ISSN: 0975-766X
CODEN: IJPTFI
Available Online through
www.ijptonline.com
Research Article

ROLE OF BIOINFORMATICS TOOLS IN IN-SILICO DRUG DESIGNING ON SARS AND TOXICITY PREDICTION

Pooja Goyal*
M.Sc Bioinformatics, Jaipur National University Jagatpura, Jaipur, Pin code- 302017
Email: poojabioinfo91@gmail.com

Received on 12-07-2013 *Accepted on 30-07-2013*

Abstract

Drug designing is an inventive process of finding new medications based on the biological properties of drug targets. These targets may be proteins, enzymes, lipid etc. Here, we discussed about structure based drug designing (in silico approach) where target molecule structure should be known.

In this project design a drug for severe acute respiratory syndrome which is caused by SARS coronavirus. Main proteinase enzyme of SARS coronavirus is consider as target (cause) for this disease. A new ligand molecule is also designed to inhibit the potency of target molecule. Here, Dimethyl sulfoxide is considered as the lead because it has the certain biological properties to inhibit the potency of main proteinase enzyme after binding on the active site of target molecule.

Several tools and softwares like: blast, procheck, ligsite, ligbuilder, modeler, swiss pdb viewer, rasmol, quantum play an important role in designing of a new drug based on in silico approach.

Key words: Basic local alignment search tool, Hydrogen bond acceptor, Hydrogen bond donor, Main proteinase enzyme, mutagenicity, Polar surface area.

Introduction

Severe acute respiratory syndrome is a respiratory disease in humans which is caused by the SARS coronavirus .The severe-acute-respiratory-syndrome coronavirus genome consists of 28 putative open reading frames in 9 mRNA transcripts. ORF1a and ORF1b, which account for about two-thirds of the genome, both encode large polyproteins**.**

First of all, understanding the mechanism of disease is very necessary to identify the disease causing agent because structure based drug designing approach totally based on the structure of target.

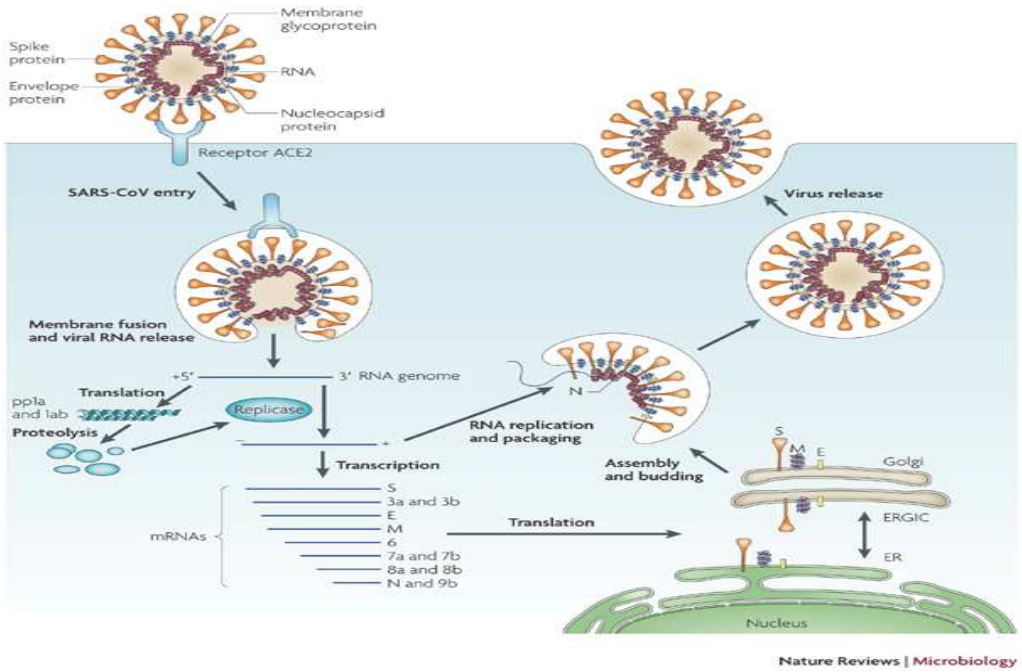


Fig.1-Mechanism of action of SARS coronavirus

A coronavirus virion lands on the cell surface at upper right and binds to cell surface receptors using its projecting spikes.

- It then enters the cell by membrane fusion with the plasma membrane or by receptor mediated endocytosis (a clathrin coated vesicle is seen floating in the region).
- The viral genome (which is +ve ss RNA, drawn as a yellow string) enters the cytoplasm.
- The genomic RNA creates a -ve version of itself and so forms a replication complex attached to membrane .
- These replication complexes then produce new genomic RNAs and (subgenomic) mRNAs (1, 5, 7).
- mRNAs code for new viral proteins, including N, M & S. The N protein joins the new genomic RNA to form new RNPs (ribonucleoproteins).
- These RNPs attach to the membrane where S (Spike) proteins and M proteins have previously assembled.
- The RNP buds into the lumen of the vesicle (lower right) and finally the membrane bound RNP and its radiating spikes detaches and comes to lie free in the lumen as an immature virion.
- These particles progress up the periphery of the Golgi apparatus. The new virus particles collect in large vesicles and are finally released onto the cell surface to start the cycle again.

Materials and Methods

Several bioinformatics tools are used in designing of a drug like-BLAST, PDB, modeler, hex, ligsite, ligbuilder, quantum, swiss pdb viewer etc.

Methodology

1. Target Identification

The major cause of the SARS is main proteinase enzyme of SARS coronavirus. After translation, proteolysis is initiated by main proteinase of SARS coronavirus.

Through NCBI the information about main proteinase of SARS coronavirus is accessed and its FASTA format is retrieved:-

>gi|39654933|pdb|1UK3|A Chain A, Crystal Structure Of Sars Coronavirus Main Proteinase (3clpro) At Ph7.6

SGFRKMAFPSGKVEGCMVQVTCGTTTLNGLWLDDTVYCPRHVICTAEDMLNPNYEDLLIRKSNHSFLVQAG

NVQLRVIGHSMQNCLLRLKVDTSNPKTPKYKFVRIQPGQTFSVLACYNGSPSGVYQCAMRPNHTIKGSFLNGS

CGSVGFNIDYDCVSFCYMHHMELPTGVHAGTDLEGKFYGPFVDRQTAQAAGTDTTITLNVLAWLYAAVING

DRWFLNRFTTTLNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAVLDMCAALKELLQNGMNGRTILGSTIL

EDEFTPFDVVRQCSGVTFQ

2. Target validation

Target validation is done through performing BLAST against the PDB database and a local similarity search is performed.It gives information about the conserved domains present in the sequence. In the result given below PDB id 2A5K A has the maximum similarity with our target sequence.

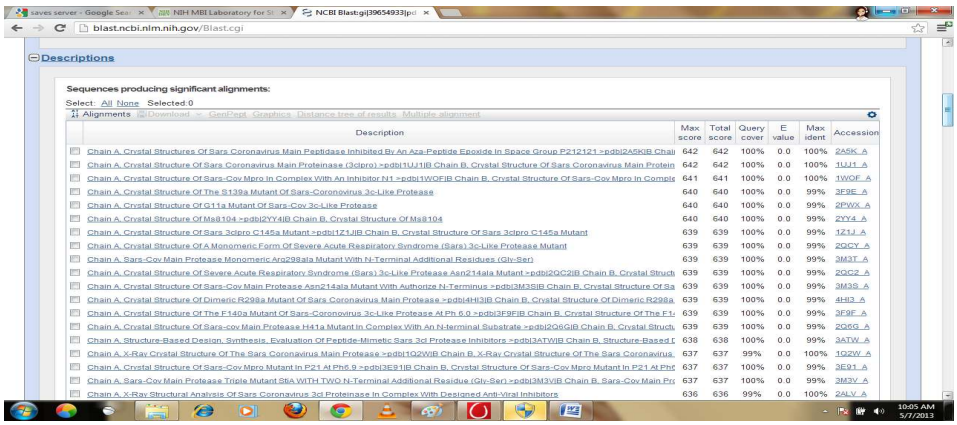


Fig.2-BLAST result represents homologous sequences

3. Structural retrieval

The structure of main proteinase is retrieved through PDB databank with the PDBid 2A5K.If structure is present in PDB. Then, retrieve it otherwise homology modeling is performed.

Homology modelling is performed through modeller prompt,through which 5 pdb files are generated.which undergo structure analysis and validation through the SAVES server.

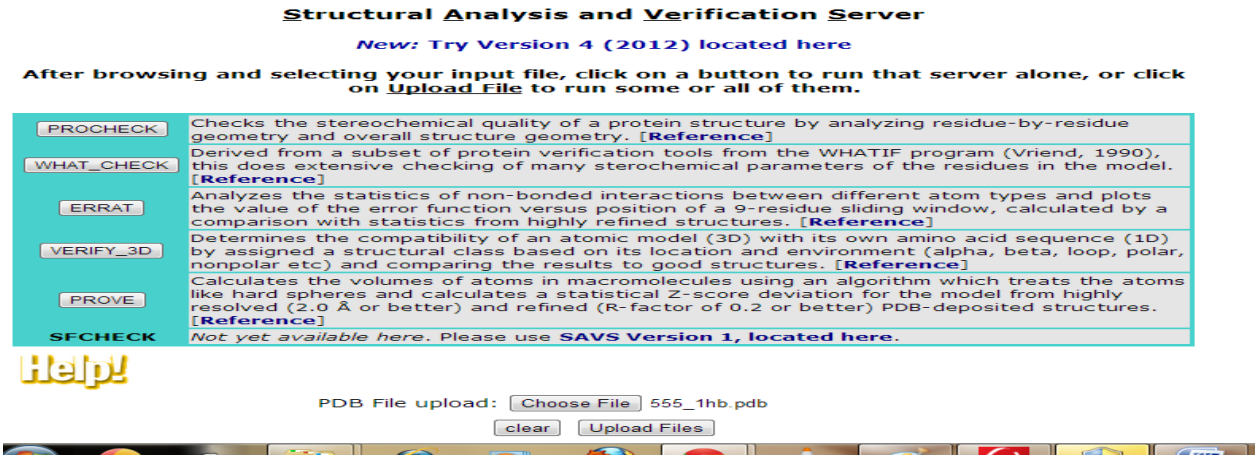


Fig.3 Saves server home page

Through the tool PROCHECK we can find out which structures need refinement.Those structures are refined in SPDBV ,in which they are removed from the disallowed region through loop building and bad contacts are removed through energy minimisation**.**

4. Best model selection

The best model selection is done on the basis of a high core value and G-factor,zero bad contacts and no amino acids should be present in the disallowed region.

Fig.4: Saves result for input file generated by modeler

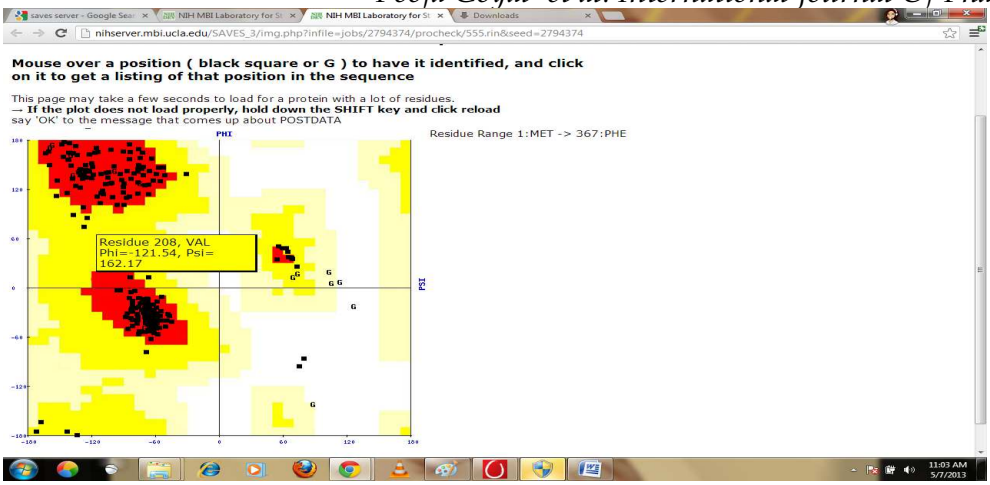


Fig.5: Ramachandran plot for input file

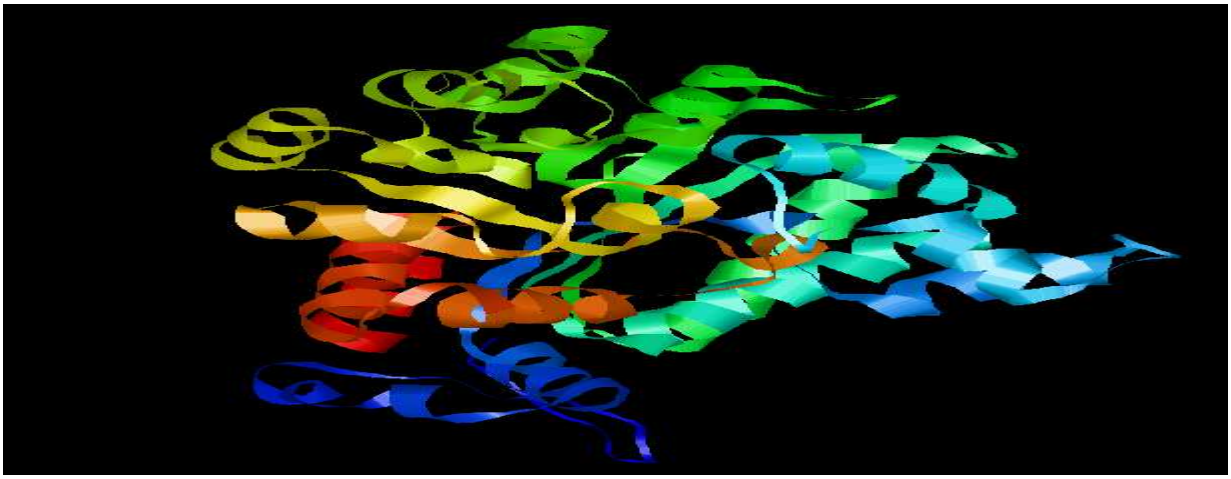


Fig.6: Final model of target sequence in Rasmol.

5. Active site identification

The active site identification is done through ligsite,which is used to find out the pockets of the best model.The minimum distance from the pocket is found out, the atom is PHE294 CD2 with a distance of 2.2 angstrom.

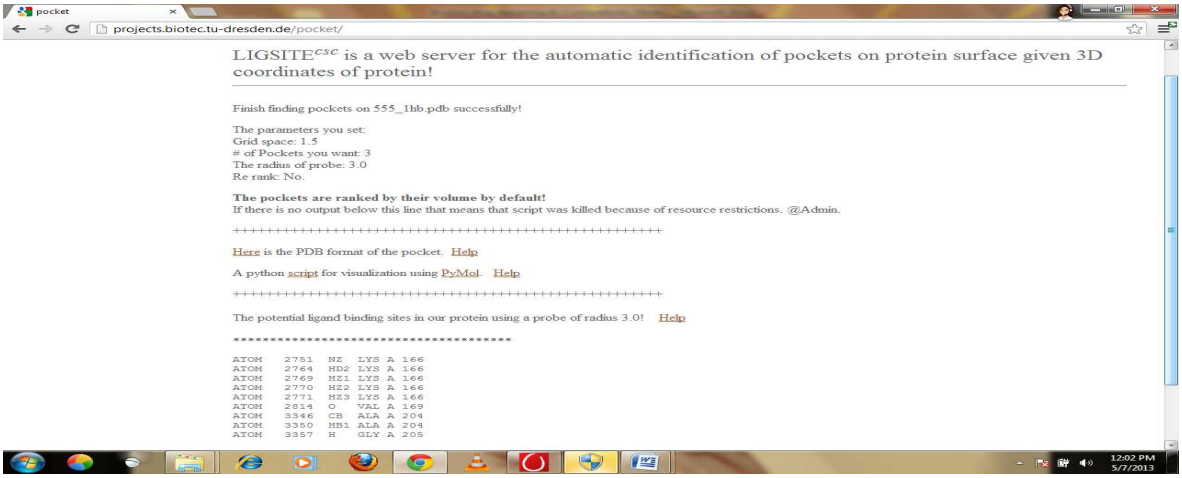


Fig.7: Pockets found by Ligsite

6. lead identification

For the lead identification , ligands (inhibitor) are benzotriazole esters i.e. benzene carbothioic-s-acid, benzoic acid, dimethyl sulfoxide.

Lead optimization follows the certain properties i.e.-

- It must be small.
- It must contain as many as possible growing site.
- It should not contain any metal atom.
- It should not contain unsaturatedness.

Dimethyl sulfoxide is considered as the lead .Its structure is then drawn in chemsketch and is converted to the PDB file using OPEN BABEL.

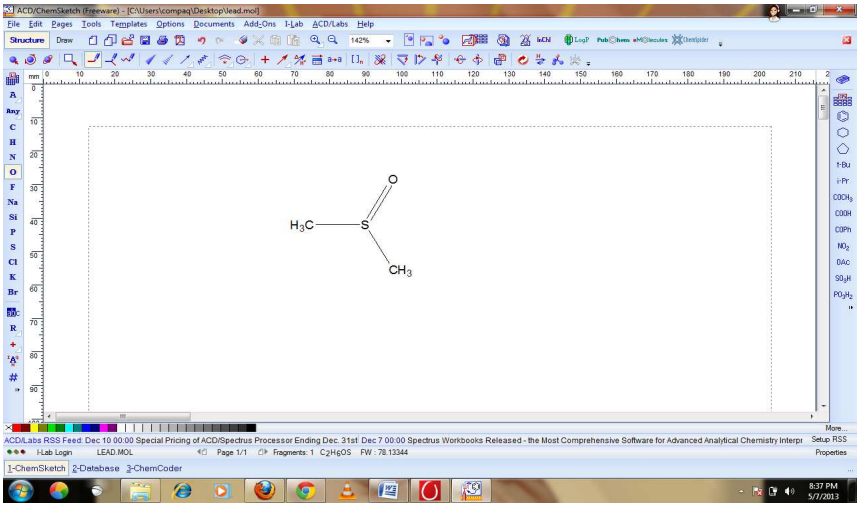


Fig.8: Lead molecule in chemsketch

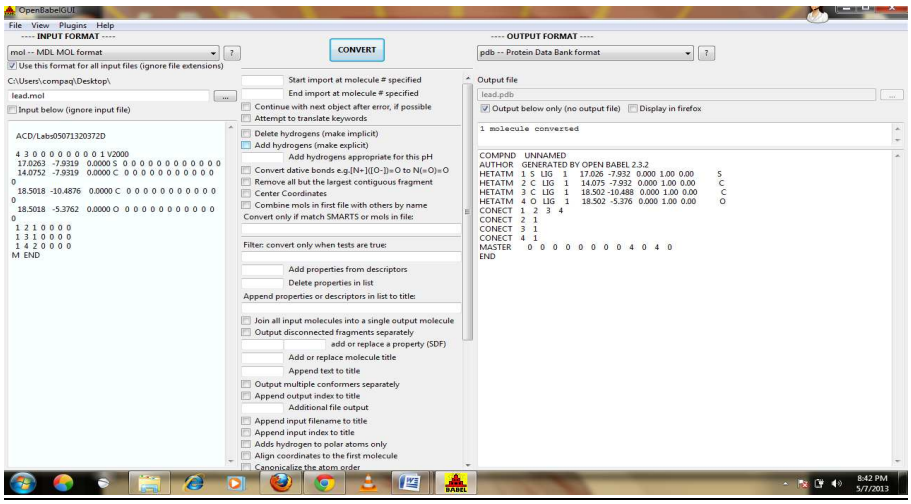


Fig.9: Lead mol file conversion by openbabelGUI

7. Lead introduction into the cavity

The lead introduction into the cavity is done through hex.The structure of protein along with the ligand is then visualised in SPDBV.

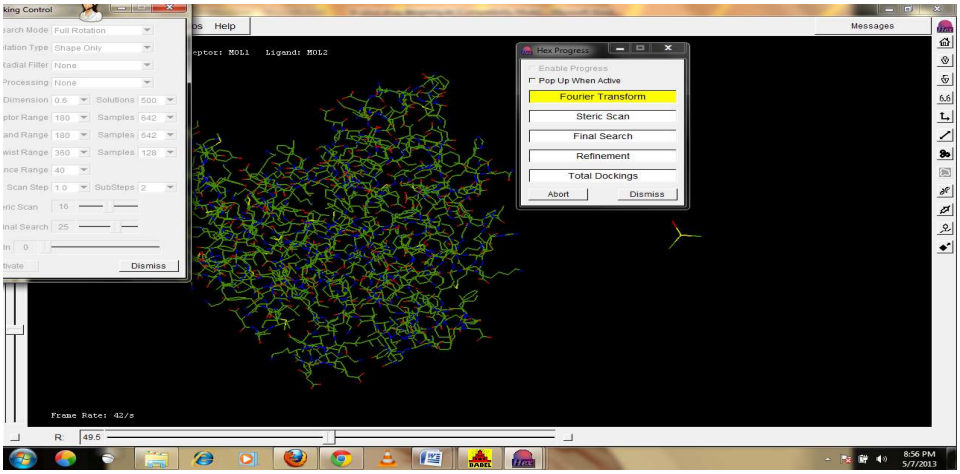


Fig.10: Hex result

The growth of the lead is performed through LIGBUILDER which generates the ligand and population file along with 5 resulting structures.

8. Docking

To explore the binding affinity,docking is performed through the software quantum. For docking , 5 files generated by ligbuilder are used as input. From these docking files we can determine the gibbs free energies as well as the rotb anr RMS,A values on the basis of which we can determine the 5 best structures as they should have minimum gibbs free energy.

Table-1 & 2:

NAME	PROTEIN	LIGAND	IC50,mol/l	Gbind,KJ/mol	E es,KJ/mol	E vdw,KJ/…	E vdw,KJ/…
Docking	2p_1hb	lig_pres1	8.86E-05	-23.58	-0.92	-40.48	8.78
Docking	2p_1hb	lig_pres2	2.59E-05	-26.7	-0.57	-39.89	-7.62
Docking	2p_1hb	lig_pres3	6.72E-05	-24.28	7.62	-38.97	-8.09
Docking	2p_1hb	lig_pres4	6.72E-05	-22.46	-1.3	-38.74	8.87
Docking	2p_1hb	lig_pres5	3.84E-05	-25.7	0.6	-38.26	-8.58

Etor, KJ/mol	Gprot	Charge	mass	flex.bond	RMS,A
26.6	0	0	493	1	46.08
6.15	0	1	465	1	44.51
-1.02	0	0	454	0	46.91
26.46	0	0	418	2	45.9
3.38	0	0	384	0	45.61

The next step is performed to ascertain the effects of the possible drug molecule on the various systems of the body as well as its toxicity.It uses various softwares such as OSIRIS, MOLSOFT,PASS, PHARMA ALGORITHMS, TOXTREE, CHEMSPIDER, most of which are online.

Toxicity prediction-

Osiris property explorer

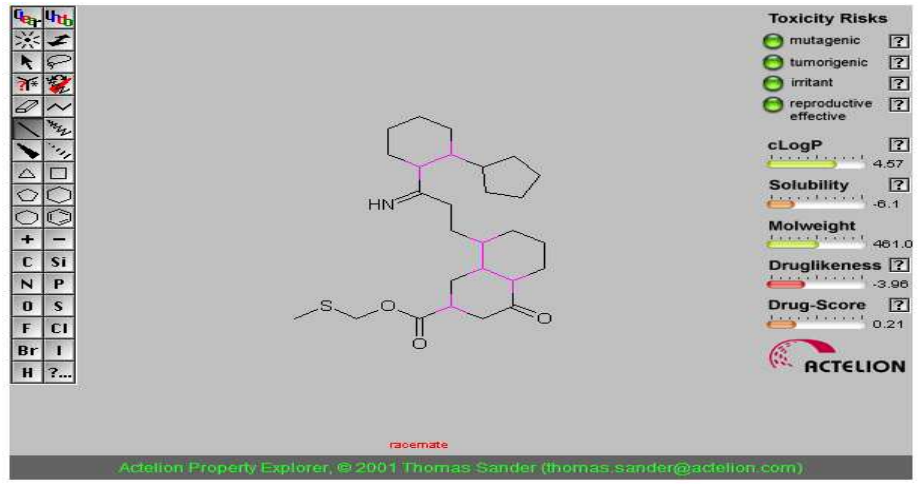


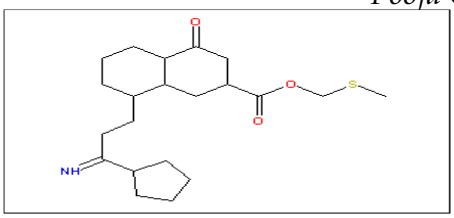
Fig.29: results predicted by Osiris property explorer

Result shows- mutagenic, tumorigenic, irritant, reproductive effective properties of drug molecule.

It also shows cLogP value, solubility, molweight, druglikeness and drug score.This software shows that the drug molecule obtained, is less effective along with some irritants.

2. Molsoft LLC

The result obtained using this software is in graphical form,which contains blue & green coloured peaks,in which blue colour indicates the drugs.That molecule can be considered as drug while green colour shows non drug.



Molecular formula: C21 H33 N O3 S
Molecular weight: 379.22
Number of HBA: 5
Number of HBD: 1
MolLogP : 4.30
MolLogS : -3.43 (in Log(moles/L)) 141.96 (in mg/L)
MolPSA : 53.03 A^2
MolVol : 408.25 A^3
Number of stereo centers: 4

Drug-likeness model score: -0.57

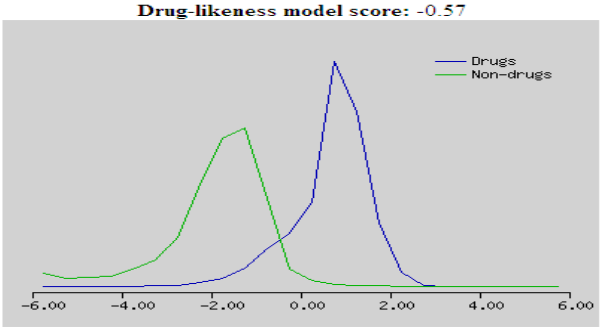


Fig.30: molsoft LLC results

Results and Discussion

On performing the toxicity prediction step for the molecules obtained from quantum software we can conclude that the drug is non mutagenic and non tumerogenic because in Osiris result it indicates green color.

According to the molsoft results, molecular weight of the drug is 379.22.Hydrogen bond acceptors are 5 and hydrogen bond donor is1. Log P value of this drug is 4.30.

This drug follows the Lipinski rule of five.

It has molecular weight less than 500 dalton.

It has water partition coeffitient less than five.

It has hydrogen bond donors less than five and not more than five hydrogen bond acceptors.

Molecular polar surface area of this drug is 53.03 A2. If any drug has less than 60 A square polar surface area means it can easily penetrate blood brain barrier.

Conclusion

This drug will become an anti-infecting agent for severe acute respiratory syndrome having greater safety and efficacy.

Acknowledgements

The author would like to express sincere thanks to the Management of Jaipur National University for her encouragement and providing necessary facilities in carryout this research work.

References

1. Marco A. Marra Steven J. M. Jones,Caroline R. Astell, Robert A. Holt , Angela Brooks-Wilson (2003). The Genome Sequence of the SARS-Associated Coronavirus.science 300:1399-1404
2. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990): Basic local alignment search tool. *J Mol Biol* 215:403–10.
3. Blundell, T.L., Sibanda, B.L., Sternberg, M.J.E., and Thornton, J.M. (1987) Knowledge-Based Prediction of Protein Structures and the Design of Novel Molecules. Nature 326:347-352.
4. Chinea G, Padron G, Hooft RWW, Sander C, Vriend G (1995): The use of position specific rotamers in model building by homology. *Proteins* 23:415–21.
5. Dayringer HE, Tramontano A, Fletterick RJ (1986): Interactive program for visualization and modelling of proteins, nucleic acids and small molecules. *J Mol Graph* 4:82–7.
6. Fiser A, Do RK, Sali A (2000): Modeling of loops in protein structures. *Protein Sci* 9:1753–73.

Corresponding Author:

Pooja Goyal*,

Email: *poojabioinfo91@gmail.com*