IN-SILICO STRUCTURE PREDICTION, ENERGY CALCULATION AND PHYLOGENETIC ANALYSIS OF MEMBRANE PROTEIN LARGE-12 (MMPL-12) OF INFECTIOUS AGENT OF HUMAN TUBERCULOSIS (*MYCOBACTERIUM TUBERCULOSIS*)

Aishwarya Tiwari[1], Anjli Katiyar[1], Swati Srivastava[1], Vijay Laxmi Saxena[1,2] and Nirupama Agarwal[3]

[1]Bioinformatics Infrastructure facility canter, D.G.P.G College, Kanpur, India.
[2]Asutosh Mookerjee Fellow of ISCA, Department of Zoology, D.G.P.G College, Kanpur, India
[3]Asutosh Mookerjee Fellow of ISCA, Department of Zoology, University of Lucknow, Lucknow, India
[*]Corresponding author E-mail: Vlaxmisaxena@gmail.com.

ABSTRACT

Mycobacterium tuberculosis is an infectious agent of human tuberculosis whose genome revealed 12 membrane proteins of MMPL family. Accurate description of amino acids of MMPL-12 is very important in prediction of structure and calculation of energy which is done with the help of Bioinformatics tools and software approach. Ramachandran plot of the ö, ø values for the amino acids in a MMPL-12 protein is further done. Using this computational approach, the total energy of MMPL-12 in Mycobacterium tuberculosis is found to be 38098.148KJ/mol. Repeating energy trends at each of the molecular, functional group, and atomic levels is also observed. Retinols and retinoic acid play essential characters in variation of gene expression and overall growth of embryo in *Mycobacterium tuberculosis*.

Keywords: Retinol; MMPL-12; Adipose tissue; P value; Computational energy

INTRODUCTION

The well-known disease Tuberculosis, commonly known as TB is caused by the bacteria *Mycobacterium tuberculosis* which usually attack the lungs, but they can also damage other parts of body. TB spreads through air when an infected person coughs, sneezes, or talks. With the spread of

environmental pollution and acquired immune deficiency syndromes (Zhu, Ou and Zheng, 2018), it is becoming a more serious problem. Previous studies have suggested role of complex MDR mechanisms to multiple genes and various mechanisms at the tissue, cellular, and molecular levels. As transcription regulating factors, noncoding RNAs have been found to regulate protein-coding gene expression at the transcriptional and post-transcriptional levels (Yan et al., 2018)**.** However, there are very few studies that assess the role of incRNAs in MDR-TB. In this study, expression profiles of both incRNAs and mRNAs in peripheral blood mononuclear cells (PBMCs) from patients with Multiple Drug-Resistant TB, patients with DS-TB, or healthy control were first detected to investigate whether incRNAs were involved in MDR-TB pathogenesis (Yan, Xu, Zhang, Wang, and Pang, 2018)**.** In gram-negative bacteria, RND proteins' function in association with accessory protein that include a periplasmic membrane fusion protein (MFP) and the channel-forming outer membrane factor (Sz»kely and Cole, 2016) were found. Biosynthesis and secretion of cell wall elements were mediated by a variety of different enzymes and proteins (Alam, Ali, Malik and Ishrat, 2016)**.** Mycobacterial membrane protein large 12 (MMPL-12) the genome of *M. tuberculosis* (Mtb), consists of 13 genes, designated as Mmpl the co-localization of Mmpl gene with the pks genes and genes responsible for lipid metabolism suggest there similar function (Chen, Arendall, Headd, 2010). The main reasons of TB getting out of control in many countries include the prevalence of human immunodeficiency viruses, multiple drug resistance, poverty, population growth, etc. Therefore, the early diagnosis and effective treatment of latent TB infection and TB are important. The established detection methods include sputum smear staining, TB-specific enzyme-linked immunospotassay (T-SPOT.TB), TB antibody (TB-Ab) test and mycobacterium TB deoxyribonucleic acid (TB-DNA) test, purified proton derivative (PPD) test, etc. (Zhu, Ou, Zheng, 2018).

The genome of Tuberculosis contains 13 genes that encode RND proteins, designated as MmpL12 (Mycobacterial membrane protein Large12). The RND co-localization of some of the Mmpl genes with genes involved in polyketide biosynthesis (post genes) and genes involved in lipid metabolism (papa, fadD) suggest a similar role of these proteins in complex lipid transport in *M. tuberculosis*. Indeed, the MmpL7 protein has been shown to be involved in transport of phthiocerol dimycocerosate (PDIM) and MmpL8 is involved in the synthesis of sulfolipid-1 by transporting a precursor of this molecule. Role in mediating *M. tuberculosis* drug resistance has also been suggested for the Mmpl proteins. In order to decipher the role of MmpL protein in *M. tuberculosis*, we have generated

mutant strains inactivated in 11 out of 12 of the mmpl genes. These strains were evaluated for altered susceptibility to antituberculosis drugs and in virulence studies using the marine model of infection. Unlike other bacteria, these proteins do not appear to play a major role in *M. tuberculosis* drug resistance. However, a role in virulence in mice was demonstrated for few of these proteins (MmpL4, MmpL8, and MmpL11) (Szekely and Cole, 2016). Mmpl proteins occur in both (slow and fast) growing stages for *Mycobacterium* species. Genes encoding Mmpl proteins were first identified in the genome of *M. tuberculosis* H37Rv.The genome of this strain encodes 13 putative Mmpl proteins-11(Szekely and Cole, 2016).

We have tried to explore the maximum information of Mmpl family by using updated bioinformatics tools and techniques including 3D- structure prediction (Chen *et al.,* 2010)

We describe the genome characteristics of the Colombian clinical isolate UT205, which was isolated from a patient with TB from Medellin, Antioquia. A comparison was carried out against the H37Rv reference genome.

MATERIAL AND METHODS

Collection of the data using PUBMED

Collecting the data for structure prediction of protein present in Mycobacterium tuberculosis of the sequence from NCBI of Mycobacterial membrane protein large

Raptor x (Fig. 2)**:** A wed-based method using Raptorx (http://raptorx.uchicago.edu/) for protein secondary structure prediction, template-based tertiary structure modelling, alignment quality assessment, and probabilistic alignment sampling. Raptor x web server delivers high-quality structural models for many targets with only remote templates. Because of this computational modelling of the three-dimensional atomic arrangement of the amino acid chain is also possible in determining the role of the protein in biological processes (Johansson, Zoete, Michielin et al., 2012)**.**

Swiss PDB Viewer 4.1.0: The Swiss-pdb viewer is an application method that runs a user-friendly interface allowing analysing proteins. Amino acid mutations, H-bonds, angles and distances between atoms are easy to obtain in from of graphic and menu interface. Moreover, Swiss-pdb viewer is tight to Swiss-Model, an automated homology modelling servers accessible from ExPASy (see for example Johansson, Zoete, Michielin, *et al.,* 2012)

Table I

GenBank	CCE37036.1
LOCUS	CCE37036
DEFINITION	mmpL12 (Mycobacterium tuberculosis UT205).
DBSOURCE	Embl accession HE608151.1
PUBMED	22404577
FASTA	>CCE37036.1 mmpL12 [Mycobacterium tuberculosis UT205]MARHDEAKAGGLFDRIGNFVVRWPLIVIGCWIAVAAALTL LLPTLQAQAAKREQAPLPPGAPSMVLQKEMSAAFQEKIET SALLLVLLTNENGLGPADEAVYRKLIENLR ADTQDKISVQDFLAVPEMKE LLASKDNKAWNLPITFAGDAASPETQAAFKRVAA IVKQTVAGTSLT VHLSGPIATVAD LTELGEKDVRIIEI GTAVSVLIILILVYRNLVTML VPLATIGASVVTAQGTLSGLAEFGLAVNMQAIVFMSAVMIGAGTDY AVFLISRYHDYVRHGEKSDMAVKKALMSIGKVITASAATVAVTF LAMVFTKLEVFSAVGPAIAVAITVSLLGAVTLLPAILTLT GRRGWIKPRRDL TSRMWRRSGVRIVRRSTIHLVGSLIVLVALAGCTLLIRFNYDDLKTVPQ HVESVKGYEAMNRHFPMNAMTPMVLFIKSPRDLRTPG ALADIEMMSREIAELPNIVMVRGLTRPNGEPLKETKVSFQAGEVG GKLDEATTLLEEHGGELDQLTGGAHQLADALAQIR NEINGAVASSSGIVNTLQAMMDLMGGDKTIRQLE NASQYVGRMRALGDNLSGTVTD AEQIATWASPMVNALNSSPVCNSDP ACRTSRAQLAAIVQAQDDGLLRSIR ALAVTLQQTQEYQT LARTVSTLDG QLKQVVSTLKAVDGLPTKLAQMQQGANALADGS AALAAGVQELVDQVKKMGS GLNEAADFLLGIKRD ADKPSMAGFNIPP QIFSRDEFKKGAQIFLSADGHAARYFVQSALNPATTEAMDQVNDILRVADSAR PNTELEDATIGLAGVPTALRDIRDYYNSDMKFIVIATIV IVFLILVILLRALVAP IYLIGSVLISYLSALGIGTLVFQLILGQEMHWSLPGLSF ILLVAIGADYNMLLISRIRDESPHGIRIGVIRTVGSTGGVITSA GLIFAASMFGLVGASINTMAQAGFTIGIGIVLDTFLVRTVTVPALTTMIG RANWWPSELGRDPSTPPTKADRWLRRVKGHRRKAPIPAPKPPHTKVVRN TNGHASKAATKSVPNGKPADLAEGNGEYLID HLRRHSLPLFGYAAMPAYDVVDGV SKPNGDG AHIGKEP VDHLLGHSLP LFG LAGLPSYDRWDDTSIGEPAVG HAGSKPDAKLST.

Pymol-v1.8.6.0.tar: PyMOL-is an open-source tool (Fig. 3 & 4) to visualize molecules. It runs on windows, Linux and macos equally well. PyMOL has competences in creating high-quality images from 3D structure; it has well established purposes for influencing structures and some elementary functions to evaluate their chemical properties. The opportunities to write scripts and plugins as well as to integrate pyMOL in custom software are vast and superior to most other programs. Pymol has been written mostly in the python language (www.python.org),while the time-critical parts of the system have been coded in c. This way, python programs intermingle most effortlessly with the pymol GUI (Rother, 2005).

Ramachandran assessment by mol probity (Fig. 5)**:** A Ramachandran plot is a visual graphic demonstration of the main-chain conformational tendencies of an amino acid. Ramachandran plotbased on experimental data is deciding whether scant data represent genuine conformations. Measured the pair-wise distances of main–chain conformational tendencies amino acids and showed the conformational relationships of amino acid are well preserved in a two dimensional. Amino acids in early and late evolutionary phases are situated in different zones in the two-dimensional map (Dahl, Bohannan, Mo, Vannucci, and Tsai, 2008) The first is a Ramachandran plot or Ramachandran map, which is simply a scatter plot of the ö, ø values for the amino acids in a single protein structure or a set of protein structures. It may be restricted to a single amino acid type and/or a single structural feature type, such as protein loops. The second is a Ramachandran distribution, a statistical representation of Ramachandran data, usually in the form of a probability density function. A probability density function gives the probability of finding an amino acid conformation in a specific range of φ, ψ values. For instance, if the function is given on a $10^\circ \times 10^\circ$ grid from $-180^\circ$to $+180^\circ$inφ, ψ (1296 values), then the distribution may give the probability per $10^\circ \times 10^\circ$region (Dunbrack *et al.,* 2013). It could also be expressed per degree squared or per radian squared. Such distributions may be derived for specific amino acid types and/or for specific structural features. There are numerous significant considerations in developing Ramachandran distributions from structural data, depending on the purpose of the derived distribution. First, while glycine and praline are usually treated separately, the other 18 amino acids are often treated as a single type. However, these amino acids are moderately different in their proportions of residues in the, β, polyproline II, and left-handed helical regions. Second, relatively different distributions are resolute when either all residues are used or only those outside the regular secondary structures of á-helices and â-sheet. The concluding assumed to be "intrinsic" favourites of the backbone, not subjective by forming specific hydrogen bonds present in regular secondary structures. Third, the quality and quantity of the data are crucial in determining distributions meant to act as eminence filters for newly determined structures or for structure prediction. As more structures have become available at higher resolutions, it is now possible to use quite large data sets with resolution cut-offs of 1.8 A$^\circ$or even better. Other filters have been used including B-factors and steric clashes to remove residues that may be model improperly or at least with considerable uncertainty within the electron density. For instance, by using higher resolution structures, B-factors, steric overlaps able to determine Ramachandran distributions with smaller "allowed" and "generously

allowed" regions than previous efforts. Fourth, most previous efforts have involved density estimation using simple histogram methods-the counts or proportion of counts of residues in non-overlapping square bind of the ö, ø space may be quite uneven (Dunbrack, Chase, Richardson *et al.*, 2013).

Clustal W Omega (Fig. 6): Clustal omega is a new multiple sequence alignment of three or more biological sequences of similar length. From the output of MSA applications. Homology can be inferred and the evolutionary relationship between the sequences studied. Tools/msa/ clustalo/pair-wise alignment, through which we examine the similarities of two sequences by searching for the alignment with the highest score. Modern phylogenetics relies on information extracted from genetic material such as DNA, RNA or protein sequences (Lassmann and Sonnhammer, 2005)**.** Graphical representation of MMPL protein is shown in Figure 1.
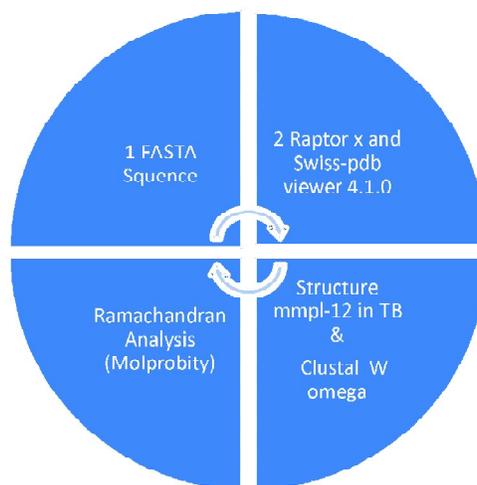

Figure 1: Graphical representation of MMPL-12 Structure prediction.

RESULTS

CCE37036.1 mmpL12 [Mycobacterium tuberculosis UT205]

M A R H D E A K A G G L F D R I G N F V V R W P L I V I G C W I A V A A A L T L L L P T L Q A Q A AKREQAPLPPGAPSMVLQKEMSAAFQEKIETSALLLVLLTNENGLGPADEAVYRKLIENL RADTQDKISVQDFLAVPEMKELLASKDNKAWNLPITFAGDAASPETQAAFKRVA AIVKQTVAGTSLTVHLSGP IATVADLTELGEKDVRIIEIGTAVSVLIILILVYRNL VTMLVPLATI GASVVTAQGTLSGLA EFGLAVNMQAIVFMSAVMIGA GTDYAVFLISRYH DYVRHGEKSD MAVKKALMSIGKVITASAATVAVT FLAMVFTKLEVFSAVGPAIAVAITVSLL GAVTLLPAILTLT GRRGWIKPRRDLTSRMWRRSGVRIVRRSTIHLVGSL IVLVALAGCTLLIRFN YDDLKTVPQH VESVKGYEAMNRHFPMNAMTPMVLFIKSPRDLRTPGALA DIEMMSREIAELPNIVM VRGLTRPNGEPLKETKVSFQAGEVGGKLDEATTLLEEHGGELDQLTGGAHQLA DALAQIRNEINGAVAS SSGIVNTLQAMMDLMGG DKTIRQLENASQY VGRMRALGD

NLSGTVTDAE QIATWASPMVNALNSSPVCNSDPACRTSRA QLAAIV QAQDDGLLR
SIRALAVTLQQTQEYQTLAR TVSTLDGQLKQVVSTLKAVDGLPTKLAQMQQGANALA
DGSAALAAGVQELVDQVKKMGSGLNEAADFLLGIKRDADKPSMAGF NIPPQIFSRDEF
K K G A Q I F L S A D G H A A R Y F V Q S A L N P A T T E A M D Q V N D I L R V A D S A R P N T E
LEDATIGLAGVPTALRDIRDYYNSDMKFIVIATIVI VFLILVILLRALV APIYLIGSVLISYL
SALGIGTLVFQLILGQEMHWSLPGL SFILLVAIGADYNMLLISRIRDESPHGI RIGVIRT
VGSTGGV ITSAGLIFAASMFGLVGASINTMAQAGFTIGIGIVLDTFLVRTVT VPALTTMIGRA
NWWPSELGRDPSTPPTKADRWLRRVKGHRRKAPIPAPKP PHTKVVRNTN GHASKAA
TKSVPNGKPADLAEGNGEYL IRRHSLPLFGYAAMPAYDVVDGVSKPNG DGAHIGK
EPVDHLLGHSLPLFGLAGLPSYDRWDDTSIGEPAVGHAGSKPDAKLST

GDT uSeqID: 124 (13) (RaptoX)

SeqID Model Name Template(s): 5khn-A

Rank P-value Score uGDT 4.37e-10

A visual summary of the above results is provided in Figure 2, Figure 3, Figure 4, Figure 5 and Figure 6.
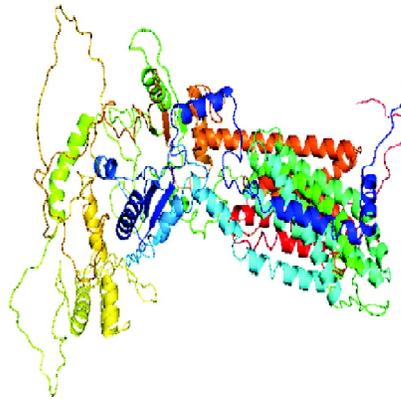

Figure 2: Mycobacterial membrane protein Large-12(MMPL-12) in
Mycobacterium tuberculosis **Using RaptorX**

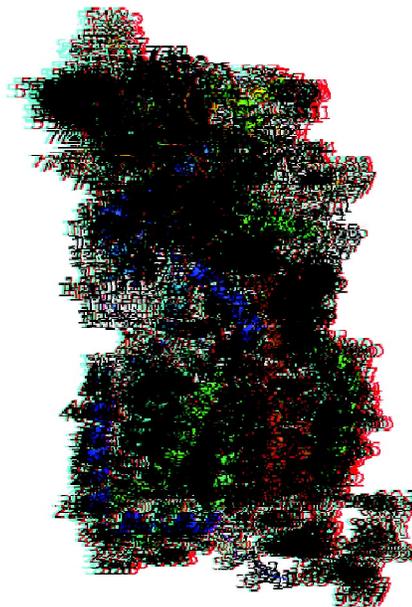**Figure 3: Position of Residues of mycobacterial membrane protein large-12**

Figure 4: Value of radius of mycobacterial membrane protein Large -12
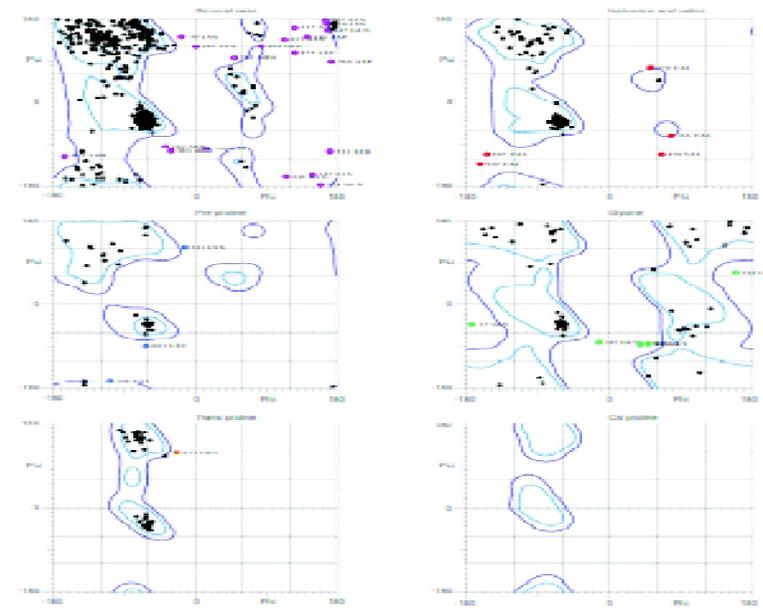*Mycobacterium tuberculosis* **Using Pymolv-1.8.6.0**

**Figure 5: Ramachandran Validation using mol probily (Vincent B. Chen,
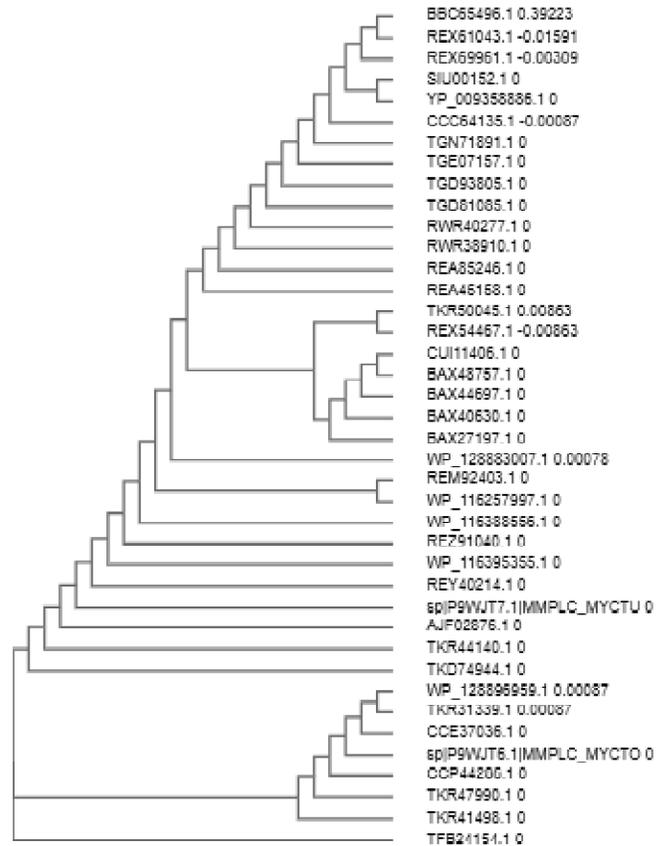W. Bryan Arendall ɪɪɪ, Jeffrey j. Headd, Daniel A. Keedy)**

Figure 6: Phylogenetic tree using crustal omega (multiple sequences Alignment)

DISCUSSION

The total energy of mycobacterial membrane protein large 12 (MMP12) is 38098.148 kj/mol by computational calculation. The calculation is based on position of amino acids, bonds, angles, torsion, improper, non-bonded and total energy of amino acids making a mycobacterial membrane protein large12 (MMPL12) in TB. The energy minimization of amino acids of MMpl12 is also calculated by Swiss Pdb Viewer. In which the minimum energy of bonds is 7039.374 angles 8960.096torsion 4160.980 improper 5645.088 nonbonded 31192.20 and total minimum energy is 38098.148 KJ/mol in Mycobacterium tuberculosis. This energy is responsible for the threshold value to express the adipose gene and other significant functions in mycobacterium tuberculosis.MMPL-12 is responsible for major adipose tissue depositions. This expression is balanced with the help of energy. A scatter plot of the φ, ψ values for the amino acids of MMPL-12 is also

analysed (Ramchandran plot). A statistical representation (Ramchandran distribution) in the form of probability of finding an amino acid conformation in a specific range of φ, ψ values is considered. First glycine and proline are usually treated separately; the other 21 amino acids are often treated as a single type. However, these amino acids are moderately different in their proportion of residues in the, β, polypro line ц, and left - handed helical regions. The "intrinsic" favourites of the backbone, not particular by forming specific hydrogen bands present in regular secondary structures.88.5% (877/991) of all residues were in favoured (98%) regions. 96.6% (957/991)of all residues were in allowed (>99.8%) regions; outliers (phi, psi) are 2 PRO (-26.0, 120.9) in validation. The position of amino acids, the position of residues of amino acids and values of bonds, torsion, improper non bonded and total energy calculation of amino acid of MMPL-12 in *Mycobacterium tuberculosis* is given below. Clustal omega multiple sequences alignment number of sequences40 input. Sequences high score slimily in the evolutionary relationship.>CCE37036.1 mmpl-12 mycobacterim tuberculosis ut205 >REA45158.1 RND transporter mmpl-12.

In-silico comparative genomics is a useful approach which can be applied for therapeutic target identification in *Mycobacterium tuberculosis*. The degree of functioning of a protein crucially depends on 3D structure. The hypothetical 3D structure can be used in future as a template for homologous protein modelling of the target. Protein conformation can be perfectly utilized for phylogeny establishment. The outcomes of computational bioinformatics tool might reveal better results in comparison to traditional ones and coveted platform for efficient genetic characterization. We can predict the function of hypothetical protein based on template protein with its unravelled similarity. Their function can be assigned as superimposition of template and target protein.

The position of amino acid of MMPL12 by Swiss-PDb viewer is shown in Table II.

Tabel II

ATOM 1	HHT (H).
ATOM 1, 64, 70, 128, 222, 252, 258, 262, 289, 296, 316, 369, 425, 431, 434, 437, 458, 459, 471, 551, 552, 555, 574, 598, 683, 711, 734, 780, 824, 879, 900, 940, 951, 980:	**MET (M).**
ATOM 2, 7, 9, 33, 35-37, 47, 49, 50, 55, 61, 72, 73, 82, 97, 100, 111, 124, 133, 139, 150, 151, 157, 158, 163, 164, 171, 184, 187, 205, 231, 236, 244, 249, 254, 260, 265, 270, 290, 294, 304, 306, 307, 310, 315, 326, 330, 332, 334, 342, 348, 394, 396, 424, 433, 452, 454, 464, 492, 502, 520, 524, 526, 528, 537, 539, 550, 567, 576, 588, 592, 595, 601,	**ALA (A).**

ATOM 1	*HHT (H).*
613, 619, 622, 623, 627, 638, 640, 653, 672, 681, 687, 689, 691, 695, 696, 698, 699, 718, 719, 729, 735, 753, 759, 763, 764, 771, 775, 779, 790, 793, 802, 807, 812, 830, 845, 848, 863, 893, 954, 976, 984 :	
ATOM 3, 15, 22, 52, 110, 161, 198, 217, 276, 282, 355, 356, 362, 363, 368, 371, 372, 376, 379, 380, 403, 427, 445, 448, 461, 473, 477, 531, 562, 573, 575, 615, 618, 634, 637, 654, 727, 746, 765, 788, 794, 814, 817, 844, 905, 907, 915, 920, 970, 983, 993:	**ARG (R).**
ATOM 4, 178, 278, 283, 384, 415, 428, 509, 521, 762, 912:	**HISA (H).**
ATOM 5,14, 98, 112, 115, 121, 136, 149, 188, 196, 268, 279, 288, 364, 407, 408, 446, 455, 500, 514, 525, 553, 558, 579, 587, 611, 629, 630, 660, 674, 692, 706, 720, 728, 730, 747, 760, 781, 785, 791, 801, 815, 818, 823, 897, 908, 965:	**ASP (D).**
ATOM 6, 53, 69, 76, 79, 91, 99, 107, 127, 130, 154, 101, 194, 201, 245, 285, 322, 417, 423, 457, 462, 465, 481, 485, 494, 501, 507, 508, 512, 533, 565, 589, 648, 703, 717, 748, 778, 798, 800, 878, 909, 999:	**GLU (E).**
ATOM 8, 51, 68, 77, 104, 116, 129, 135, 138, 160, 167, 195, 286, 292, 293, 300, 320, 360, 410, 420, 442, 484, 487, 498, 559, 664, 671, 679, 709, 710, 726, 731, 750, 751, 825:	**LYSH (L).**
ATOM 10, 11, 17, 29, 60, 93, 95, 148, 172, 181, 193, 203, 230, 238, 242, 247, 264, 266, 284, 299, 328, 341, 354, 357, 374, 387, 397, 421, 451, 474, 480, 493, 496, 497, 510, 511, 518, 519, 536, 543, 556, 557, 572, 578, 583, 631, 661, 675, 686, 693, 700, 714, 724, 736, 752, 761, 805, 808, 854, 865, 867, 876, 885, 895, 913, 917, 923, 926, 927, 933, 942, 945, 955, 959, 961, 982, 992:	**GLY (G).**
ATOM 13, 19, 74, 122, 146, 159, 246, 257, 272, 313, 318, 324, 404, 429, 440, 490, 721, 737, 744, 749, 756, 767, 826, 836, 871, 888, 836, 941, 956, 967:	**PHE (F).**
ATOM 16, 26, 28, 32, 78, 106, 117, 144, 165, 183, 199, 200, 202, 210, 211, 213, 229, 255, 263, 274, 298, 302, 331, 335, 349, 359, 377, 383, 390, 402, 441, 456, 463, 469, 530, 534, 544, 561, 591, 624, 636, 725, 739, 743, 755, 786, 804, 816, 827, 829, 832, 834, 838, 841, 850, 853, 858, 866, 874, 889, 894, 903, 906, 914, 916, 919, 929, 935, 948, 958, 960, 962, 981:	**ILE (I).**
ATOM 12, 25, 38, 40, 41, 45, 57, 66, 83, 84, 85, 87, 88, 94, 105, 109, 123, 131, 132, 142, 175, 179, 189, 192, 209, 212, 214, 219, 223, 226, 240, 243, 248, 273, 295, 314, 321, 339, 340, 345, 345, 346, 350, 352, 365, 385, 389, 392, 395, 400, 401, 409, 439, 447, 453, 466, 475, 483, 499, 505, 506, 506, 513, 516, 523, 527, 548, 554, 564, 577, 581, 602, 621, 632, 633, 639, 643, 652, 659, 663, 670, 676, 680, 690, 697, 704, 715, 722, 723, 757, 772, 787, 799, 806, 813, 837, 839, 842, 843, 846, 852, 857, 861, 864, 869, 873, 875, 883, 886, 890, 891, 901, 902, 934, 943, 964, 968, 977, 991:	**LEU (L).**

ATOM 1	HHT (H).
ATOM 20, 21, 27, 34, 65, 86, 101, 119, 125, 162, 166, 170, 177, 186, 197, 206, 208, 215, 220, 224, 233, 234, 250, 256, 261, 271, 281, 291, 301, 309, 311, 317, 323, 327, 333, 337, 343, 375, 378, 386, 391, 393, 412, 416, 419, 438, 470, 472, 488, 495, 538, 545, 517, 585, 599, 607, 625, 641, 656, 666, 667, 673, 701, 705, 708, 768, 783, 789, 809, 828, 833, 835, 840, 847, 856, 870, 892, 918, 922, 928, 944, 963, 969, 972, 974:	**VAL (V).**
ATOM 23, 31, 140, 358, 370, 594, 881, 986, 987:	**TRP (W).**
ATOM 30, 398, 608, 614:	**CYSH (C).**
ATOM 24, 43, 56, 58, 59, 62, 96, 126, 143, 153, 182, 225, 329, 347, 361, 413, 430, 436, 444, 450, 467, 478, 482, 597, 606, 612, 677, 732, 740, 741, 774, 795, 810, 849, 884, 911, 975, 988:	**PRO (P).**
ATOM 44, 80, 89, 113, 145, 155, 169, 173, 176, 185, 190, 204, 221, 228, 235, 239, 267, 303, 308, 312, 319, 336, 344, 351, 353, 366, 382, 399, 411, 435, 449, 476, 486, 503, 504, 517, 547, 560, 584, 586, 593, 616, 642, 646, 651, 655, 658, 669, 678, 776, 777, 797, 803, 811, 831, 868, 921, 925, 930, 950, 957, 966, 971, 973, 978, 979:	**THR (T).**
ATOM 46, 48, 54, 67, 75, 114, 120, 156, 168, 237, 253, 414, 491, 515, 522, 529, 549, 563, 569, 590, 620, 626, 628, 644, 645, 647, 650, 662, 665, 682, 684, 685, 702, 707, 742, 754, 769, 782, 872, 877, 953:	**GLN (Q).**
ATOM 63, 71, 81, 118, 134, 152, 174, 180, 207, 232, 241, 259, 275, 287, 297, 305, 325, 338, 367, 373, 381, 388, 418, 443, 460, 489, 540-542, 568, 582, 596, 604, 605, 610, 617, 635, 657, 668, 694, 713, 733, 745, 758, 770, 792, 822, 855, 859, 862, 882, 887, 904, 910, 924, 931, 939, 947, 989:	**SER (S).**
ATOM 18, 90, 92, 108, 137, 141, 218, 251, 405, 426, 432, 468, 479, 532, 535, 546, 566, 580, 600, 603, 609, 688, 716, 738, 773, 784, 796, 821, 899, 949, 985:	**ASN (N).**
ATOM 102, 216, 269, 277, 280, 406, 422, 570, 649, 766, 819, 820, 851, 860, 898:	**TYR (Y).**

References

1. Alam, A., Ali, S., Malik, Z., Ishrat, R.(2016). In Silico Analysis of Mmpl Gene Family of Mycobacterium Tuberculosis: a Novel Target for Anti-Tb Drugs. *Int. J. Sci Res*, 5(1), 210-218.

2. Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr. 2010 Jan;66(Pt 1):12-21. doi: 10.1107/S0907444909042073. Epub 2009 Dec 21. PMID: 20057044; PMCID: PMC2803126.

3. Dahl, D.B., Bohannan, Z., Mo, Q,, Vannucci, M and Tsai J. Assessing side-chain perturbations of the protein backbone: a knowledge-based classification of residue Ramachandran space. *J Mol Biol.*, 378(3):749-58. doi: 10.1016/j.jmb.2008.02.043. Epub 2008 Feb 29. PMID: 18377931; PMCID: PMC2440669.

4. Dunbrack, R. , Chase, F. , Richardson, J.S., Keedy, D.A. and Richardson, D.C. (2013). The Plot Thickens: More Data, More Dimensions, More Uses. In M. Bansal & N. Srinivasan (Eds.), Biomolecular Forms and Functions: A Celebration of 50 Years of the Ramachandran Map. Singapore; Hackensack, World Scientific Pub. Co.

5. Johansson, M.U., Zoete, V., Michielin, O. *et al.* (2012). Defining and searching for structural motifs using DeepView/Swiss-PdbViewer. BMC Bioinformatics 13, 173. https://doi.org/10.1186/1471-2105-13-173

6. Lassmann, T & Sonnhammer, E. (2005). KALIGN: An accurate and fast multiple sequence alignment algorithm. *BMC bioinformatics.* 6. 298. 10.1186/1471-2105-6-298.

7. Rother, K. (2005). Introduction to PyMOL. *Methods Mol. Biol. Clift. Nj,* 635(8),1–32.

8. Sz»kely, R., Cole, S.T.(2016). Micro Commentary Mechanistic insight into mycobacterial MmpL protein function. *Molecular Microbiology.* 99(5), 831–834.

9. Wang, S., Li, W., Liu, S. and Xu, J. (2016). RaptorX-Property: a web server for protein structure property prediction. *Nucleic Acids Res.,* 44(W1), W430–W435.

10. Yan, H., Xu, R., Zhang, X., Wang, Q., and Pang, J.(2018). Identifying differentially expressed long non-coding RNAs in PBMCs in response to the infection of multidrug-resistant tuberculosis. *Infect. Drug Resist.,* 11, 945–959.

11. Zhu, F., Ou, Q., Zheng, J (2018). Application Values of T-SPOT.TB in Clinical Rapid Diagnosis of Tuberculosis.," *Iran. J. Public Health,* 47(1), 18–23.