Indian J. Applied & Pure Bio. Vol. 37(1), 275-283 (2022). **ISSN: 0970-2091**
A web of Science Journal

Structure-based Drug Design on *Mycobacterium tuberculosis* (Mtb) targeting unsaturated-phospholipid methyltransferase - An *In silico* analysis.

Harsha Upadhyay and Jaishree Sikka*

Department of Botany
P.M.B. Gujrati Science College Indore-452001 (India)
jaiusikka@hotmail.com
harshadubey2008@gmail.com

Abstract

Bacteria control the biophysical properties of their membranes to allow them to thrive in a wide range of physical environments[6]. *Mycobacterium tuberculosis* (Mtb) contains an unsaturated-phospholipid methyltransferase (UPM) which is associated with the transferase family, especially those who move methyltransferase one-carbon group[16]. Deactivation of the UPM gene coding region is necessary to weaken the activity of Mtb bacteria; therefore, disrupting the function of the UPM was the study goal. This study helps to understand the effective inhibitors, which can inhibit the function of UPM. The protein sequence of UPM has been retrieved from the Swiss-Prot database. Structural similarity search has been performed to find templates by standalone BLAST against the PDB database. BLAST shows the protein 3D structure of the UPM, hence it has been downloaded from the PDB database and docking studies have been carried out with the same.

Bacteria control the biophysical properties of their membranes to allow them to thrive in a wide range of physical environments[6]. *Mycobacterium tuberculosis* (Mtb) contains an unsaturated-phospholipid methyltransferase (UPM) which is associated with the transferase family, especially those who move methyltransferase one-carbon group[16]. *Mycobacterium tuberculosis* (Mtb) belongs to the *Mycobacterium* genus. It is a pathogenic bacteria species causing Tuberculosis in human lungs[9]. Robert Koch in 1882 discovered Mtb, which is an obligate aerobic Gram-positive mycobacterium[4]. In Mtb genome approx 40,000 genes are present and many of the genes functions are still unknown[15, 17]. Of the 4411522 base pair long genome of *Mtb*, 34660.1 base pairs are functional, which is 90.80% of the total genome[5]. Mtb genome has a high content of G+C approx 65.60%. Compared to other species of *Mycobacterium*, only six pseudogenes have been found in the genome of Mtb, indicating that throughout evolution, the Mtb genome has lost many

*Head of the department

functional and non-functional regions[7]. Evolutionary studies of lost areas help differentiate between *Mycobacterium* species as well as track evolutionary paths that geographically separate Mtb strains. Genomic evolutionary studies of each strain show that Mtb has undergone several changes in the genome, due to which it also shows high drug resistance[8]. This research has attempted to find out which product's structural and functional Mtb gene domain encodes. Other common names for UPM include cyclopropane synthetase. Weakening of *Mtb* requires inactivation of the UPM gene coding region, so the study goal was to inhibit the function of UPM. UPM induced catalytic reaction which Equation showed in below.

Protein sequences fetched from the Swiss Prot tool[2], based on a review of UPM literature from PubMed, PMC, Oxford journals, Nature journals, *etc*., which have the elite of annotation protein sequence data, provides the characterization of protein sequences including their functions. The UPM query sequence has a total of 302 amino acids with a molecular weight of 34660.01. Protein sequence analysis by EMBOSS, Pepwheel, and Pepstats. Predicted the primary structure of the protein with the assist of ProtParm software[3] and the secondary structure with the help of SOPMA software[12,14]. Protein 3D structure of UPM was obtained from PDB database[1,13]. With the assist of the standalone blastp found that the UPM sequence for modeling in the PDB database was similar to the template and evolutionary studies performed by Philippe. For drug comparison, Gentamicin, Kanamycin, Prednisone, Rifampin, and Streptomycin, data were obtained from the PubChem database and the drugs have been screened by JMol software[10]. Molegro Virtual Docker software was used for docking the 3D structure of the UPM. With the assist of ADMET software, checked the effectiveness of drugs at the target site[11]. In this work, studied the drug actions of Gentamicin, Kanamycin, Prednisone, Rifampin, and Streptomycin on UPM.

The sequence of UPM was analyzed by EMBOSS Pepwheel and Pepstats. The resulting Pepwheel showed molecular arrangements and with the help of pepstats the amino acid properties such as molecular weight=34660.01, Isoelectric point=4.9118, total numbers, Mole%, DayhoffStat, *etc*. were analyzed. These are shown in Figure 1 and 2. Prediction of protein primary structure by Protparam, by which the number of amino acids (302), the molecular weight of protein sequence (34660.1), Theoretical pI (5.11), amino acid composition (C-1555, H-2363, N-409, O) -466, s) -13) and total atom number (4806), negatively charged(Asp + Glu)-(47) and positively charged (Arg + Lys)-(34) residues, *etc*. are shown in Figure 3. All the values helped understand Protein sequence length, chemical, and physical properties. Protein secondary structure prediction was done SOPMA, which determined helix, β bridge, random coil, ambiguous states, etc. are shown in Figure 4 and graph indicating helix, coil, and turn shown at particular length of the sequence are shown in Figure 5 which helped in understanding binding sites and the number of cavity. UPM and drugs 3D structure analyzed by JMole viewer, which determined atom arrangement, bond length, bond width, angle value and surface of the atom, etc are shown in Figure 6. Results from blastp which shows 99% identity and 99% similarity between UPM and PDB database of bacteria are shown in Figure 7.

Results from Molegro Virtual Docker showed ligand binding site or drug target site. In that bond energy, affinity, torsions, and rerank scores are shown in Figures 8, 9, 10, 11 and 12. By the ADMET study, found effects of drugs on the human body based on various categories are shown in Figure 13. The highest binding average energy with UPM v/s Kanamycin was 141.5768 and Rifampin was 140.2142.

Mycobacterium tuberculosis has an unsaturated-phospholipid methyltransferase enzyme belonging to the family of transferases[16]. Deactivation of the UPM gene coding region is necessary to weaken the activity of Mtb bacteria. Hence, study target was to stop the function of UPM. This study would help to understand the effective inhibitors, which can stop the function of UPM. Gentamycin, Kanamycin, Prednisone, Rifampin, and Streptomycin drugs were prescribed by the physicians against Mtb. Study target was to find out the most stable, effective, and fewer side effects on the human body. Two of the four drugs namely Kanamycin and Rifampin are the most effective drugs to stop the function of UPM, which are calculated by docking software Molegro Virtual Docker and ADMET.

Equation

S-adenosyl-L-methionine + phospholipid olefinic fatty acid → S-adenosyl-L-homocysteine+ phospholipid cyclopropane fatty acid

Equation 1: UPM induced catalytic reaction

Figures :

Residue	Number	Mole%	DayhoffStat	Residue	Number	Mole%	DayhoffStat	Residue	Number	Mole%	DayhoffStat
A = Ala	22	7.285	0.847	I = Ile	13	4.305	0.957	S = Ser	16	5.298	0.757
B = Asx	0	0.000	0.000	K = Lys	17	5.629	0.853	T = Thr	19	6.291	1.031
C = Cys	4	1.325	0.457	L = Leu	28	9.272	1.253	V = Val	14	4.636	0.702
D = Asp	25	8.278	1.505	M = Met	9	2.980	1.753	W = Trp	5	1.656	1.274
E = Glu	22	7.285	1.214	N = Asn	8	2.649	0.616	X = Xaa	0	0.000	0.000
F = Phe	16	5.298	1.472	P = Pro	13	4.305	0.828	Y = Tyr	16	5.298	1.558
G = Gly	20	6.623	0.788	Q = Gln	10	3.311	0.849	Z = Glx	0	0.000	0.000
H = His	8	2.649	1.325	R = Arg	17	5.629	1.149				

PEPSTATS of YP_001281795.1 from 1 to 302

Molecular weight = 34660.01 **Residues** = 302

Average Residue Weight = 114.768 **Charge** = -9.0

Isoelectric Point = 4.9118

A280 Molar Extinction Coefficient = 48930

A280 Extinction Coefficient 1mg/ml = 1.41

Probability of expression in inclusion bodies = 0.722

Property	Residues	Number	Mole%
Tiny	(A+C+G+S+T)	81	26.821
Small	(A+B+C+D+G+N+P+S+T+V)	141	46.689
Aliphatic	(I+L+V)	55	18.212
Aromatic	(F+H+W+Y)	45	14.901
Non-polar	(A+C+F+G+I+L+M+P+V+W+Y)	160	52.980
Polar	(D+E+H+K+N+Q+R+S+T+Z)	142	47.020
Charged	(B+D+E+H+K+R+Z)	89	29.470
Basic	(H+K+R)	42	13.907
Acidic	(B+D+E+Z)	47	15.563

Figure- 1 Protein sequence analysis by EMBOSS Pepstats

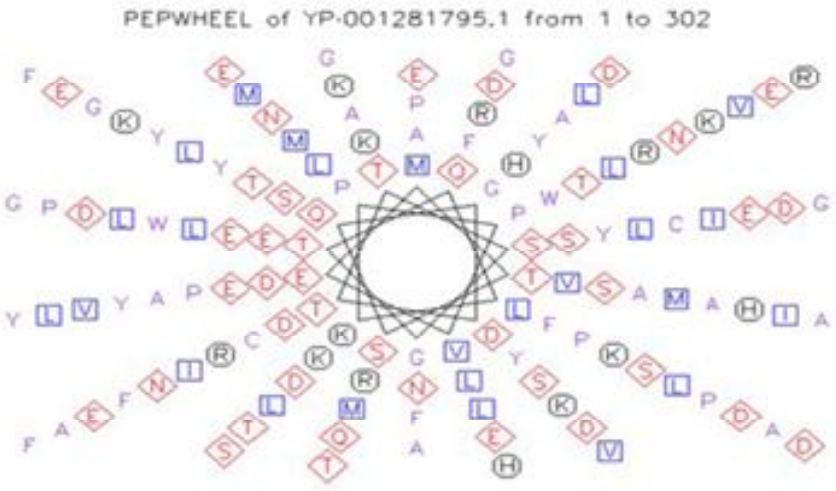


Figure- 2 Protein sequence analysis by EMBOSS Pepwheel

```
      10         20         30         40         50         60
MTSQGDTTSG TQLKPPVEAV RSHYDKSNEF FKLWLDPSMT YSCAYFERPD MTLEEAQYAK
      70         80         90        100        110        120
RKLALDKLNL EPGMTLLDIG CGWGSTMRHA VAEYDVNVIG LTLSENQYAH DKAMFDEVDS
     130        140        150        160        170        180
PRRKEVRIQG WEEFDEPVDR IVSLGAFEHF ADGAGDAGFE RYDTFFKKFY NLTPDDGRML
      190        200        210        220        230        240
LHTITIPDKE EAQELGLTSP MSLLRFIKFI LTEIFPGGRL PRISQVDYYS SNAGWKVERY
      250        260        270        280        290        300
HRIGANYVPT LNAWADALQA HKDEAIALKG QETYDIYMHY LRGCSDLFRD KYTDVCQFTL

VK
```

Number of amino acids: 302	**Atomic composition:**
Molecular weight: 34660.1	Carbon C 1555
Theoretical pI: 5.11	Hydrogen H 2363
Total number of atoms: 4806	Nitrogen N 409
Formula: C1555H2363N409O466S13	Oxygen O 466
	Sulfur S 13

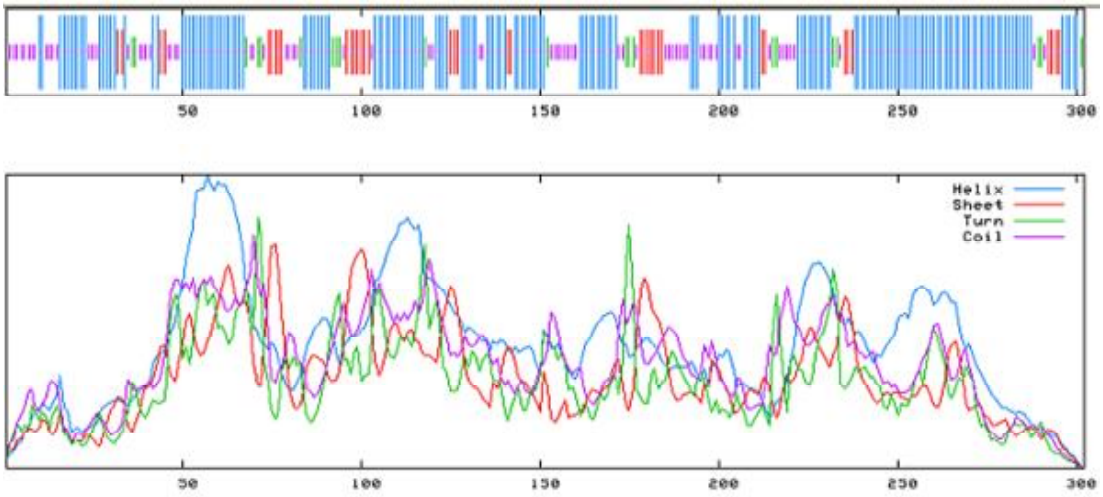
Total number of negatively charged residues (Asp + Glu): 47

Total number of positively charged residues (Arg + Lys): 34

Figure- 3 Protein Primary structure prediction by Prot Param

```
         10         20         30         40         50         60         70
          |          |          |          |          |          |          |
MTSQGDTTSGTQLKPPVEAVRSHYDKSNEFFKLWLDPSMTYSCAYFERPDMTLEEAQYAKRKLALDKLNL
eccccccccchhcccchhhhhhhhcccchhhhheehcttccccchheecccchhhhhhhhhhhhhhhhhhtc
EPGMTLLDIGCGWGSTMRHAVAEYDVNVIGLTLSENQYAHDKAMFDEVDSPRRKEVRIQGWEEFDEPVDR
cttceeeeeccccthhhhhhhhhttteeeeeeechhhhhhhhhhhhhhtcchhhheeehhhhhcchhhhh
IVSLGAFEHFADGAGDAGFERYDTFFKKFYNLTPDDGRMLLHTITIPDKEEAQELGLTSPMSLLRFIKFI
heehhhhhhhhhtccccccchhhhhhhhhhhcctttceeeeeeecccccchhhccccchhhhhcchhh
LTEIFPGGRLPRISQVDYYSSNAGWKVERYHRIGANYVPTLNAWADALQAHKDEAIALKGQETYDIYMHY
hheecttccccchhhhhhhhhttceeehhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhh
LRGCSDLFRDKYTDVCQFTLVK
hhhhhhhhcttceeeehhhhht

Sequence length :   302

SOPMA :
   Alpha helix      (Hh) :   171 is  56.62%
   3(10) helix      (Gg) :     0 is   0.00%
   Pi helix         (Ii) :     0 is   0.00%
   Beta bridge      (Bb) :     0 is   0.00%
   Extended strand  (Ee) :    38 is  12.58%
   Beta turn        (Tt) :    21 is   6.95%
   Bend region      (Ss) :     0 is   0.00%
   Random coil      (Cc) :    72 is  23.84%
   Ambigous states  (?)  :     0 is   0.00%
   Other states          :     0 is   0.00%
```

Figure- 4. Protein Secondary structure by SOPMA




```
rameters :
  Window width              :  17
  Similarity threshold      :   8
  Number of states          :   4
```

Figure- 5. Protein Secondary structure by SOPMA

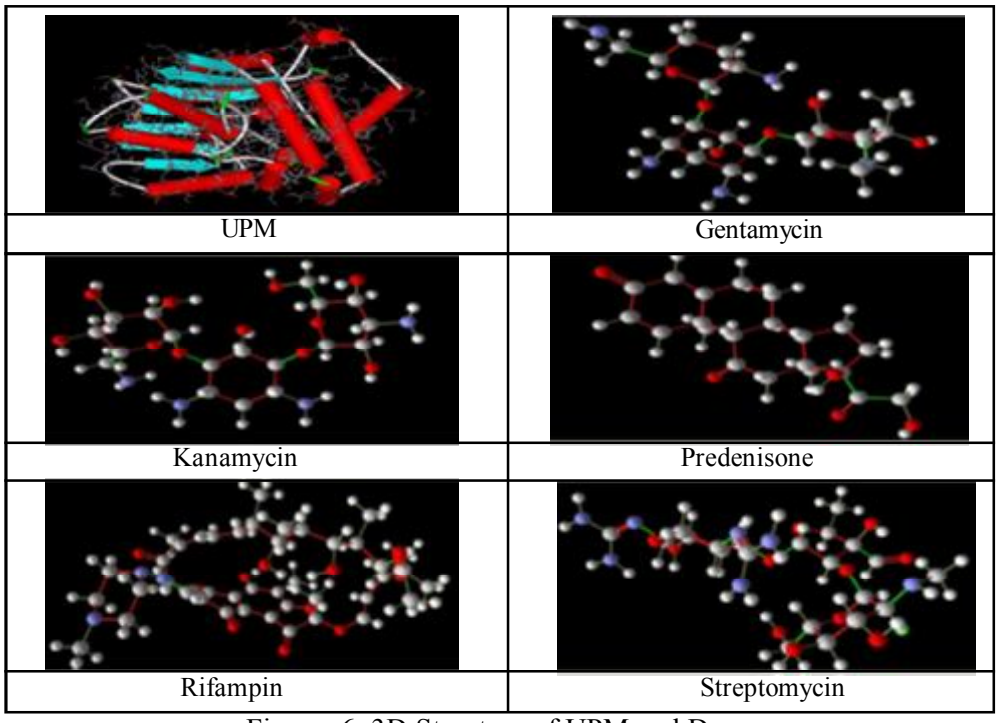


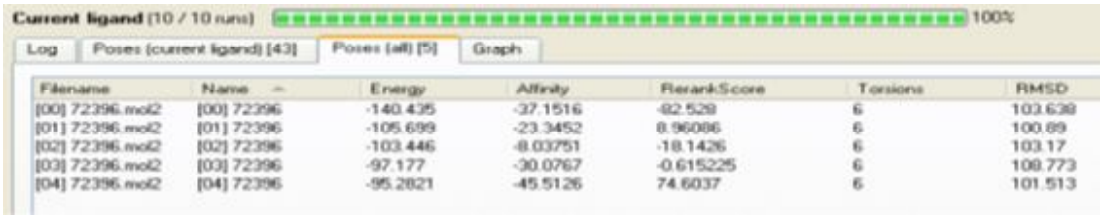
Figure- 6. 3D Structure of UPM and Drugs

Result of Blast

```
>pdb|1KPI|A Chain A, Crystal Structure Of Mycolic Acid Cyclopropane Synthase
           Cmaa2 Complexed With Sah And Dddmab
          Length = 302

 Score =  622 bits (1605), Expect = e-179
 Identities = 301/302 (99%), Positives = 301/302 (99%)

Query: 1    MTSQGDTTSGTQLKPPVEAVRSHYDKSNEFFKLWLDPSMTYSCAYFERPDMTLEEAQYAK 60
             MTSQGDTTSGTQLKPPVEAVRSHYDKSNEFFKLWLDPSMTYSCAYFERPDMTLEEAQYAK
Sbjct: 1    MTSQGDTTSGTQLKPPVEAVRSHYDKSNEFFKLWLDPSMTYSCAYFERPDMTLEEAQYAK 60

Query: 61   RKLALDKLNLEPGMTLLDIGCGWGSTMRHAVAEYDVNVIGLTLSENQYAHDKAMFDEVDS 120
             RKLALDKLNLEPGMTLLDIGCGWGSTMRHAVAEYDVNVIGLTLSENQYAHDKAMFDEVDS
Sbjct: 61   RKLALDKLNLEPGMTLLDIGCGWGSTMRHAVAEYDVNVIGLTLSENQYAHDKAMFDEVDS 120

Query: 121  PRRKEVRIQGWEEFDEPVDRIVSLGAFEHFADGAGDAGFERYDTFFKKFYNLTPDDGRML 180
             PRRKEVRIQGWEEFDEPVDRIVSLGAFEHFADGAGDAGFERYDTFFKKFYNLTPDDGRML
Sbjct: 121  PRRKEVRIQGWEEFDEPVDRIVSLGAFEHFADGAGDAGFERYDTFFKKFYNLTPDDGRML 180

Query: 181  LHTITIPDKEEAQELGLTSPMSLLRFIKFILTEIFPGGRLPRISQVDYYSSNAGWKVERY 240
             LHTITIPDKEEAQELGLTSPMSLLRFIKFILTEIFPGGRLPRISQVDYYSSNAGWKVERY
Sbjct: 181  LHTITIPDKEEAQELGLTSPMSLLRFIKFILTEIFPGGRLPRISQVDYYSSNAGWKVERY 240

Query: 241  HRIGANYVPTLNAWADALQAHKDEAIALKGQETYDIYMHYLRGCSDLFRDKYTDVCQFTL 300
             HRIGANYVPTLNAWADALQAHKDEAIALKGQET DIYMHYLRGCSDLFRDKYTDVCQFTL
Sbjct: 241  HRIGANYVPTLNAWADALQAHKDEAIALKGQETCDIYMHYLRGCSDLFRDKYTDVCQFTL 300

Query: 301  VK 302
             VK
Sbjct: 301  VK 302
```

Figure- 7. Structure similarity searching by standalone blastp against PDB database

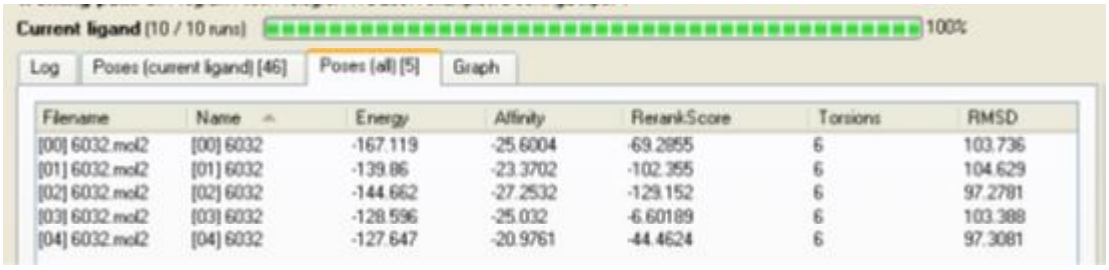
Molecular Docking by Molegro Virtual Docker Docking Result



Filename	Name	Energy	Affinity	RerankScore	Torsions	RMSD
[00] 72396.mol2	[00] 72396	-140.435	-37.1516	-92.529	6	103.638
[01] 72396.mol2	[01] 72396	-105.699	-23.3452	8.96086	6	100.89
[02] 72396.mol2	[02] 72396	-103.446	-8.03751	-18.1426	6	103.17
[03] 72396.mol2	[03] 72396	-97.177	-30.0767	-0.615225	6	109.773
[04] 72396.mol2	[04] 72396	-95.2821	-45.5126	74.6037	6	101.513

Average Energy: -108.40782

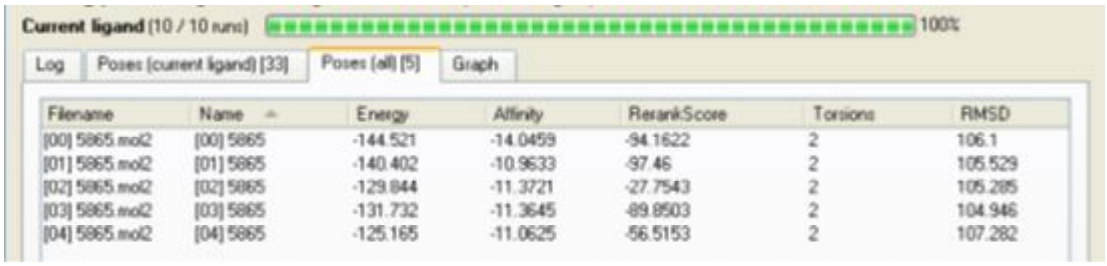
Figure- 8. Molecular Docking UPM vs Gentamycin



Filename	Name	Energy	Affinity	RerankScore	Torsions	RMSD
[00] 6032.mol2	[00] 6032	-167.119	-25.6004	-69.2855	6	103.736
[01] 6032.mol2	[01] 6032	-139.86	-23.3702	-102.355	6	104.629
[02] 6032.mol2	[02] 6032	-144.662	-27.2532	-129.152	6	97.2781
[03] 6032.mol2	[03] 6032	-128.596	-25.032	-6.60189	6	103.388
[04] 6032.mol2	[04] 6032	-127.647	-20.9761	-44.4624	6	97.3081

Average Energy: -141.5768

Figure-9. Molecular Docking UPM vs Kanamycin



Filename	Name	Energy	Affinity	RerankScore	Torsions	RMSD
[00] 5865.mol2	[00] 5865	-144.521	-14.0459	-94.1622	2	106.1
[01] 5865.mol2	[01] 5865	-140.402	-10.9633	-97.46	2	105.529
[02] 5865.mol2	[02] 5865	-129.844	-11.3721	-27.7543	2	105.285
[03] 5865.mol2	[03] 5865	-131.732	-11.3645	-89.8503	2	104.946
[04] 5865.mol2	[04] 5865	-125.165	-11.0625	-56.5153	2	107.282

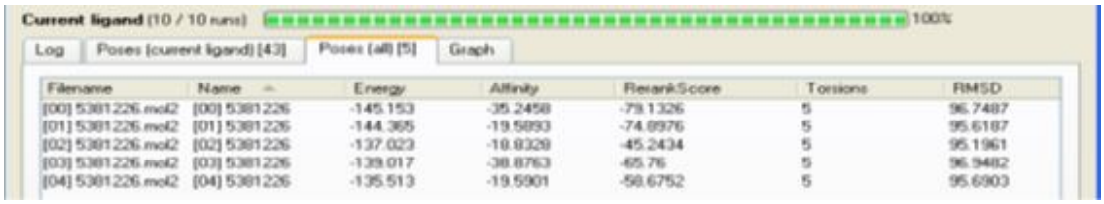
Average Energy: -134.3328

Figure- 10. Molecular Docking UPM vs Prednisone

Filename	Name	Energy	Affinity	RerankScore	Torsions	RMSD
[00] 19649.mol2	[00] 19649	-147.594	-17.7441	-75.1455	9	97.9086
[01] 19649.mol2	[01] 19649	-139.914	-12.6839	-32.2553	9	98.6107
[02] 19649.mol2	[02] 19649	-136.781	-17.8446	-66.5858	9	97.9113
[03] 19649.mol2	[03] 19649	-118.643	-15.495	-69.6922	9	97.5517
[04] 19649.mol2	[04] 19649	-115.742	-33.0819	-59.618	9	98.8068

Average Energy: -131.5348

Figure- 11. Molecular Docking UPM vs Streptomycin



Average Energy: -140.2142

Figure- 12. Molecular Docking UPM vs Rifampin

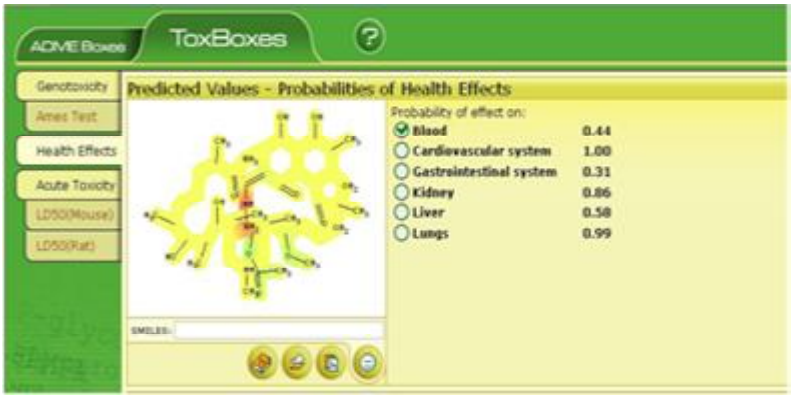


Figure- 13. ADMET studies of effective drug by ADME Boxes Tox Boxes

References :

1. Anand P, S Sankaran, S Mukherjee, K Yeturu, R Laskowski, and A Bhardwaj, *et al.* (2011). *PLoS ONE* 6(10): e27044.
2. Bairoch, Amos and Boeckmann Brigitte (1993). *Nucleic acids research 21:* 3093-6.
3. Bhramar Dutta, Aparna Banerjee, Priyanka Chakraborty and Rajib Bandopadhyay (2018). *J Genet Eng Biotechnol 16*(2): 749–756.
4. Christian Sohlenkamp and Otto Geiger (2016). *FEMS Microbiology Reviews 40*(1): 133–159.
5. Cole S.T. (1998). *Nature 393:* 537–544.
6. Dennis W. Grogan and John E. Cronan (1997). *Microbiology And Molecular Biology Reviews 61*(4): 429–441.
7. Guilhot C., B. Gicquel and C. Martín (1992). *FEMS Microbiol Lett. 77:* 181–186.
8. Houben RMGJ and PJ Dodd (2016). *PLoS Med. 13:* e1002152.
9. Issar Smith (2003). *Clin Microbiol Rev. 16*(3): 463–496.
10. Leonardo G. Ferreira, Ricardo N. dos Santos, Glaucius Oliva, and D. Adriano Andricopulo (2015). *Molecules 20*(7): 13384–13421.
11. Lipinski C.A., F. Lombardo, B.W. Dominy and P.J. Feeney (2012). *Adv. Drug Deliv. Rev. 64:* 4–17.
12. N. V. Pradeep, S. Anupama, K. Ankitha

and J. Pooja (2012). *Advances in Life Science and Technology 4:* 27-34.

13. Peter W. Rose, Andreas Prlić, Ali Altunkaya, Chunxiao Bi, Anthony R. Bradley, Cole H. Christie, *et al.* (2017). *Nucleic Acids Res 45:* D271–D281.
14. Rezaei Marzieh, Rabbani Khorasgani Mohammad, Zarkesh-Esfahani Sayyed Hamid, Emamzadeh Rahman and Abtahi Hamid (2019). *Infectious Disorders - Drug Targets 18*(1): 10.2174/1871526518666180709121653.
15. Varalakshmi D. Vissa, Rama Murthy Sakamuri, Wei Li, Patrick J. Brennan (2009). *Indian J Microbiol. 49*(1): 11–47.
16. Vattipally B Sreenu, Pankaj Kumar, Javaregowda Nagaraju and Hampapathalu A Nagarajaram (2006). *BMC Genomics 7*(78).
17. Vivek Rao, Feng Gao, Bing Chen, William R. Jacobs Jr. and Michael S. Glickman (2006). *J Clin Invest 116*(6): 1660–1667.