Establishing Novel Drug Leads For Bubonic Plague Using *In-Silico* Approach

Pratik Dherange*, Shraddha Lavhale, Preenon Bagchi

Institute of Biosciences and Technology, MGM University, Chht.

Sambhajinagar, India

pratikdherange91@gmail.com

Abstract. Bubonic plague is an infection spread mostly to humans by infected fleas that feed on rodents which is also known as 'Black death'. It killed millions of Europeans during the middles ages. Bubonic Plague is an infectious disease caused by a specific type of bacterium called Yersinia pestis. Which is also know as Y.pestis can affect humans spread mainly by fleas. Prevention doesn't include a vaccine, but does involve reducing your exposure to mice, rats, squirrels and other animals that may be infected. The deaths exceeded 25 million people during the middle ages. The plague is rare now. Only a few thousand people around the world get it each year most of the cases are in Africa, India.

The FASTA sequence of the genes-receptors of the above disease were retrieved from Genbank database. Next, we modelled the 3D structure of the FASTA protein sequence using modeller. The best model was selected using Ramachandran plot. Phytocompounds from medicinal plants is considered Emblica Officinalis, Terminalia Chebula, Terminalia Bellirica, Asphaltum Punjabianum, Azadirachta Indica as considered as novel drug leads is retrieved from PUBCHEM database. The phytocompounds are checked for drug-like properties using molinspiration software. The compounds having no violations was considered for further docking studies.

Thannilignan, Ellagic acid, Quercetin, Termilignan docks best with both the receptors. Hence these phytocompounds can be used as ligands for the Bubonic Plague.

Keywords: Bubonic Plague, Yersinia pestis, Bioinformatics, Docking, Phytocompound.

1 Introduction

Bubonic plague is an infection spread mostly to and by humans by infected fleas which travel on rodents. These rodents are also known as 'black death' since they killed millions of Europeans [1]. This infectious disease is caused by Yersinia pestis [2] (Fig. 1). This plague is rare now but each year most of the cases are reported in India [3].

R. Somashekhar et al. (eds.), *Proceedings of the International Conference on Advances in Nano-Neuro-Bio-Quantum (ICAN 2023)*,
Advances in Intelligent Systems Research 69,
https://doi.org/10.2991/978-94-6463-294-1_11

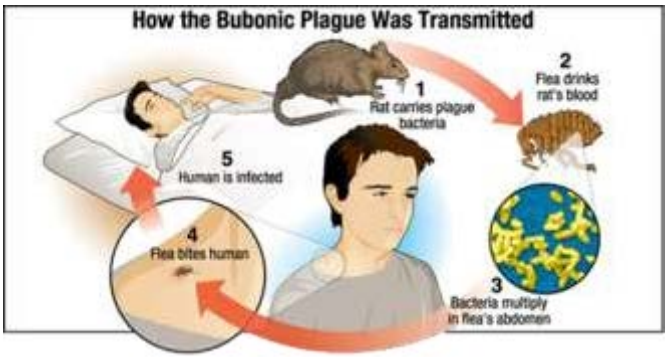


Fig. 1: Bubonic plague transmission (reproduced from https://www.nextgurukul.in/questions-answers-forum/question/academic/bubonic-plague/29980)

Genes involved

1. T3SS:

Type 3 Secretion Systems (T3SSS) are complex bacterial structures, that enables some bacteria to directly inject effector proteins into host cells, facilitating colonization [4]. Yersinia pestis, the causative agent of plague, possesses a number of virulence mechanisms that allows it to survive and proliferate during its interaction with the host. Yersinia pestis, the etiologic agent of plague, causes a variety of serious diseases in humans and animals. The clinical syndromes in humans include bubonic, pneumonic, and septicemic plague[5].

2. RovA:

The pathogenic species of yersinia contain the transcriptional regulator RovA[6]. RovA regulates expression of the invasion factor invasin (inv), which mediates translocation across the intestinal epithelium[7].

research to any language, breaking down language barriers and enabling more effective collaboration.

2. Materials and Methodology

The FASTA sequence of the T3SSA and ROVA gene receptor was retrieved from Genbank. Their 3d structure was modeled using modelled using modeler[8]. The phytocompounds of the plants Emblica officinalis, Terminalia chebula, Terminalia bellirica, Asphaltum punjabianum and Azadirachta indica were selected and their SMILES were retrieved from PubChem database. The selected compounds were docked using the best receptor model.

3. Results

The receptors' amino acid sequence were retrieved from Genbank and their homologous templates selected using BLAST & download from RCSB PDB (Table 1).

Table 1: Receptors' Genbank Accession number with their homologous templates.

Gene	Genbank Accession number	Homologous templates
RovA	OW971834.1	5AOXF
T3SS	Q7AR18.1	6NJOA 2OBLA 4NPHA

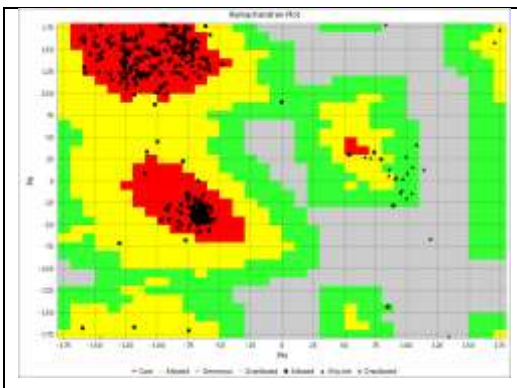
The receptors' 3d structures were modeled using Modeller. Modeller generated 5 models (Table 2). The models were verified using VADAR Ramachandran Plot server [9] (Fig. 2).

Table 2(a): Ramachandran plot analysis of RovA

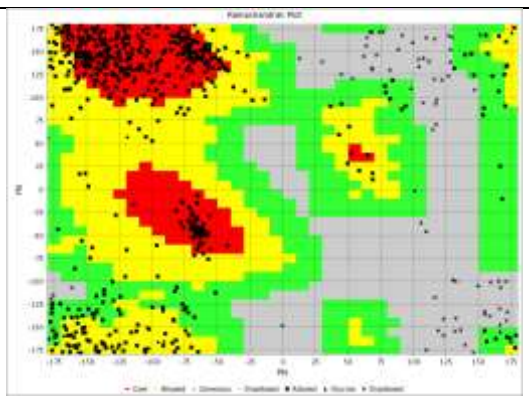
	# res in phipsi core	# res in phipsi allowed	# res in phipsi generous	# res in phipsi outside	
T3SS.B99990001	421 (95%)	16 (3%)	2 (0%)	0 (0%)	Selected
T3SS.B99990002	415 (94%)	17 (3%)	4 (0%)	3 (0%)	
T3SS.B99990003	423 (96%)	11 (2%)	4 (0%)	1 (0%)	
T3SS.B99990004	418 (95%)	19 (4%)	2 (0%)	0 (0%)	
T3SS.B99990005	414 (94%)	22 (5%)	2 (0%)	1 (0%)	

Table 2(b): Ramachandran plot analysis of T3SS

TEMPLATE	# res in phipsi core	# res in phipsi allowed	# res in phipsi generous	# res in phipsi outside	
ROVA.B99990001	356 (52%)	187 (27%)	81 (11%)	58 (8%)	
ROVA.B9999000	432 (63%)	144 (21%)	49 (7%)	57 (8%)	
ROVA.B99990003	325 (7%)	170 (24%)	94 (13%)	93 (13%)	
ROVA.B99990004	349 (51%)	166 (24%)	86 (12%)	81 (11%)	
ROVA.B99990005	405 (59%)	158 (23%)	68 (9%)	51 (7%)	selected



RovA



T3SS

Fig. 2: Ramachandran Plot of RovA & T3SS

PLANT : EMBLICA OFFICINALIS	
PHYTOCOMPOUND	
ASCORBIC ACID	C([C@@H]([C@@H]1C(=C(C(=O)O1)O)O)O)O
ELLAGIC ACID	C1=C2C3=C(C(=C1O)O)OC(=O)C4=CC(=C(C(=C43)OC2=O)O)O
RUTIN	C[C@H]1[C@@H]([C@H]([C@H]([C@@H](O1)OC[C@@H]2[C@H]([C@@H]([C@H]([C@@H](O2)OC3=C(OC4=CC(=CC(=C4C3=O)O)O)C5=CC(=C(C=C5)O)O)O)O)O)O)O)O
QUERCETIN	C1=CC(=C(C=C1C2=C(C(=O)C3=C(C=C(C=C3O2)O)O)O)O)O
CATECHOL	C1=CC=C(C(=C1)O)O

Plant: *TERMINALIA CHEBULA*	
Phytocompound	SMILES
CHEBULIC ACID	C1=C2C(=C(C(=C1O)O)O)[C@@H]([C@H](OC2=O)C(=O)O)[C@H](CC(=O)O)C(=O)O
GALLIC ACID	C1=C(C=C(C(=C1O)O)O)C(=O)O
ELLAGIC ACID	C1=C2C3=C(C(=C1O)O)OC(=O)C4=CC(=C(C(=C43)OC2=O)O)O
EUGENOL	COC1=C(C=CC(=C1)CC=C)O
METHYL GALLATE	COC(=O)C1=CC(=C(C(=C1)O)O)O

PLANT : TERMINALIA BELLIRICA	
PHYTOCOMPOUND	SMILES
TERMILIGNAN	COC1=CC(=C(C=C1)CC(=C)C(=C)CC2=CC=C(C=C2)O)O
THANNILIGNAN	COC1=CC(=C(C=C1)CC(CO)(C(=C)CC2=CC=C(C=C2)O)O)O
ELLAGIC ACID	C1=C2C3=C(C(=C1O)O)OC(=O)C4=CC(=C(C(=C43)OC2=O)O)O
CHEBULIC ACID	C1=C2C(=C(C(=C1O)O)O)[C@@H]([C@H](OC2=O)C(=O)O)[C@H](CC(=O)O)C(=O)O

PLANT : ASPHALTUM PUNJABIANUM	
PHYTOCOMPOUND	
FLAVAN	C1CC2=CC=CC=C2OC1C3=CC=CC=C3

ISOFLAVONE	C1=CC=C(C=C1)C2=COC3=CC=CC=C3C2=O
QUERCETIN	C1=CC(=C(C=C1C2=C(C(=O)C3=C(C=C(C=C3O2)O)O)O)O)O
MYRICETIN	C1=C(C=C(C(=C1O)O)O)C2=C(C(=O)C3=C(C=C(C=C3O2)O)O)O

PLANT : AZADIRACHTA INDICA	
PHYTOCOMPOUND	
LIMONOIC ACID	C[C@]1(CC[C@H]2[C@]([C@@]13[C@H](O3)C(=O)O)(C(=O)C[C@@H]4[C@@]2([C@@H](OC4(C)C)CC(=O)O)CO)C)[C@H](C5=COC=C5)O
GLYCERIDES	CCCCCCCCCCCCCCCCCC(=O)OC(CO)COC(=O)CCCCCCCCCCCCCCCC
BETA- SITOSTENOL	CC[C@H](CC[C@@H](C)[C@H]1CC[C@@H]2[C@@]1(CC[C@H]3[C@H]2CC=C4[C@@]3(CC[C@@H](C4)O)C)C)C(C)C
QUERCETIN	C1=CC(=C(C=C1C2=C(C(=O)C3=C(C=C(C=C3O2)O)O)O)O)O

Table 3: SMILES notation of the phytocompounds

Further, the phytocompounds of the plants Emblica officinalis, Terminalia chebula, Terminalia bellirica, Asphaltum punjabianum and Azadirachta indica were selected and their SMILES were retrieved from PubChem (Table 3).

Using molinspiration the drug-likeliness of the above compounds were ascertained (Table 4).

Phytocompound	miLogP	TPSA	natoms	MW	nON	nOHNH	nrotb	volume	Nviolations
ASCORBIC ACID	-1.40	107.22	12	176.12	6	4	2	139.71	0
ELLAGIC ACID	0.94	141.33	22	302.19	8	4	0	221.78	0
RUTIN	-1.06	269.43	43	610.52	16	10	6	496.07	3
QUERCETIN	1.68	131.35	22	302.24	7	5	1	240.08	0
CATECHOL	0.99	40.46	8	110.11	2	2	0	100.08	0

CHEBULIC ACID	-1.14	198.89	25	356.24	11	6	5	273.51	2
GALLIC ACID	0.59	97.98	12	170.12	5	4	1	135.10	0
ELLAGIC ACID	0.94	141.33	22	302.19	8	4	0	221.78	0
EUGENOL	2.10	29.46	12	164.20	2	1	3	162.14	0
METHYL GALLATE	0.85	86.99	13	184.15	5	3	2	152.63	0
TERMILIGNAN	4.58	49.69	22	296.37	3	2	6	286.10	0
THANNILIGNAN	2.79	90.15	24	330.38	5	4	7	307.71	0
ELLAGIC ACID	0.94	141.33	22	302.19	8	4	0	221.78	0
CHEBULIC ACID	-1.14	198.89	25	356.24	11	6	5	273.51	2
FLAVAN	4.09	9.23	16	210.28	1	0	1	204.03	0
ISOFLAVONE	3.54	30.21	17	222.24	2	0	1	200.00	0
QUERCETIN	1.68	131.35	22	302.24	7	5	1	240.08	0
MYRICETIN	1.39	151.58	23	318.24	8	6	1	248.10	1
LIMONOIC ACID	0.48	167.03	36	506.55	10	4	6	446.14	1
GLYCERIDES	10.05	72.84	42	596.98	5	1	36	664.20	2

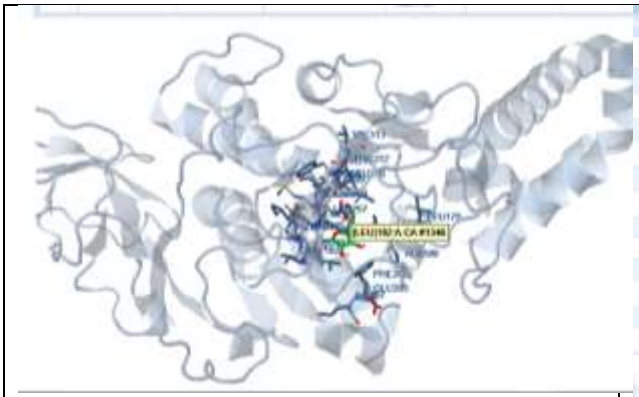
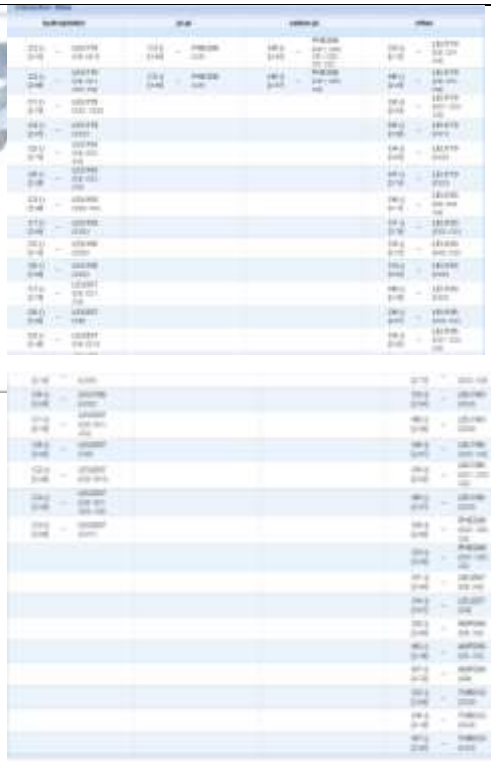
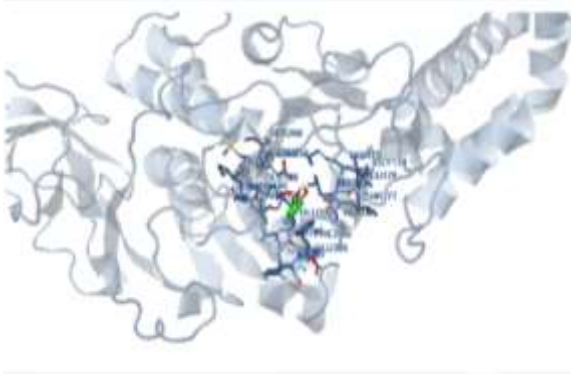
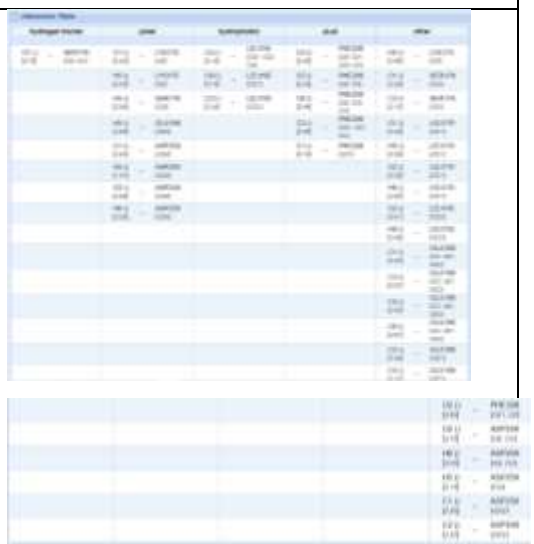
BETA-SITOSTENOL	8.62	20.23	30	414.72	1	1	6	456.52	1
QUERCETINS	1.68	131.35	22	302.24	7	5	1	240.08	0

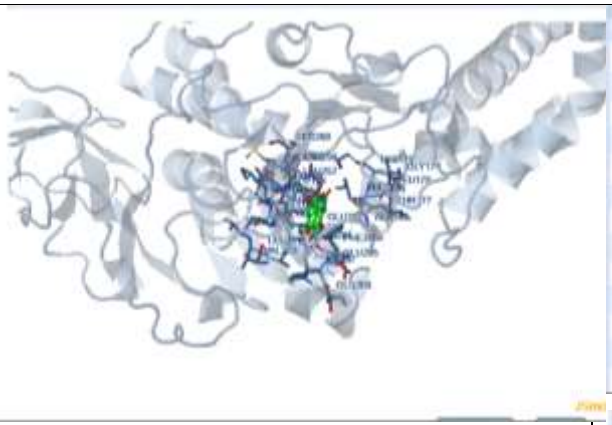
Table 4: ADME properties of phytocompounds using molinspiration

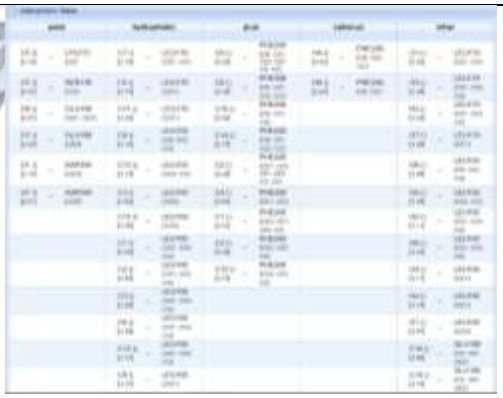
The compounds having nviolations 0 are selected and docked with the selected receptors [10] from Table 2 (Table 5, Fig. 3).

Phyto-componds	Docking score	No .of interactions	Docking yes/no
ASCORBIC ACID	-13.39kcaL/MOL	53	YES
CATECHOL	-0.61kcal/mol	38	YES
ELLAGIC ACID	-0.06Kcal/mol	81	YES
EUGENOL	-0.01Kcal/mol	39	YES
FLAVAN	-0.06kcal/mol	54	YES
GALLIC ACID	-0.82Kcal/mol	56	YES
ISOFLAVONE	-0.13Kcal/mol	50	YES
METHYL GALLATE	-1.15Kcal/mol	59	YES
QUERCETIN	-0.28Kcal/mol	80	YES
TERMILIGNAN	+0.39Kcal/mol	79	YES
THANNILIGNAN	+0.19Kcal/mol	90	YES

Table 5(a): Docking analysis of T3SS

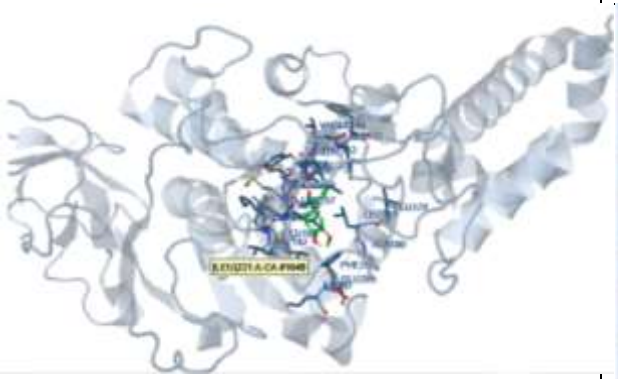
	
ASCORBIC ACID DOCKING	ASCORBIC ACID DOCKING INTERACTING AMINO ACIDS
	
CATECHOL ACID DOCKING	CATECHOL ACID DOCKING INTERACTING AMINO ACIDS

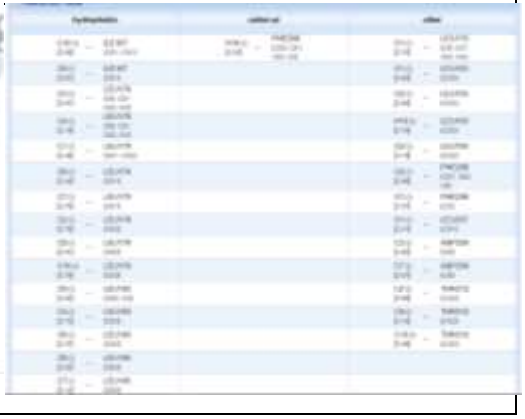


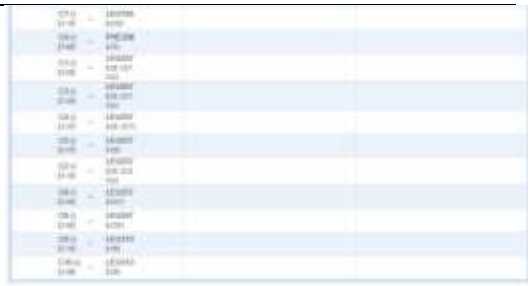
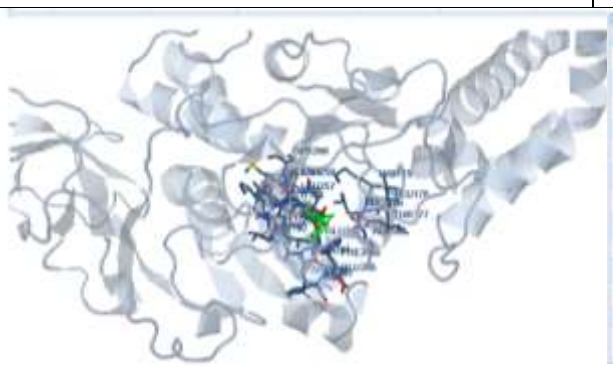
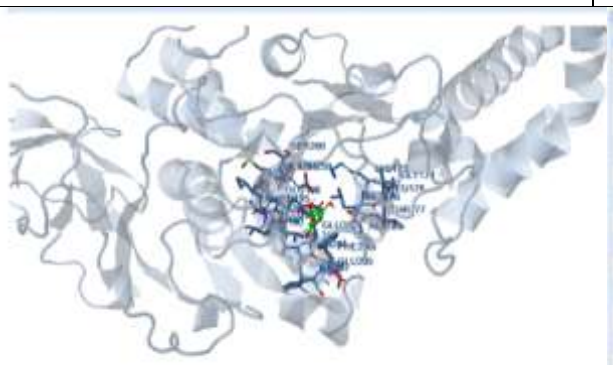


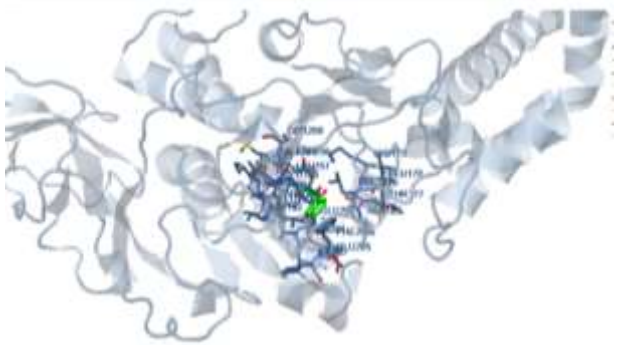
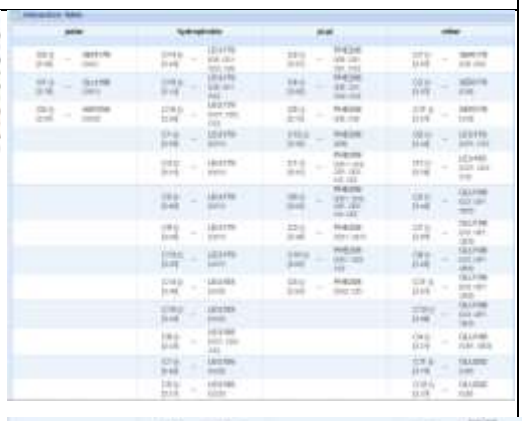
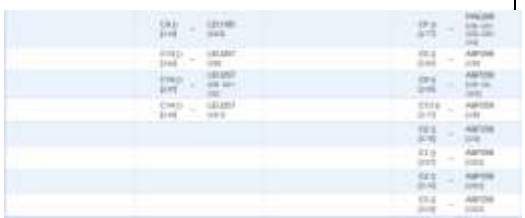
ELLAGIC ACID DOCKING

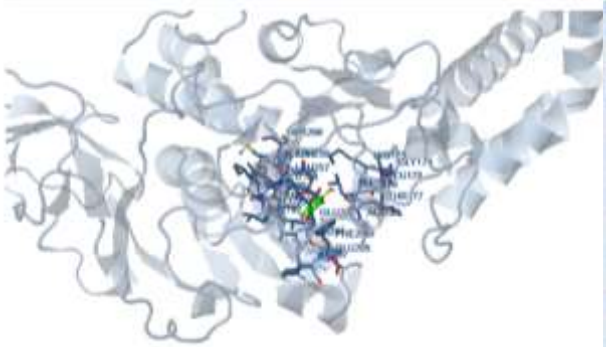
ELLAGIC ACID DOCKING INTERACTING AMINO ACIDS



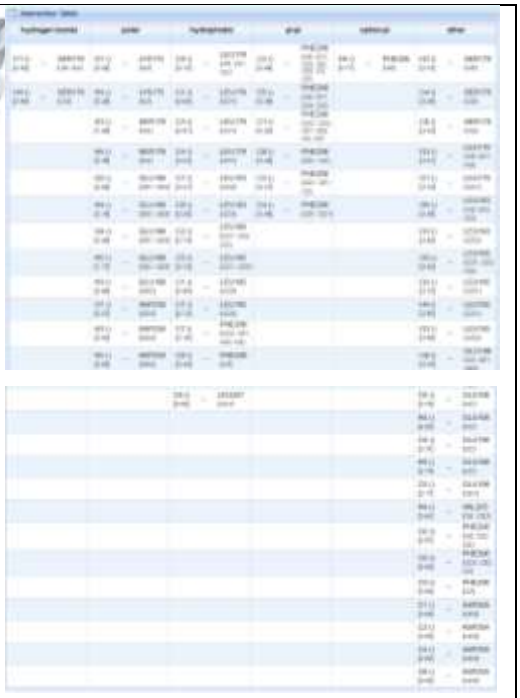


	
EUGENOL ACID DOCKING	EUGENOL ACID DOCKING INTERACTING AMINO ACIDS
	 
FLAVAN ACID DOCKING	FLAVAN ACID DOCKING INTERACTING TWO AMINO ACIDS
	

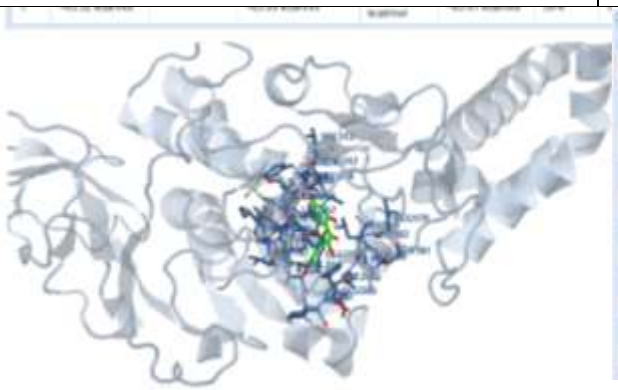
	
GALLIC ACID DOCKING	GALLIC ACID DOCKING INTERACTING AMINO ACIDS
	 
ISOFLAVONE ACID DOCKING	ISOFLAVONE ACID DOCKING INTERACTING AMINO ACIDS

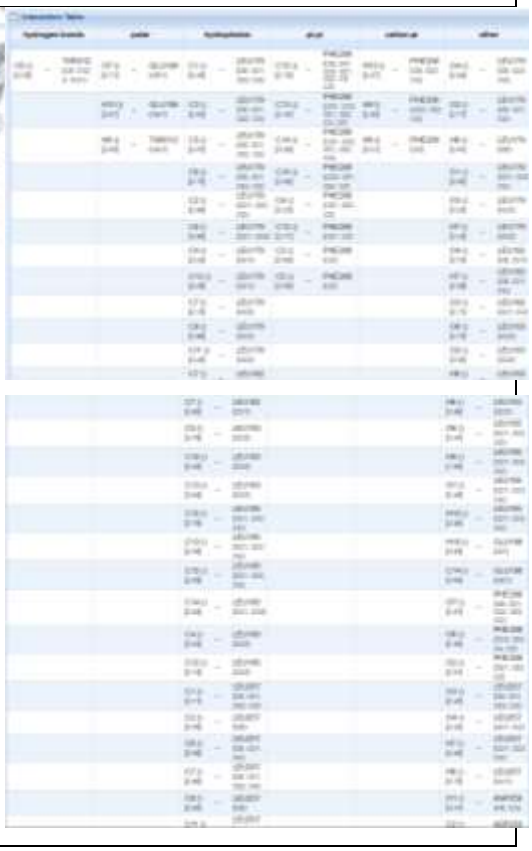


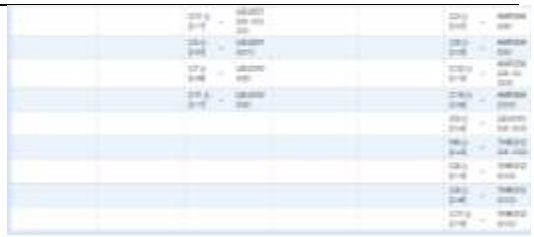
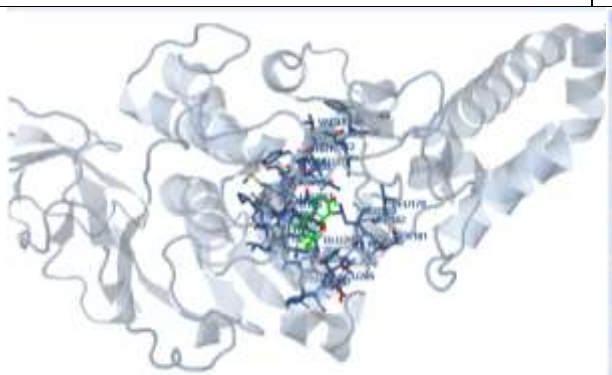
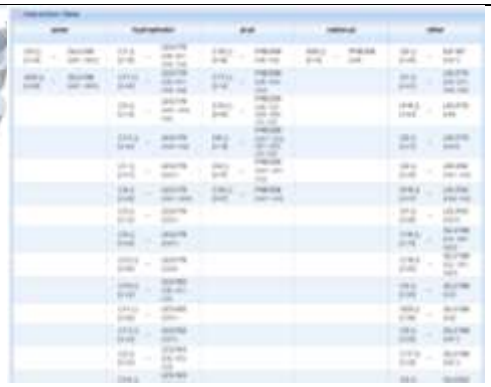
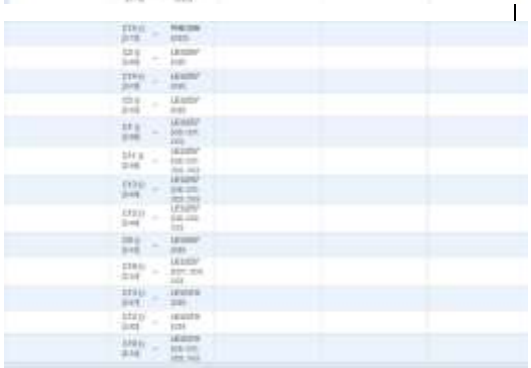
METHYL GALLATE ACID DOCKING

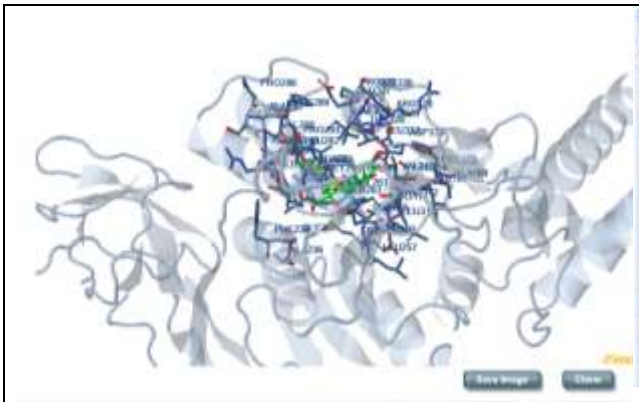


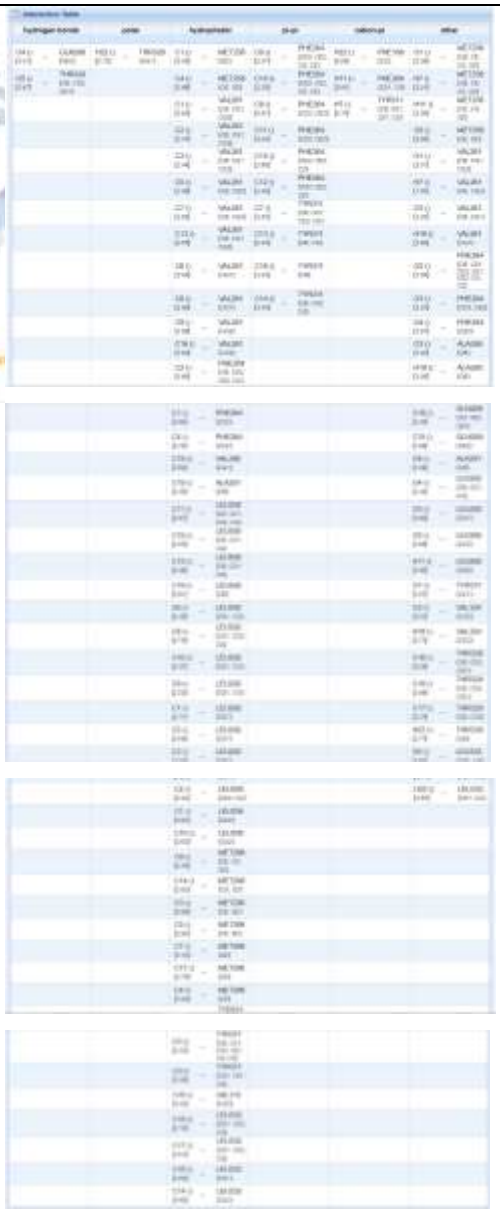
METHYL GALLATE ACID DOCKING INTERACTING AMINO ACIDS





	
QUERCETIN ACID DOCKING	QUERCETIN ACID DOCKING INTERACTING AMINO ACIDS
	  
TERMILIGNAN ACID DOCKING	TERMILIGNAN ACID DOCKING INTERACTING AMINO ACIDS





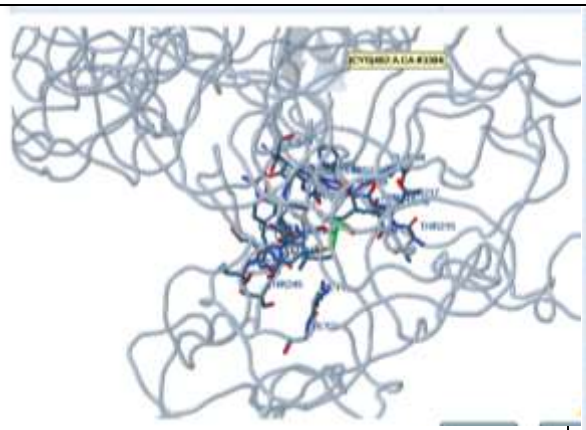
THANNILIGNAN ACID DOCKING	THANNILIGNAN DOCKING INTERACTING AMINO ACIDS

Fig. 3(a): Docking results of T3SS.

Table 5(b): Docking analysis of ROVA

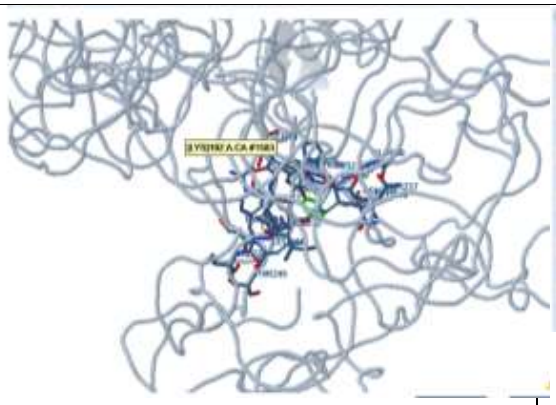
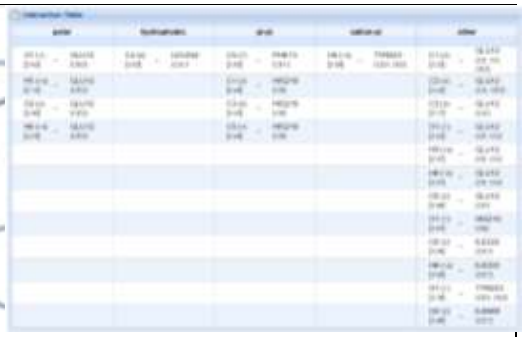
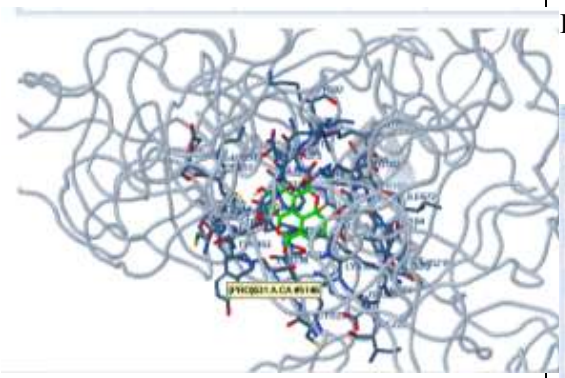
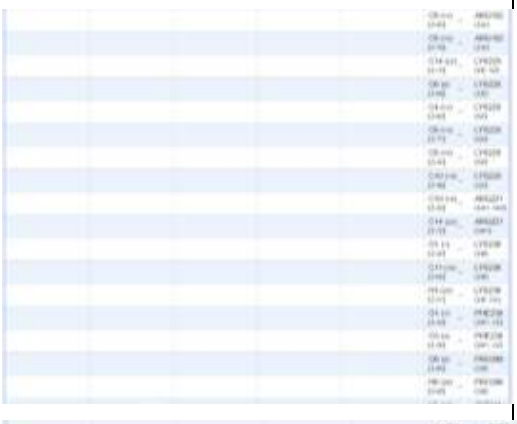
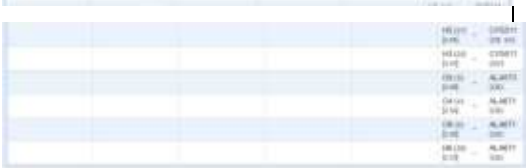
Phyto-componds	Docking score	No .of interactions	Docking yes/no
ASCORBIC ACID	-2.83kcal/mol	32	YES
CATECHOL	-4.89kcal/mol	22	YES

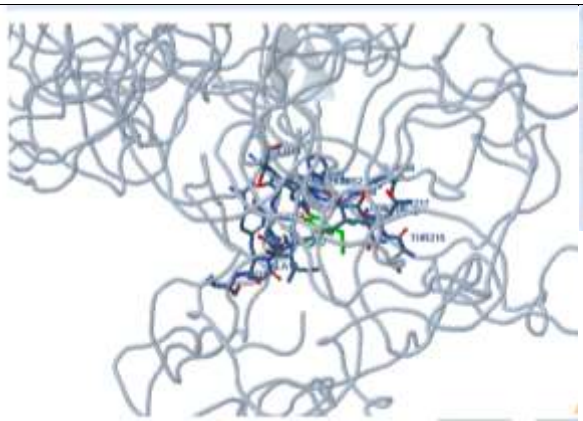
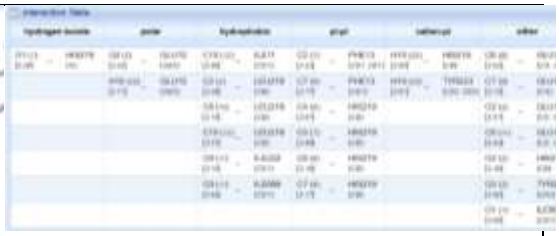
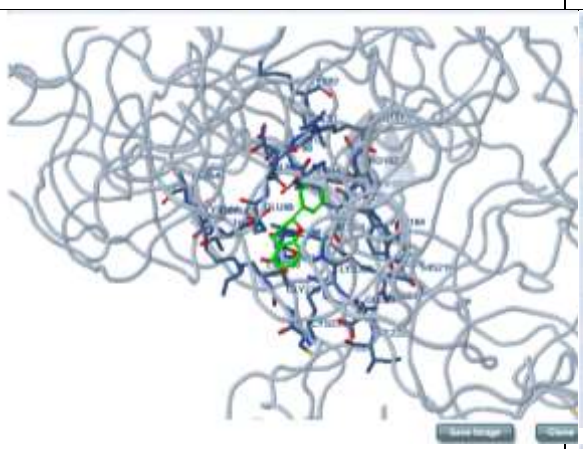
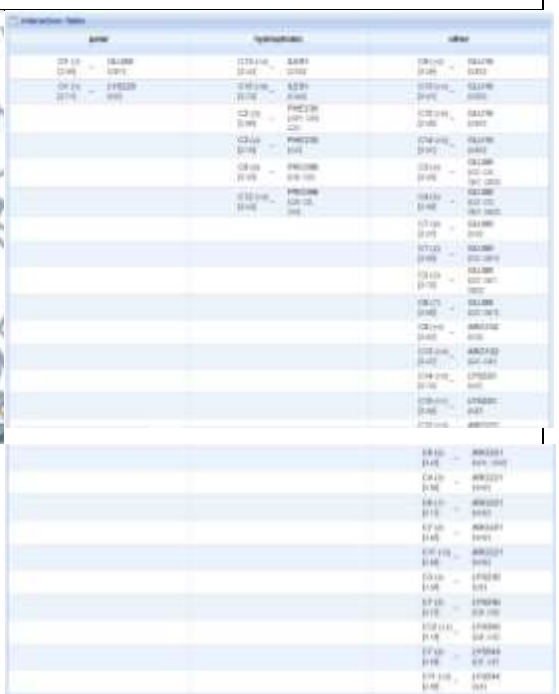
ELLAGIC ACID	-2.13kcaL/mol	62	YES
EUGENOL	-5.74kcal/mol	24	YES
FLAVAN	-7.16kcal/mol	33	YES
GALLIC ACID	-4.87kcal/mol	26	YES
ISOFLAVONE	-7.14kcal/mol	38	YES
METHYL GALLATE	-4.63kcal/mol	29	YES
QUERCETIN	-4.32 kcal/mol	70	YES
TERMILIGNAN	-6.23 kcal/mol	57	YES
THANNILIGNAN	-5.90 kcal/mol	69	YES

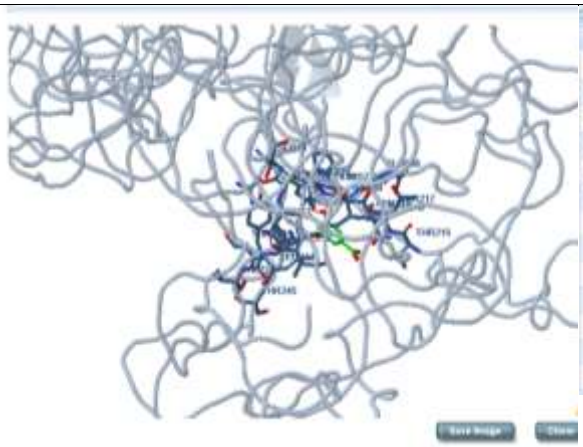
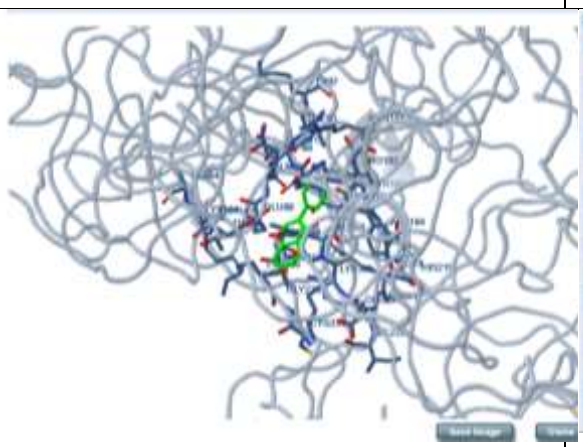


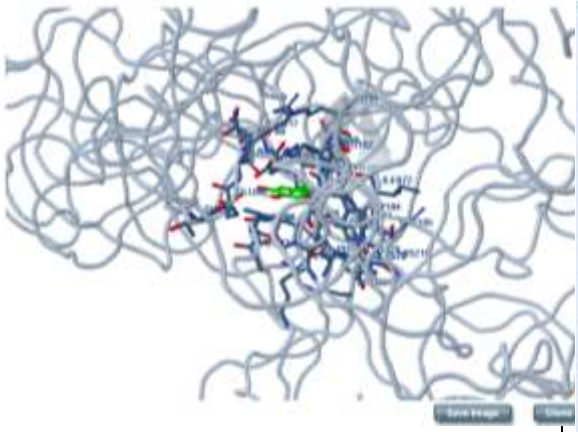
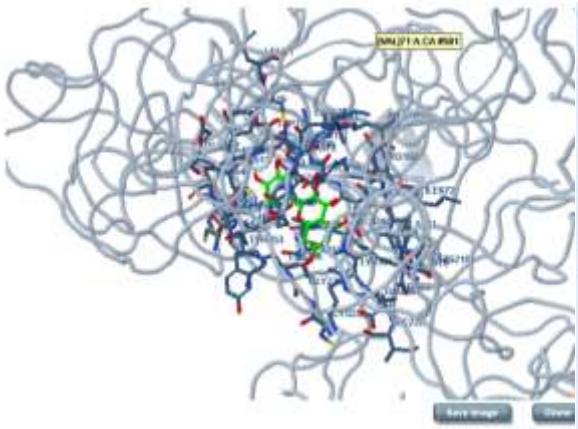
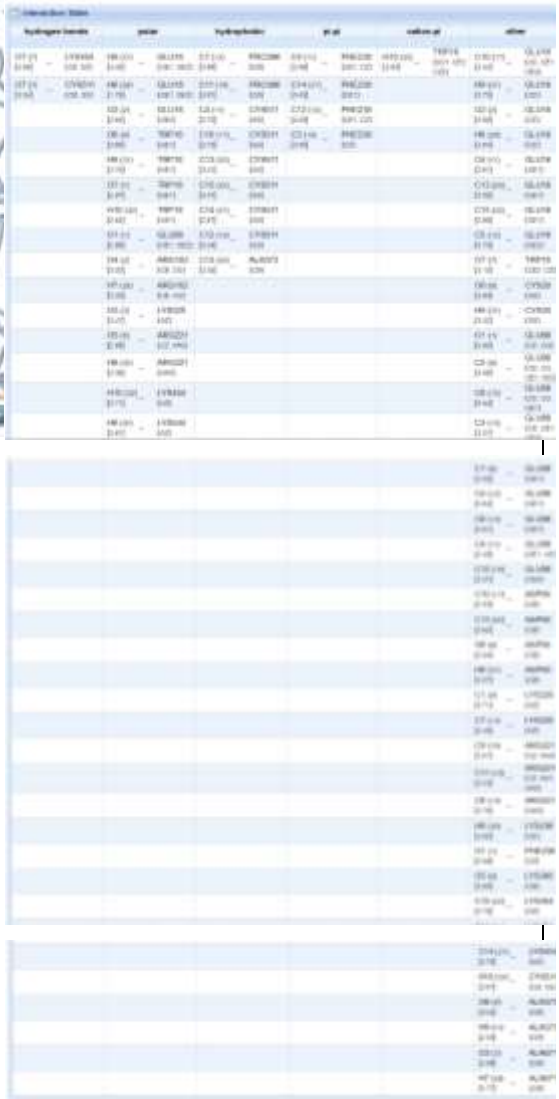
ASCORBIC ACID DOCKING

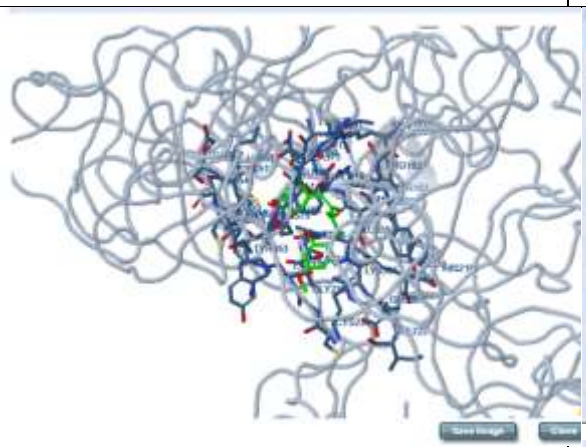
ASCORBIC ACID DOCKING INTERACTING AMINO ACIDS

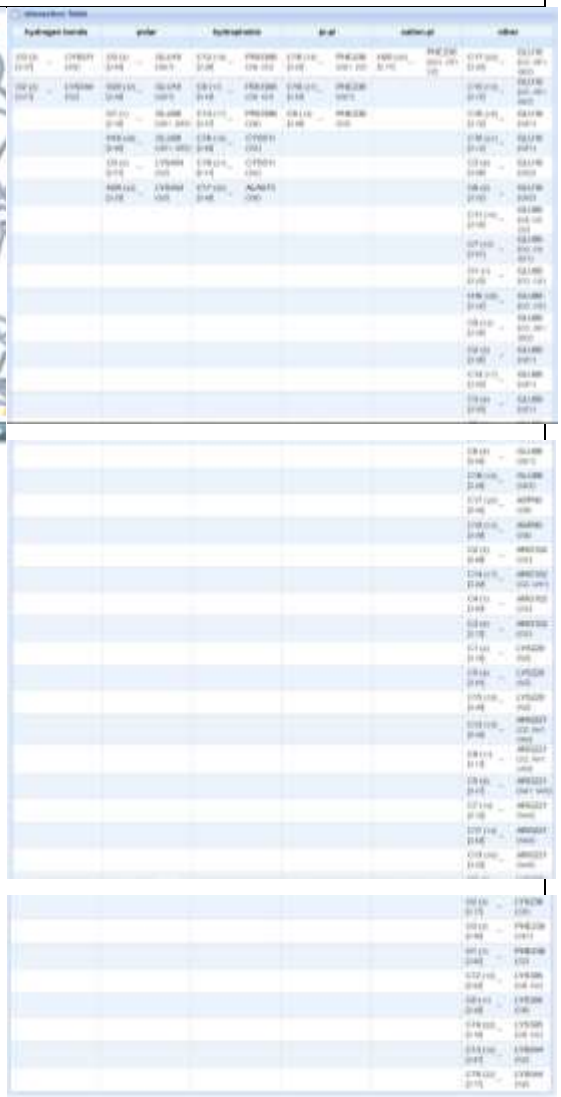
	
CATECHOL ACID DOCKING	CATECHOL ACID DOCKING INTERACTING AMINO ACIDS
	  
ELLAGIC ACID DOCKING	ELLAGIC ACID DOCKING INTERACTING AMINO ACIDS

	
EUGENOL ACID DOCKING	EUGENOL ACID DOCKING INTERACTING AMINO ACIDS
	
FLAVAN ACID DOCKING	FLAVAN ACID DOCKING INTERACTING AMINO ACIDS

	
GALLIC ACID DOCKING	GALLIC ACID DOCKING INTERACTING AMINO ACIDS
	 
ISOFLAVONE ACID DOCKING	ISOFLAVONE ACID DOCKING INTERACTING TWO AMINO ACIDS

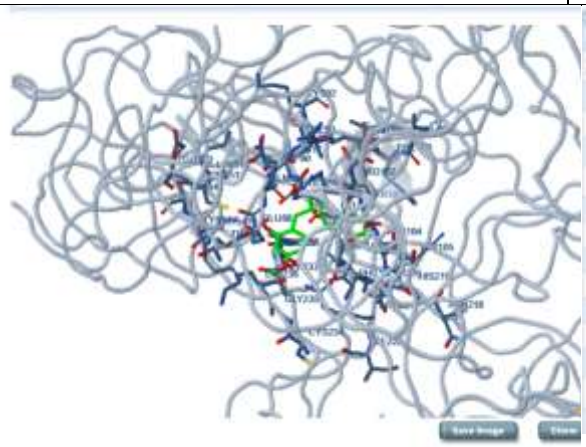
	
METHYL GALLATE ACID DOCKING	METHYL GALLATE ACID DOCKING INTERACTING AMINO ACIDS
	
QUERCETIN ACID DOCKING	QUERCETIN ACID DOCKING INTERACTING AMINO ACIDS

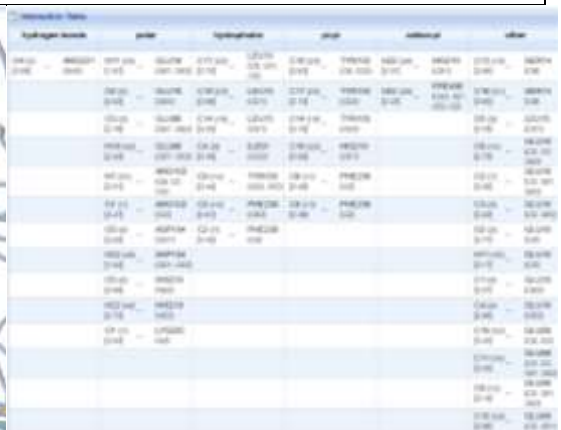




TERMILIGNAN ACID DOCKING

TERMILIGNAN ACID DOCKING INTERACTING AMINO ACIDS





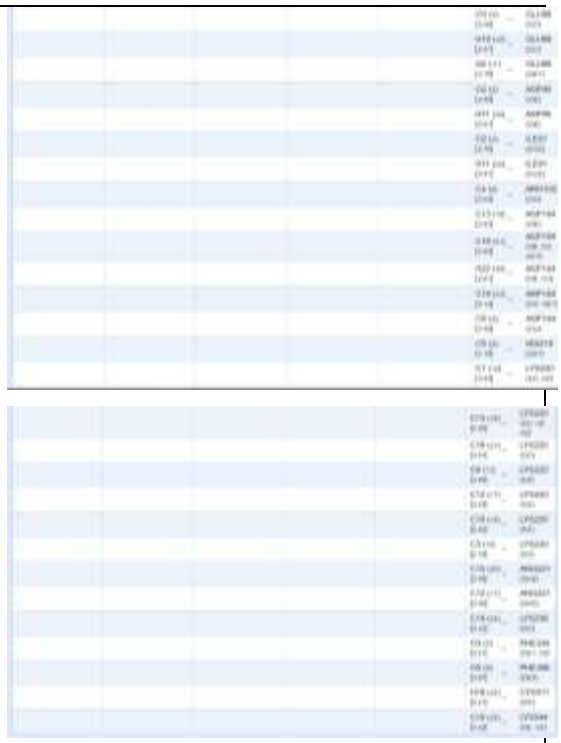
	
THANNILIGNAN ACID DOCKING	THANNILIGNAN ACID DOCKING INTERACTING AMINO ACIDS

Fig. 3(b): Docking results of RovA

5. Discussion

As per the results it is seen that T3SS docks with thannilignan with a docking score of +0.19kcal/mol with 90 interaction and ROVA docks with thannilignan with a docking score -5.90 kcal/mol with 69 interaction.

T3SS docks with Ellagic acid with a docking score of -0.06 kcal/mol with 81 interaction and ROVA docks with Ellagic acid with a docking score -2.13 kcal/mol with 62 interaction.

T3SS docks with quercetin with a docking score of -0.28 kcal/mol with 80 interaction and ROVA docks with quercetin with a docking score -4.32 kcal/mol with 70 interaction.

T3SS docks with termilignan with a docking score of -0.39 kcal/mol with 79 interaction and ROVA docks with termilignan with a docking score -6.23 kcal/mol with 57 interaction.

6. Conclusion

As per the results it is seen that Thannilignan, Ellagic acid, Quercetin Termilignan . T3SS docks with quercetin with docking score of -0.28kcal/mol with 80 interactions.ROVA docks with qurecitin with docking score -4.32kcal/mol with 70 interaction.

T3SS docks with termilignan with docking score -0.39kcal/mol with 79 interactions. ROVA docks with termilignan with docking score -6.23kcal/molwith 57 interactions.

Hence these phytocompounds can be used as ligands for the Bubonic Plague

References

1. Glatter KA, Finkelman P. History of the Plague: An Ancient Pandemic for the Age of COVID-19. Am J Med. 2021 Feb;134(2):176-181. doi: 10.1016/j.amjmed.2020.08.019. Epub 2020 Sep 24. PMID: 32979306; PMCID: PMC7513766.
2. Riedel S. Plague: from natural disease to bioterrorism. Proc (Bayl Univ Med Cent). 2005 Apr;18(2):116-24. doi: 10.1080/08998280.2005.11928049. PMID: 16200159; PMCID: PMC1200711.
3. Barbieri R, Signoli M, Chevé D, Costedoat C, Tzortzis S, Aboudharam G, Raoult D, Drancourt M. Yersinia pestis: the Natural History of Plague. Clin Microbiol Rev. 2020 Dec 9;34(1):e00044-19. doi: 10.1128/CMR.00044-19. PMID: 33298527; PMCID: PMC7920731.
4. Coburn B, Sekirov I, Finlay BB. Type III secretion systems and disease. Clin Microbiol Rev. 2007 Oct;20(4):535-49. doi: 10.1128/CMR.00013-07. PMID: 17934073; PMCID: PMC2176049.
5. Perry RD, Fetherston JD. Yersinia pestis--etiologic agent of plague. Clin Microbiol Rev. 1997 Jan;10(1):35-66. doi: 10.1128/CMR.10.1.35. PMID: 8993858; PMCID: PMC172914.
6. Cathelyn JS, Crosby SD, Lathem WW, Goldman WE, Miller VL. RovA, a global regulator of Yersinia pestis, specifically required for bubonic plague. Proc Natl Acad Sci U S A. 2006 Sep 5;103(36):13514-9. doi: 10.1073/pnas.0603456103. Epub 2006 Aug 28. PMID: 16938880; PMCID: PMC1569194.
7. Heroven AK, Nagel G, Tran WH, Parr S, Dersch P, 2004, RovA is autoregulated and antagonizes H-NS-mediated silencing of invasin and rovA expression in Yersinia pseudotuberculosis, Molecular microbiology, 53, DO - 10.1111/j.1365-2958.2004.04162.x
8. Webb B and Sali A. Comparative Protein Structure Modeling Using Modeller. Current Protocols in Bioinformatics 54, John Wiley & Sons, Inc., 5.6.1-5.6.37, 2016.
9. Hollingsworth SA, Karplus PA. A fresh look at the Ramachandran plot and the occurrence of standard structures in proteins. Biomol Concepts. 2010 Oct;1(3-4):271-283. doi: 10.1515/BMC.2010.022. PMID: 21436958; PMCID: PMC3061398
10. Bikadi, Z., Hazai, E. Application of the PM6 semi-empirical method to modeling proteins enhances docking accuracy of AutoDock J. Cheminf. 1, 15 (2009)