Set-B **(**Table 4**)** residues bind with the LMT (Dodecyl β-D-maltoside) (Fig. 6-R).

MmpL10 (1002 residues long protein) Set-A **(**Table 4**)** residues bind with the ZLD (**Linezolid)** (Fig. 6-S). Set-B **(**Table 4**)** residues bind with the DM2 (DOXORUBICIN). Doxorubicin or Adriamycin also has antibiotic activity, providing a simple resistance assay for screening of environmental bacterial (Wiederhold, 2005). There are various research that have proved the antimicrobial activity of Doxorubicin (Fig. 6-T) (Sader et al., 2001) (Westman et al., 2012).

MmpL11 (966 residues long protein) Set-A **(**Table 4**)** residues show the bonding relationship with human neutrophil peptides (HNPs) (Fig. 6-U). Set-B **(**Table 4**)** residues bind with the LBF (Leptomycin B an unsaturated, branched-chain fatty acid, and is an important tool in the study of nuclear export. It show an antifungal, antibiotic and anti-tumour activity (Fig. 6-V) (Di Marco et al., 1975).

MmpL12 (1146 residues long protein) Set-A (Table 4) residues bind with ERY (Erythromycin A). Erythromycin A is a potent antibiotic long-recognized as a therapeutic option for bacterial infections (Gumpert et al., 1982). Erythromycin acts by inhibiting elongation at the transpeptidation step, specifically aminoacyl translocation from the A-site to P-site (Fig. 6-W). Set-B **(**Table 4**)** residues bind with LMT (Dodecyl β- maltoside (Fig. 6-X).

MmpL13a (303 residues long protein) Set-A **(**Table 4**)** residues bind with BCL (BACTERIOCHLOROPHYLLA). Bacteriochlorophylls are photosynthetic pigments that occur in various phototrophic bacteria (Fig. 6-Y). Set-B **(**Table 4**)** residues bind with VDY (CALCIDIOL) the major circulating metabolite of vitamin D3 (cholecalciferol). It is formed in the liver and is the best marker of the body's vitamin D stores. It is efficient in the cure of rickets and osteomalacia, both in azotemic and non- azotemic patients. Calcifediol also has mineralizing properties (Fig. 6-Z).

MmpL13b (470 residues long protein) Set-A (Table 4) residues bind with CVM (CYMAL-4) *C*22 *H*40 *O*11

Volume : 5 | Issue : 1 | JANUARY 2016 • ISSN No 2277 - 8179

(Fig. 6-A"). Set-B (Table 4) residues bind with the LMT (**Dodecyl β-D-maltoside** (Fig. 6-B").

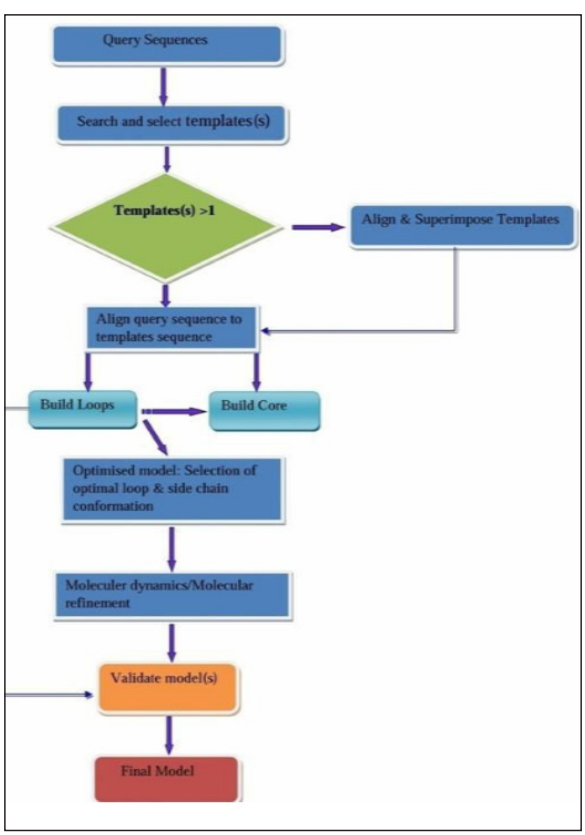
Based on these result, we successfully annotated the functions of 13 protein of MmpL family, which can be further used as lead for designing experimental approach geared towards evolution of exact function of the gene. Mostly proteins are predicted as drug-transmembrane transporter, ligand-gated ion channel activity, lipid homeostasis, ER-nucleus signalling pathway, molting cycle and many more. MmpL protein family carry out a miscellaneous variety of functions and we used as main drug target of pharmaceutical agent, we predict the functional sites on the protein of MmpL family as well as their possible ligands too. These MmpL families may serve as potential diagnostic target; and may be a milestone in the emerging in field of drug discovery development.

CONCLUSION

In silico analysis of MmpL protein family of *Mycobacterium tuberculosis* described here provides a simple & precise method to predicting 3D structure and assigning multi-functions to proteins of MmpL family. Our study facilitates a rapid identification of binding sites in these proteins and also predicted possible ligands for MmpL protein family which is a potential therapeutic target. We expect that this protein family may play a significant role in host-pathogen interaction and in new anti-TB drugs development.

ACKNOWLEDGEMENTS

MZM and SA are financially supported by Indian Council of Medical Research, New Delhi, India under SRF (Senior Research Fellowship)



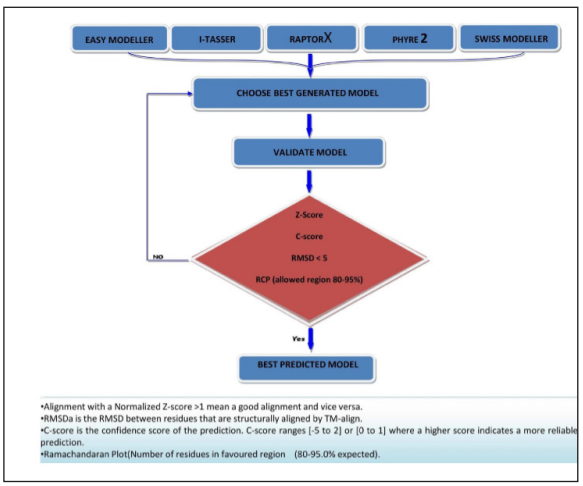


Fig: 3

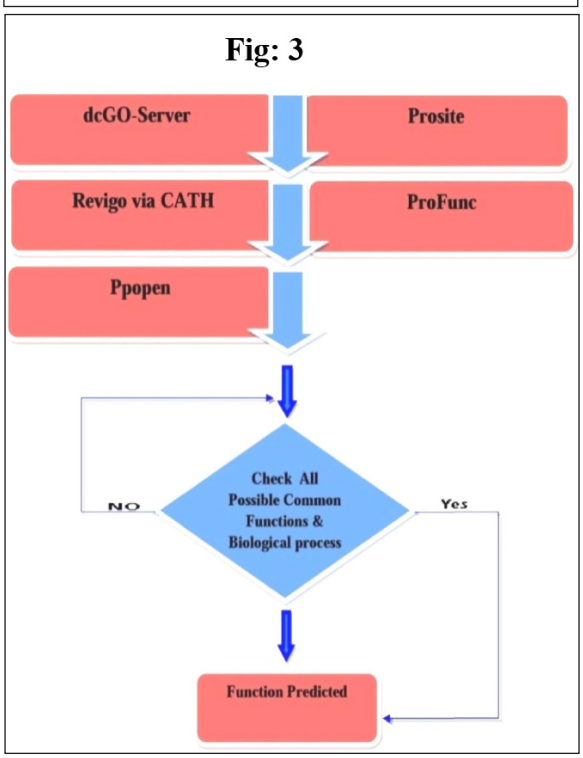


Fig: 4a

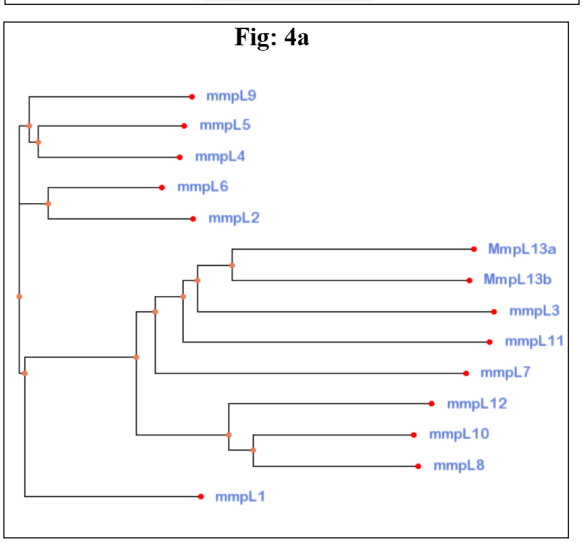
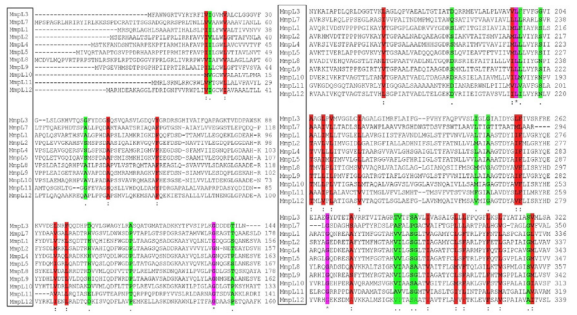
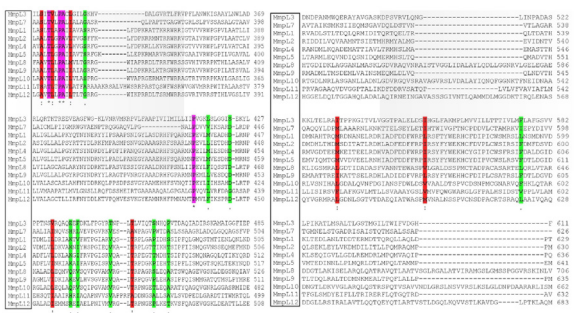
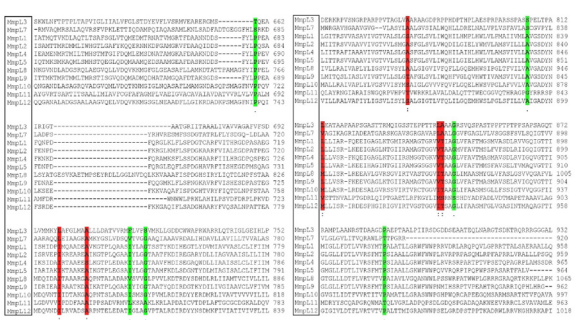


Fig: 4b







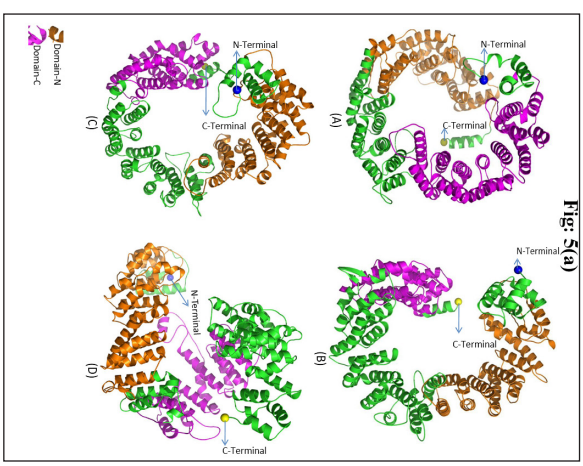

Domain-N
Domain-C
N-Terminal
C-Terminal
(C)
(A)
(D)
(B)

Fig: 5(a)

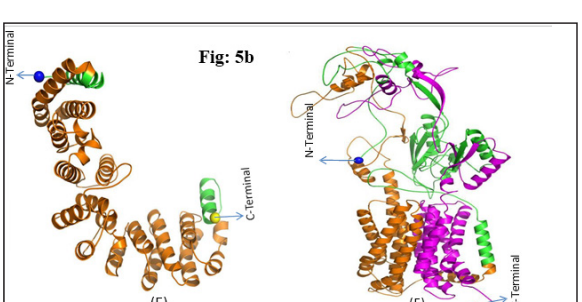
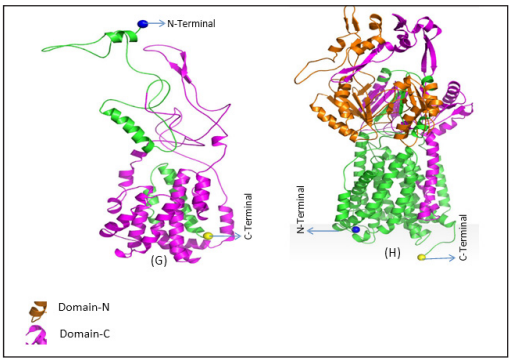
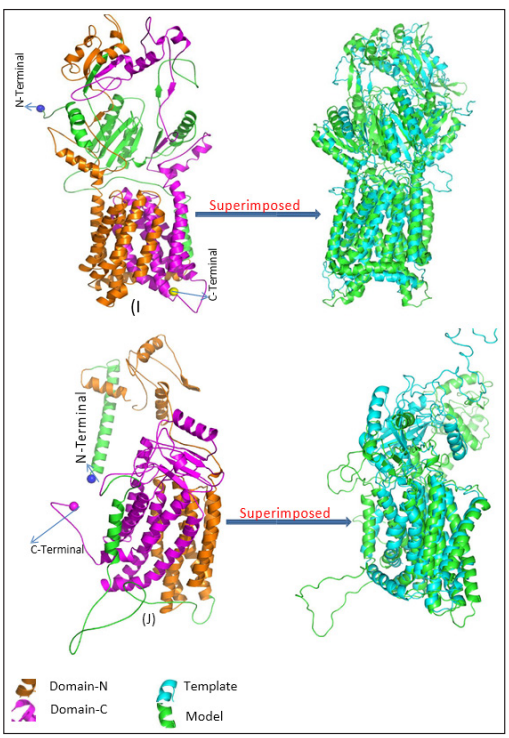


Fig: 5b
N-Terminal
C-Terminal
(E)
(F)


N-Terminal
C-Terminal
(G)
(H)
Domain-N
Domain-C


N-Terminal
Superimposed
C-Terminal
(I)
(J)
Domain-N
Domain-C
Template
Model

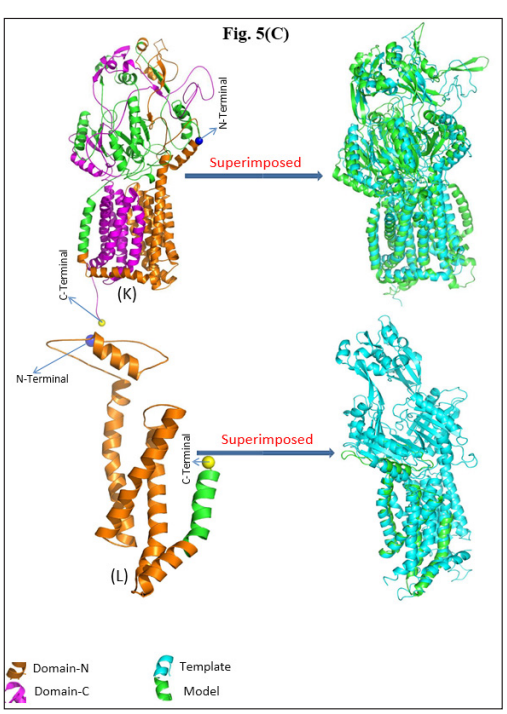
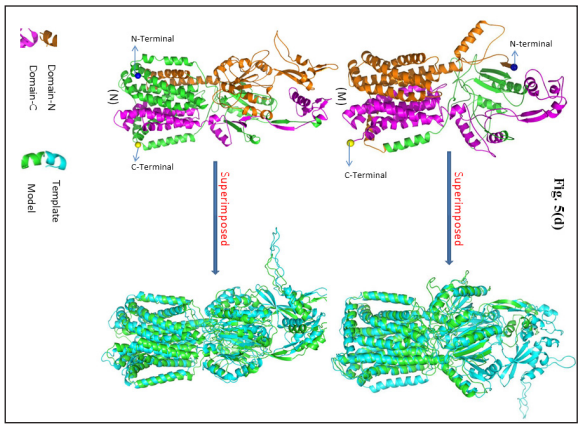
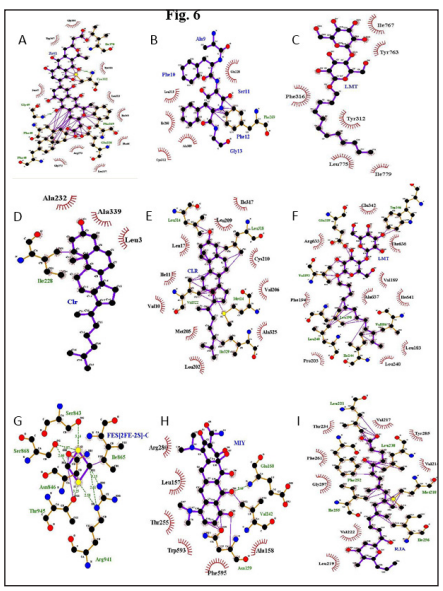
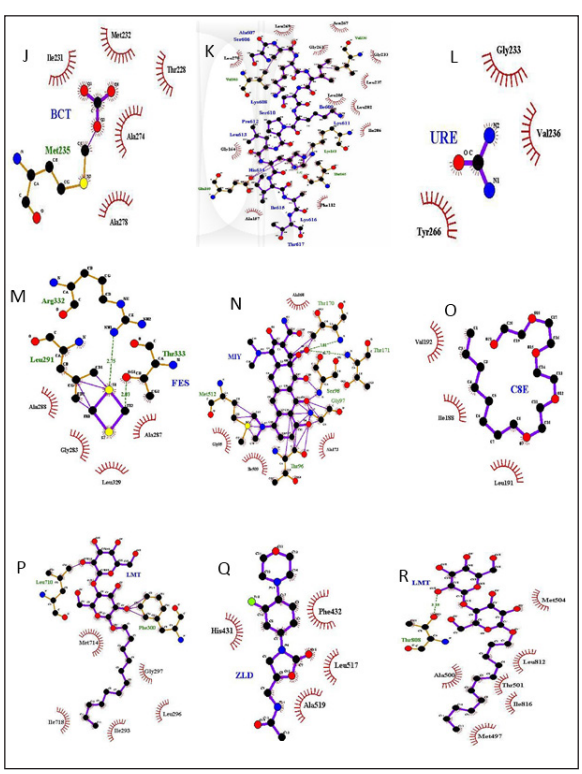


Fig. 5(C)
N-Terminal
Superimposed
C-Terminal
(K)
(L)
Domain-N
Domain-C
Template
Model

Volume : 5 | Issue : 1 | JANUARY 2016 • ISSN No 2277 - 8179







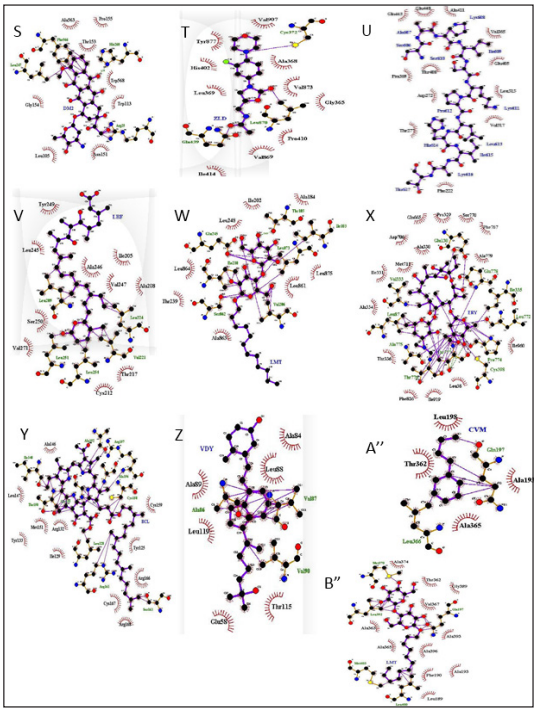


Table 1 Physiochemical properties of MmpL family proteins in Mycobacterium tuberculosis

S.NO.	GeneName	UniprotID.	Rvnumber	Molecular mass(Da)	Isoelectric point.	Grand mean of hydropathicity (Gravy)	Protein length (aa)	Locatio n(s)	Functionalcategory.	Function
1	*MmpL1*	P9WJY9	Rv0402c	104806	7.1217	2877	958	480.333	cell wall and cellprocesses.	Unknown. Thought to be involved inDrugtransmembrane.
2	*MmpL2*	P9WJY7	Rv0507	106202	7.1355	2907	968	297.199	cell wall and cellprocesses.	Unknown. Thought to be involved inDrugtransmembrane.
3	*MmpL3*	P9WJY5	Rv0206c	100892	9.9071	2833	944	244.454	cell wall and cellprocesses.	Unknown. Thought to be involved inDrugtransmembrane.
4	*MmpL4*	P9WJY3	Rv0450c	103202	7.1179	2904	967	238.588	cell wall and cellprocesses.	Unknown. Thought to be involved inDrugtransmembrane.
5	*MmpL5*	P9WJY1	Rv0676c	104785	5.9272	2895	964	172.586	cell wall and cellprocesses.	Unknown. Thought to be involved inDrugtransmembrane.
6	*MmpL6*	P9WJY8	Rv1557	42389.5	6.5701	1194	397	1761.74	cell wall and cellprocesses.	Unknown. Thought to be involved inDrugtransmembrane.
7	*MmpL7*	P9WJY7	Rv2942	97121.9	9.0903	2763	920	3285.07	cell wall and cellprocesses.	Involved in translocation ofphthioceroldimycocerosate (dim) in
8	*MmpL8*	P9WJY5	Rv3823c	115996	9.6476	3270	1089	4288.26	cell wall and cell processes	Thought to be involved in the transportoflipids, it has been shown to be requiredfor production of a sulfatedglycolipid.
9	*MmpL9*	P9WJY3	Rv2339	104615	7.1687	2859	952	2614.69	cell wall and cellprocesses.	Unknown. Thought to be involved inDrugtransmembrane.
10	*MmpL10*	P9WJY1	Rv1183	106383	8.9457	3009	1002	1321.22	cell wall and cellprocesses.	Unknown. Thought to be involved inDrugtransmembrane.
11	*MmpL11*	P9WJY9	Rv0202c	102502	9.4881	2901	966	238.392	cell wall and cellprocesses.	Unknown. Thought to be involved inDrugtransmembrane.
12	*MmpL12*	P9WJY6	Rv1222c	122598	8.5126	3441	1146	1714.17	cell wall and cellprocesses.	Unknown. Thought to be involved inDrugtransmembrane.
13	*MmpL13a*	O06545	Rv1143	32545.8	8.5722	912	303	1272.42	cell wall and cellprocesses.	Unknown. Thought to be involved inDrugtransmembrane.
14	*MmpL13b*	O06546	Rv1146	50027.4	10.8818	1413	470	1273.35	cell wall and cellprocesses.	Unknown. Thought to be involved inDrugtransmembrane.

Table 2 Multiple functions of MmpL family proteins in Mycobacterium tuberculosis

MmpL1	. Inorganic cation, hydrogen ion transmembrane transporter activity. . Oxidoreductase activity, acting on the CH-OH groups of donors, NAD or NADP as accepter. . Drug transmembrane transporter. . Hydroxyl methyl glutaryl-CoA-reductase	. Endoplasmic Reticulum nucleus signalling pathway. . Lipid homeostasis. . Sulfolipid biosynthetic process. . Negative regulation of multicellular organismol process and positive regulation of transcription from RNA-Polymerase II promoter. . Detoxification of copper ion.

MmpL2	. Lipoprotein particle receptor activity. . Copper ion, carbohydrate and ADP binding. . Sterol response element binding. . Silver ion transmembrane transporter activity. . Ligand-gated ion channel activity. . Ubiquitin-Protein trasferase activity. Sequence-specific DNA binding transporter . factor activity.	. Positive regulation of ERK1 & ERK2 cascade. . Growth of symbiont in host cell. . Kertino cycle proliferation. . RasGTPas activator activity. . Regulationof Protein Phosphorylation. . Mitochondrian degradation by induced vacuole formation. . Pharnyl pumping. . Cell Proliferation growth.
MmpL3	. Protein binding. . Ca-ion binding. . Sequence-specific DNA binding transcription factor activity. . Inorganic cation transmembrane transporter activity. . Drug transporter transmembrane. . Copper ion exporter. It is involved in mycolic acid . transport in mycobacteria[09].	. Lipid transporter. . Growth of symbiont in host cell. . Sulfolipid biosynthetic process. . Target of the antituber-culer pyrrole derivatives BM212. . Molting cycle. . Lipid homeostasis. . Important for mycobacterial growth and hold the participation of this transporter in the translocation of trehalosemonomycolate across the plasma membrane[10].
MmpL4	. Growth in symbiont in host cell. . Drug transmembrane transport. . Active erasin of host immune response. . 0[6]. . Oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as accepter. . Detoxification of silver ion. . Hydroxymethylglutaryl-coA reductase activity . Lipoprotein particle receptor activity. . Copper ion and Drug binding. . Ag^+ transmembrane transporter activity.	. ER-nucleus signalling pathway. . Lipid homeostasis. . Drug transmembrane transport. . Antibacterium type cell wall biogenesis. . Mammary gland epithelial cell differentiation. . Kertinocycle proliferation. . Lipoprotein Metabolism. . Positive regulation of ERK1 & ERK2 cascade. . Cellular copper ion homeostasis.
MmpL5	. Ion transmembrane transporter activity. . Hydrogen ion transmembrane transporter activity. . Respond to organic substance. . Drug-transmembrane transporter. . Lipoprotein particle receptor activity. . Receptor & Transporter activity. . Copper ion binding.	. Lipid homeostasis. . ER-nucleus signalling pathway. . Mammary gland epithelial cell differentiation. . Kertinocycle proliferation. . Actinobcterium type cell wall biogenesis.
MmpL6	. Hydrogen ion transmembrane transporter activity. . Oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as accepter. . Respond to organic substance. . Lipid and Drug-transmembrane transportation. . Hydroxymethylglutaryl-coA reductase activity . Unfolded protein binding and Alcohol binding. . Silver ion transmembrane transporter activity.	. ER-nucleus signalling pathway. . Growth of symbiont in host cell. . Kertinocycle proliferation. . Lipoprotein Metabolism. . Actinobcterium type cell wall biogenesis. . Mammary gland epithelial cell differentiation. . Lipid homeostasis. . Sulfolipid biosynthetic process.
MmpL7	. Inorganic cation transmembrane transport activity. . Signal transducer activity. Molting cycle. . Negative regulation of growth. . Negative regulation of growth of multicellular organismol process. . Positive regulation of Transcription from RNA polymerase II promoter. . MmpL7 is essential for virulence, presumably because it transports polypeptide virulence factors such as phthioceroldimycocerosate (PDIM) to the cell wall[11].and reduced the growth kinetics and lethality[6]. . ATPase activity, coupled to transmembrane movement of substance.	. Steroid metabolic process. . ER-nucleus signalling pathway. . Growth of symbiont in host cell. . Substrate adhesion dependent cell spreading. . Kertinocycle proliferation. Lipoprotein Metabolism. Lipoprotein Metabolism. . Transporter of the RND family[10]. . MmpL7 is a potential substrate for kinase PknD[23]. . Responsible for isoniazid efflux in Mycobacterium smegmatis[11].

MmpL8	· Alcohol binding. · Hydrogen ion transmembrane transporter activity. · Oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as accepter. · Lipid homeostasis and transport. · Heme-binding. · ATPase activity, coupled to transmembrane movement of substance. · Glycosphigolipid binding. · Receptor Activity.	· Type IV pilus dependent moyility. · Pilus-assembly. · Substrate adhesion dependent cell spreading. · Kertinocycle proliferation. · Lipoprotein Metabolism. · Muscle, organ development. · Its play a vital role in sulfolipid-1 biosynthesis and Mycobacterium tuberculosis virulence[12].
	· Extra cellular matrix structure constituents. · Signal transducer activity.	
MmpL9	· Alcohol binding. · Hydrogen ion transmembrane transporter activity. · Oxireductase activity, acting on the CH_OH group of donor, NAD or NADP as accepter. · Lipid homeostasis and transport. Drug transmembrane transport. · Molting cycle, collagen &cuticulin based cuticle. · Lysosomal transport. · Protein homodimerization activity. · Hydroxymethylglutaryl-coA reductase activity.	· Positive regulation of ERK1 & ERK2 cascade. · Growth of symbiont in host cell. Kertinocycle proliferation. · RasGTPas activator activity. Kertinocycle proliferation. Lipoprotein Metabolism. · ER-nucleus signalling pathway. · Cholesterol homeostasis.
MmpL10	· Oxireductase activity. · Lytic envelop. · Nuclear envelop. · Cellular copper ion homeostasis. · Response to silver ion. · NADPH and Heparin binding. · Zinc ion, B-amyloid and actin binding. · Protein-kinase inhibitor activity. · Chromatin binding. · Anti-oxidant activity.	· Perinuclear region of cytoplasm. · ER nucleus signalling pathway. · Response to host immune response. · Kertinocycle proliferation. · Lipoprotein Metabolism. · Sulfolipid biosynthetic process.
MmpL11	· Inorganic cation transmembrane transport activity. · Alcohol binding. Oxidoreductase activity · Response to host immune response. · Keratinocyte proliferation. · Lipoprotein metabolism. Hydroxymethylglutaryl-coA reductase activity. · Sterol sensor activity.	· ER nucleus signalling pathway. · Kertinocycle proliferation. · Lipoprotein Metabolism. Sulfolipid biosynthetic process · Positive regulation of ERK1 & ERK2 cascade. · Actinobacterium type cell wall biogenesis. · Lipid homeostasis. Lipid transfer
MmpL12	· Alcohol binding. · Hydrogen ion transmembrane transporter activity. · Copper ion and silver ion transmembrane transporter activity. · Drug-hydrogen antiporter activity. · Anti-oxidant activity. · Structural molecule activity. · Beta-amyloid binding. · Actin, Ca^{+} binding.	· Molting cycle. · Positive regulation of Transcription from RNA polymerase II promoter. · Steroid metabolic process. · Lipid homeostasis. · Sulfolipid biosynthetic process. · Oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as accepter. · Sulfolipid biosynthetic process.
MmpL13a	· Alcohol binding. · H+ transmembrane transporter activating. · Oxidoreductase activity. · Active erasin of host immune response. · Ag+ transmembrane transporter activity. · Hydroxymethylglutaryl-coA reductase activity · Drug-binding. · Lipoprotein particle receptor activity.	· Positive regulation of ERK1 & ERK2 cascade. · Actinobacterium type cell wall biogenesis. · Kertinocycle proliferation. · Autophasy. · Growth of symbiont in host cell. · ER-nucleus signalling pathway. · Lipid homeostasis.
MmpL13b	· Alcohol binding. · Oxidoreductase & transporter activity. · Organic substance transporter. · Response to host immune response. · Unfolded protein binding. · Drug and receptor binding. · Hydroxymethylglutaryl-coA reductase activity. · Ag+ transmembrane transporter activity.	· Lipoprotein metabolism. · Positive regulation of ERK1 & ERK2 cascade. · Actinobacterium type cell wall biogenesis. · Kertinocycle proliferation. · Growth of symbiont in host cell. · Lipid transport.

Table 3 This table shows the best templates, RMSD-value and percentage of residue in allowed region used by various tools to generate 3D structure.

Tools/ Server	Protein	Best Templates	RMSD	Residues in Core region (%)
I-TESSAR	MmpL1	4k0eA, 3ne5A	1.75, 4.8	84.7
RaptorX	MmpL2	3k07A	3.33	92.7
RaptorX	MmpL3	3aqpA	2.91	92.3
Easy Modeller	MmpL4	3w9i, 2hrt, 4dx5	2.94	92.8
Swiss Model	MmpL5	2v50B, 3w9jB, 3w9jC	3.33	88.4
I-TESSAR	MmpL6	4c0oA, 2x1gG	1.30, 3.04	80.
I-TESSAR	MmpL7	3ne5A, 4k0eA	3.32, 1.55	81.3
RaptorX	MmpL8	3k07A	3.30	89.4
I-TESSAR	MmpL9	3ne5A ,4c0oA,	3.44, 1.66	80.9
Easy Modeller	MmpL10	2j8s, 2zy9,3w9i	2.96	93.5
I-TESSAR	MmpL11	1u6gC, 3a6pF	3.53, 2.59	81.3
Easy Modeller	MmpL12	3aqp, 3ne5,3w9i	3.23	92.9
Swiss model	MmpL13 a	4dntC, 2v50A,2v50B	2.51	95.8
Easy Modeller	MmpL13 b	3aqp, 3ne5, 3w9j	2.77	93.2

Table 4 This table shows the binding residue position and possible Ligands in two groups Set A (red) and Set B (green).

Protein	Possible Ligands	Residues
MmpL1	HNPs, ZCT	9,11,12,13,265,268,269,273,312,315,316 | 48,49,227,228
MmpL2	CLR, LMT	228,232,339,338,345 | 312,316,763,764,767,768,771,775
MmpL3	CLR, LMT	206,210,213,314,318,324 |244,248,632,633,636,637,640,644
MmpL4	FES, MIY	841,842,843,847,866,867,868,869|157,158,159,160,242,244,255,593,595
MmpL5	RJA, BCT	217,218,221,231,234,235,249,252,259,264|275,277
MmpL6	HNPs, URE	156,157,159,160,163,164,205,233,236,237,263,266,269,270,273,300,30 |229,230,258,266,267
MmpL7	FES, MIY	288,290,292,329,332,333 | 96,168,171,173,509,512
MmpL8	C8E, LMT	188,191,192 | 293,296,297,300,706,707,710,711,714,718
MmpL9	ZLD, LMT	431,432 | 500,504,804,805,808,809,812,816
MmpL10	ZLD, DM2	872,873 | 27,70,107,152,154,155,247,563,566
MmpL11	HNPs, LBF	269,270,272,273,276,277,280,317,365,368,369,405,408,411,412,415,440 ,443 |198,202,205,217,220,240,246,247,250,254
MmpL12	ERY, LMT	75,87,130,131,133,292,663,665,705,764,787,928|186,206,209,239,24886 1.863
MmpL13 a	BCL, VDY	128,129,132,133 |115,119,122
MmpL13 b	CVM, LMT	389,392 |189,194,391,392,395,396,400,404

Figure Caption

Fig: 1 Basic steps involved in homology Modelling.

Fig: 2 Diagram show the steps involved to choose the best built model generated by various tools

Fig: 3Diagram show the protocol of the selection of the common appropriate function and biological process are predicted from above tools. All the possible functions and biological process are shown in next page in the tabular form.

Fig: 4(a) Cladogram of MmpL family, Tree is split into three clusters has closely related to sequences, while MmpL1 just behave like an out group and ancient protein which divided into to give rise to cluster 1(MmpL4, 5 and 9), cluster 2 (MmpL6 and 2) and clusters 3 (MmpL3, 7, 8, 10, 11, 12, and 13).

Fig: 4(b) This is the MSA of proteins of MmpL family. Fully conserved regions are shown in magenta colour indicated with '*'. Colon (:) indicates conservation between groups of strongly similar properties are shown in red colour and in green colour a dot (.) indicates conservation between groups of weakly similar properties.

Fig: 5 a The Structure A, B, C and D respectively Represent the models of MmpL 1, 7, 9 & 11 as obtain from I-Tasser. N and C terminals shown in blue and yellow sphere respectively. It explicitly depicts the domains N and ranging from (35-279) and (645-1008) respectively with different color i.e. Orange & magenta.

Fig: 5 bThe Structure E, F, G and H respectively Represent the models of MmpL 6, 10, 13b & 12 respectively. Structure 10, 12 and 13b obtain from easy modeler while 6 is obtain from I-Tasser. The structure E and G have single domain i.e. domain-N and domain-C respectively.

Fig: 5c The Structure I, J, K and L respectively Represent the models of MmpL 2, 3, 8 & 13a respectively. Model 2, 3 and 8 as obtain from RaptorX and model 13b is built by Swiss model. These all models are superimposed with their template.

Fig: 5d The Structure M and N respectively Represent the models of MmpL 5 & 4 respectively. Model 4 obtain from easy modeler and model 5 is built by Swiss model. These all models are superimposed with their template.

Fig: 6 It outlines the interacting residues of MmpL's two sets (Highest score) with different ligands in giving a formal understanding of the molecular mechanism. MmpL1 interacting with human neutrophil peptide and ZCT ligands in fig. 6A & B respectively; MmpL2 and MmpL3 (fig. 6C & E and 6D & F) interacting with CLR and LMT respectively; Mmpl4 interacting with FES and MIY in fig: 6G & H respectively. MmpL5 interacting with RJA and BCT in fig: 6I& J respectively. MmpL6 interacting with URE and HNPs in fig: 6K & L respectively. MmpL7 interacting with FES and MIY in fig: 6M & N respectively. MmpL8 interacting with LMT and C8E in fig: 6O & P respectively. MmpL9 interacting with ZLD and LMT in fig: 6Q & R respectively. MmpL10 interacting with ZLD and DM2 in fig: 6S & T respectively. MmpL11 interacting with HNPs and LBF in fig: 6U & V respectively. MmpL12 interacting with ERY and LMT in fig: 6W & X respectively. MmpL13a interacting with BCL and VDY in fig: 6Y & Z respectively. MmpL13b interacting with CVM and LMT in fig: 6A" & B" respectively.

Volume : 5 | Issue : 1 | JANUARY 2016 • ISSN No 2277 - 8179

REFERENCE

Ahmad, S., 2011. Pathogenesis, Immunology, and Diagnosis of Latent Mycobacterium tuberculosis Infection. Clin. Dev. Immunol. 2011, 1–17. doi:10.1155/2011/814943 Ali, S., Malik, M.Z., Alam, M.J., Ishrat, R., Singh, R.K.B., 2014. Evolutionary Trace Analysis of p53 Protein: A Statistical Analysis of Conserved Amino Acids in p53 Protein. J. Bioinforma. Intell. Control 3, 194– 201. doi:10.1166/jbic.2014.1087 Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T.G., Bertoni, M., Bordoli, L., Schwede, T., 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res. 42, W252–258. doi:10.1093/nar/gku340 Burzlaff, A., Kalesse, M., Kasper, C., Scheper, T., 2003. Multi parameter in vitro testing of ratjadone using flow cytometry. Appl. Microbiol. Biotechnol. 62, 174–179. doi:10.1007/s00253-003-1300-0 Converse, S.E., Mougous, J.D., Leavell, M.D., Leary, J.A., Bertozzi, C.R., Cox, J.S., 2003. MmpL8 is required for sulfolipid-1 biosynthesis and Mycobacterium tuberculosis virulence. Proc. Natl. Acad. Sci. 100, 6121–6126. doi:10.1073/pnas.1030024100 Cosper, N.J., Eby, D.M., Kounosu, A., Kurosawa, N., Neidle, E.L., Kurtz, D.M., Iwasaki, T., Scott, R.A., 2009. Redox-dependent structural changes in archaeal and bacterial Rieske-type [2Fe-2S] clusters. Protein Sci. 11, 2969–2973. doi:10.1110/ps.0222402 Di Marco, A., Arcamone, F., Zunino, F., 1975. Daunomycin (Daunorubicin) and Adriamycin and Structural Analogues: Biological Activity and Mechanism of Action, in: Corcoran, J.W., Hahn, F.E., Snell, J.F., Arora, K.L. (Eds.), Mechanism of Action of Antimicrobial and Antitumor Agents. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 101–128. Fang, H., Gough, J., 2013. dcGO: database of domain-centric ontologies on functions, phenotypes, diseases and more. Nucleic Acids Res. 41, D536–D544. doi:10.1093/nar/gks1080 Fiser, A., Do, R.K., Sali, A., 2000. Modeling of loops in protein structures. Protein Sci. Publ. Protein Soc. 9, 1753–1773. doi:10.1110/ps.9.9.1753 Gromiha, M.M., Yabuki, Y., Suresh, M.X., Thangakani, A.M., Suwa, M., Fukui, K., 2009. TMFunction: database for functional residues in membrane proteins. Nucleic Acids Res. 37, D201–D204. doi:10.1093/nar/gkn672 Gumpert, J., Dornberger, K., Smith, T.H., 1982. Antimicrobial activities of daunorubicin and adriamycin derivatives on bacterial and protoplast type L-form cells of Bacillus subtilis 170, Escherichia coli B, and Proteus mirabilis VI. Structure--activity relationship. Z. Für Allg. Mikrobiol. 22, 687–692. Heo, L., Shin, W.-H., Lee, M.S., Seok, C., 2014. GalaxySite: ligand-binding-site prediction by using molecular docking. Nucleic Acids Res. 42, W210–W214. doi:10.1093/nar/gku321 Kaczanowski, S., Zielenkiewicz, P., 2010. Why similar protein sequences encode similar three-dimensional structures? Theor. Chem. Acc. 125, 643–650. doi:10.1007/s00214-009-0656-3 Källberg, M., Margaryan, G., Wang, S., Ma, J., Xu, J., 2014. RaptorX server: a resource for template-based protein structure modeling. Methods Mol. Biol. Clifton NJ 1137, 17–27. doi:10.1007/978-1-4939- 0366-5_2 Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., Sternberg, M.J.E., 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protoc. 10, 845–858. doi:10.1038/nprot.2015.053 Khan, F.I., Govender, A., Permaul, K., Singh, S., Bisetty, K., 2015. Thermostable chitinase II from Thermomyces lanuginosus SSBP: Cloning, structure prediction and molecular dynamics simulations. J. Theor. Biol. 374, 107–114. doi:10.1016/j.jtbi.2015.03.035 Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26, 283–291. doi:10.1107/S0021889892009944 Laskowski, R.A., Swindells, M.B., 2011. LigPlot+: Multiple Ligand–Protein Interaction Diagrams for Drug Discovery. J. Chem. Inf. Model. 51, 2778–2786. doi:10.1021/ci200227u Laskowski, R.A., Watson, J.D., Thornton, J.M., 2005. ProFunc: a server for predicting protein function from 3D structure. Nucleic Acids Res. 33, W89–W93. doi:10.1093/nar/gki414 Lubec, G., Afjehi-Sadat, L., Yang, J.-W., John, J.P.P., 2005. Searching for hypothetical proteins: theory and practice based upon original data and literature. Prog. Neurobiol. 77, 90–127. doi:10.1016/j.pneurobio.2005.10.001 Miner, M.D., Chang, J.C., Pandey, A.K., Sassetti, C.M., Sherman, D.R., 2009. Role of cholesterol in Mycobacterium tuberculosis infection. Indian J. Exp. Biol. 47, 407–411. Pasca, M.R., Guglierame, P., De Rossi, E., Zara, F., Riccardi, G., 2005. mmpL7 gene of Mycobacterium tuberculosis is responsible for isoniazid efflux in Mycobacterium smegmatis. Antimicrob. Agents Chemother. 49, 4775–4777. doi:10.1128/AAC.49.11.4775-4777.2005 Reddy, N.S., Rao, A.S., Chari, M.A., Kumar, V.R., Jyothi, V., Himabindu, V., 2011. Synthesis and Antibacterial Activity of Urea and hiourea Derivatives at C-8 Alkyl Chain of Anacardic Acid Mixture Isolated from a Natural Product Cashew Nut Shell Liquid (CNSL). Int. J. Org. Chem. 01, 167–175. doi:10.4236/ijoc.2011.14025 Rost, B., Yachdav, G., Liu, J., 2004. The PredictProtein server. Nucleic Acids Res. 32, W321–W326. doi:10.1093/nar/gkh377 Sader, H.S., Gales, A.C., Jones, R.N., 2001. Antimicrobial activity of linezolid against Gram-positive cocci isolated in Brazil. Braz. J. Infect. Dis. Off. Publ. Braz. Soc. Infect. Dis. 5, 171–176. Sali, A., Blundell, T.L., 1993. Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. 234, 779–815. doi:10.1006/jmbi.1993.1626 Sigrist, C.J.A., Cerutti, L., de Castro, E., Langendijk-Genevaux, P.S., Bulliard, V., Bairoch, A., Hulo, N., 2010. PROSITE, a protein domain database for functional characterization and annotation. Nucleic Acids Res. 38, D161–D166. doi:10.1093/nar/gkp885 Stephens, D.E., Khan, F.I., Singh, P., Bisetty, K., Singh, S., Permaul, K., 2014. Creation of thermostable and alkaline stable xylanase variants by DNA shuffling. J. Biotechnol. 187, 139–146. doi:10.1016/j.jbiotec.2014.07.446 Westman, E.L., Canova, M.J., Radhi, I.J., Koteva, K., Kireeva, I., Waglechner, N., Wright, G.D., 2012. Bacterial Inactivation of the Anticancer Drug Doxorubicin. Chem. Biol. 19, 1255–1264. doi:10.1016/j.chembiol.2012.08.011 Wiederhold, N.P., 2005. Antibacterial activity of linezolid and vancomycin in an in vitro pharmacodynamic model of Gram-positive catheter-related bacteraemia. J. Antimicrob. Chemother. 55, 792–795. doi:10.1093/jac/dki106 Yang, J., Zhang, Y., 2015. I-TASSER server: new development for protein structure and function predictions. Nucleic Acids Res. 43, W174–181. doi:10.1093/nar/gkv342