Establishing Ligands for Tuberculosis Gene Receptor Using *In-Silico* Methods

Shraddha Lavhale*, Rehan Khan and Preenon Bagchi

Institute of Biosciences and Technology, MGM University, Chht. Sambhajinagar,

India

Lavhaleshraddha4@gmail.com

Abstract. ***Mycobacterium tuberculosis*** is a species of pathogenic bacteria of the family Mycobacteriaceae and the causative agent of tuberculosis. First discovered in 1882 by Robert Koch, *M. tuberculosis* has an unusual, waxy coating on its cell surface primarily due to the presence of mycolic acid. There are two forms tuberculosis (TB): Latent TB and Active TB. TB bacteria can live in the body without making you sick. This is called latent TB infection. People with latent TB infection: Have no symptoms. Active TB disease is a contaginous. That means it can be spreads from one person to another. It is most often spread through the air. The germs may enter the air when person with TB disease of the lungs or throat cough or seenzes people nearby breathe Active in these germs and get infected. The FASTA sequence of the (genes) was retrieved from Genbank. 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 *Ocimum sanctum*, *Tinospora cordifolia*, *Curcuma longa*, *Eclipta prostrate*, *Elettaria cardamomum* 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 violation was considered for further docking studies. The phytocompounds kaempferol, fisetin , piperin having least docking score and most interactions is considered as the drug leads for Tuberculosis. Further receptor ligand binding assay studies will be done to establish the compound as drug for the above disease.

Keywords: Melioidosis, Burkholderia pseudomallei, Whitmore's disease, Bioinformatics, Docking, Phytocompound

1 Introduction

Tuberculosis (TB) is a leading cause of death worldwide. It is an infectious disease usually caused *by Mycobacterium tuberculosis* bacteria [1]. Tuberculosis is an airborne disease, spread from one person to another person through the air. It is also known as Koach Basillus, since Robert Koch 1st observed it in 1882 [2]. The bacterium belongs to the family mycobacteriace. Mycobacterium is a small, aerobic, non- motile bacillus. In nature, the bacterium can grow only within the cells of a host, but it can be cultured in the laboratory. [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_18

1.1 How the genes cause the disease

NRAMP1 (Natural Resistance Asssociated Macrophage Protein 1) is a protein encode by the SLC11A1 gene. When a mutation of NRAMP-1 gene yields a nonfunctional NRAMP-1 protein, there is an inhibition on the intracellular killing mechanism of M. tuberculosis in macrophage [4].
CCL2 (C-C motif chemokine ligand-2) was evidenced to be associated with tuberculosis susceptibility in some ethnic groups [5].

A. Objectives

Computer aided drug design
In this work, the receptor gene of receptor NRAMP1 and CCL2 is taken and their 3d structure is modeled. Further, the phytocompounds from *Ocimum sanctum*, *Tinospora cordifolia*, *Curcuma longa*, *Eclipta prostrate*, *Elettaria cardamomum* are selected, screened with molinspiration based on the principles of ADME and docked with the gene receptors.

B. Abbreviations

1. et. al. : Etalia (and associate)
2. TB: Tuberculosis
3. MTB: Mycobacterium Tuberculosis
4. NCBI: National center of Biotechnology Information
5. BLAST: Basic Local Alignment Search Tool
6. FASTA: FAST-ALL
7. NRAMP1: Natural Resistance-Associated Macrophage Protein1
8. CCL2: C-C Motif Chemokine Ligand 2
9. PTB: Pulmonary Tuberculosis
10. HTS :High Throughput Sequence

2 METHODOLOGY

The FASTA sequence of the receptors NRMP-1 and CCL2) was retrieved from genbank database. Their 3d structure was modeled using modeler [6] and verified using Ramachandran Plot [7].
The SMILES of the photocompounds, *Ocimum sanctum*, *Tinospora cordifolia*, *Curcuma longa*, *Eclipta prostrate*, *Elettaria cardamomum* are retrieved from PubChem and ADME screening is done. Further, the selected compounds are docked with modeler generated best .

RESULTS

The amino acid sequence of the receptors CCL2 and NRAMP1 are retrieved and using BLAST, their homologous templates are downloaded from RCSB PDB (Table 1).

Gene receptor	Genbank accession number	Homologous templates
NRAMP1	OU343079.1	3JB9_B 7D58_D 6ZXF_F
CCL2	CP034501.1	5AOX_F 4UE5_E 7OBQ_V

The amino acid sequence of the receptors CCL2 and NRAMP1 are retrieved and using BLAST, their homologous templates are downloaded from RCSB PDB (Table 1).

Table 1: Genbank accession number of the receptors with their homologous templates

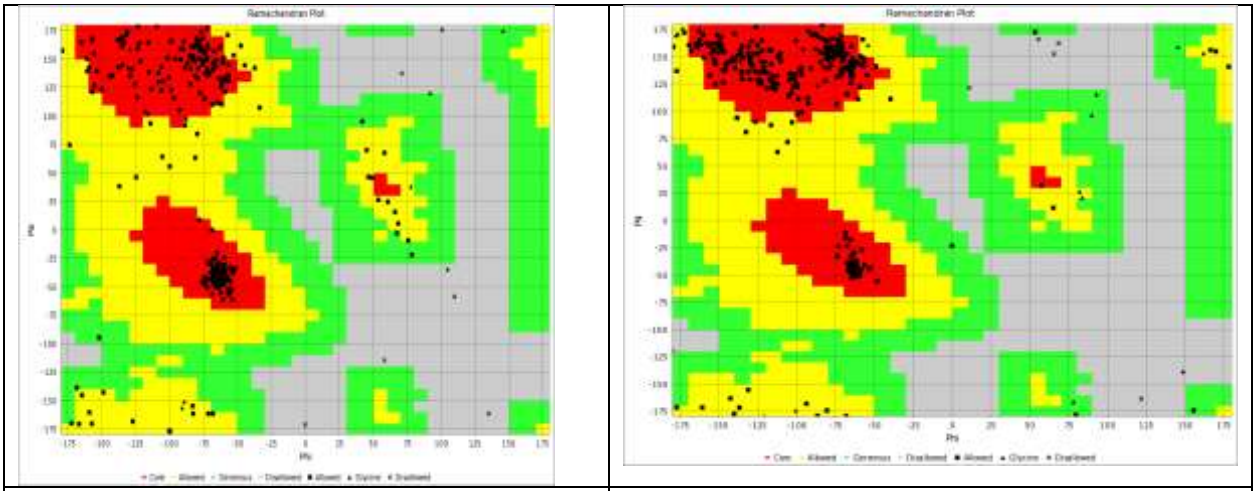
The modeler generated modes are verified using Ramachandran plot server (Table 2 and Fig. 1).

Table 2(a): Ramachandran plot analysis of CCL2

	# res in phipsi core	# res in phipsi allowed	# res in phipsi generous	# res in phipsi outside	
Model1	212 (75%)	55 (19%)	9 (3%)	3 (1%)	
Model2	228 (81%)	40 (14%)	7 (2%)	4 (1%)	
Model3	217 (77%)	48 (17%)	9 (3%)	5 (1%)	
Model4	229 (82%)	42 (15%)	5 (1%)	3 (1%)	selected
Model5	218 (78%)	53 (18%)	3 (1%)	5 (1%)	

Table 2(b): Ramachandran plot analysis of NRAMP1

	# res in phipsi core	# res in phipsi allowed	# res in phipsi generous	# res in phipsi outside	
Model1	268 (83%)	41 (12%)	10 (3%)	3 (0%)	
Model2	208 (64%)	56 (17%)	28 (8%)	30 (9%)	
Model3	248 (77%)	46 (14%)	16 (4%)	12 (3%)	
Model4	228 (70%)	61 (18%)	20 (6%)	13 (4%)	
Model5	287 (89%)	27 (8%)	3 (0%)	5 (1%)	selected



CCL2 **NRAMP1**

Fig. 1: Ramachandran Plot of the best model

The medicinal plants, *Ocimum sanctum*, *Tinospora cordifolia*, *Curcuma longa*, *Eclipta prostrate*, *Elettaria cardamomum* are selected and their SMILES are retrieved from PubChem (Table 3).

Table 3: **Phytocomponds and their SMILES from Pubchem**

Plant name: ocimum sanctum (Saha and Ghosh, 2012)	
Phytocompound	SMILES
Eugenol	COC1=C(C=CC(=C1)CC=C)O
Methyleugenol	COC1=C(C=C(C=C1)CC=C)OC
Methylchavicol	COC1=CC=C(C=C1)CC=C
Beta-Elemene	CC(=C)C1CCC(C(C1)C(=C)C)(C)C=C
Beta-caryophyllene	CC1=CCCC(=C)C2CC(C2CC1)(C)C
Caryophyllene oxide	CC1(CC2C1CCC3(C(O3)CCC2=C)C)C
Isocaryophyllene	CC1=CCCC(=C)C2CC(C2CC1)(C)C

Plant name: Tinospora cordifolia (Pattanayak et.al., 2010)	
Phytocompound	SMILES
Eugenol	COC1=C(C=CC(=C1)CC=C)O
Carvacrol	CC1=C(C=C(C=C1)C(C)C)O
Linalool	CC(=CCCC(C)(C=C)O)C

Plant name: curcuma longa (Ammon andWahl, 1991)	
Phytocompounds	SMILES
Myricetin	C1=C(C=C(C(=C1O)O)O)C2=C(C(=O)C3=C(C=C(C=C3O2)O)O)O
Qurecetin	CC1C(C(C(C(O1)OC2=C(OC3=CC(=CC(=C3C2=O)O)O)C4=CC(=C(C=C4)O)O)O)O)O
Vasicine	C1CN2CC3=CC=CC=C3N=C2C1O
Piperine	C1CCN(CC1)C(=O)C=CC=CC2=CC3=C(C=C2)OCO3
Ascorbic acid	C(C(C1C(=C(C(=O)O1)O)O)O)O
Plant name: Eclipta prostrate (Feng et.al, 2019)	
Phytocompounds	SMILES
Qurecetin	CC1C(C(C(C(O1)OC2=C(OC3=CC(=CC(=C3C2=O)O)O)C4=CC(=C(C=C4)O)O)O)O)O
(-)-epicatechin	C1C(C(OC2=CC(=CC(=C21)O)O)C3=CC(=C(C=C3)O)O)O
(-)epigallocatechin	C1C(C(OC2=CC(=CC(=C21)O)O)C3=CC(=C(C(=C3)O)O)O)OC(=O)C4=CC(=C(C(=C4)O)O)O
Kaempferol	C1=CC(=CC=C1C2=C(C(=O)C3=C(C=C(C=C3O2)O)O)O)O
Luteolin	C1=CC(=C(C=C1C2=CC(=O)C3=C(C=C(C=C3O2)O)O)O)O
Fisetin	C1=CC(=C(C=C1C2=C(C(=O)C3=C(O2)C=C(C=C3)O)O)O)O

Plant name: Elettaria cardamomum (Yahyazadeh et.al, 2011)	
Phytocompounds	SMILES
Alpha-terpinyl acetate	CC1=CCC(CC1)C(C)(C)OC(=O)C
1,8-cineole	CC1(C2CCC(O1)(CC2)C)C
Alpha-terpineol	CC1=CCC(CC1)C(C)(C)O
Linalyl acetate	CC(=CCCC(C)(C=C)OC(=O)C)C
Terpinen-4-ol	CC1=CCC(CC1)(C(C)C)O
Beta-pinen	CC1(C2CCC(=C)C1C2)C
Beta-selinene	CC(=C)C1CCC2(CCCC(=C)C2C1)C

Further, using molinspiration their drug- like properties are identified and screened based on no violations from Lipinski's rule of five (Table 4).

Table 4: ADME property detection of phytocomponds using molinspiration (molinspiration)

Phytocompounds	mil.ogP	TPSA	natoms	MW	nON	nOHNH	nrotb	volume	nviolation
Eugenol	2.10	29.46	12	164.20	2	1	3	162.14	0
methyleugenol	2.41	18.47	13	178.23	2	0	4	179.67	0
methylchavicol	2.82	9.23	11	148.21	1	0	3	154.12	0
Beta-Elemene	5.37	0.00	15	204.36	0	0	3	235.23	1
Beta-caryophyllene	5.17	0.00	15	204.36	0	0	0	229.95	1
Caryophyllene oxide	4.14	12.53	16	220.36	1	0	0	234.01	0
Isocaryophyllene	5.17	0.00	15	204.36	0	0	0	229.95	1
Eugenol	2.10	29.46	12	164.20	2	1	3	162.14	0
carvacrol	3.81	20.23	11	150.22	1	1	1	158.57	0
Linalool	3.21	20.23	11	154.25	1	1	4	175.59	0
myricetin	1.39	151.58	23	318.24	8	6	1	248.10	1
Qurecetin	0.64	190.28	32	448.38	11	7	3	363.95	2
Vasicine	1.04	35.83	14	188.23	3	1	0	173.66	0
Piperine	3.33	38.78	21	285.34	4	0	3	267.74	0
Ascorbic acid	-1.40	107.22	12	176.12	6	4	2	139.71	0
Qurecetin	0.64	190.28	32	448.38	11	7	3	363.95	2
(-)-epicatechin	1.37	110.37	21	290.27	6	5	1	244.14	0
(-)epigallocatechin	2.25	197.36	33	458.38	11	8	4	367.57	4
kaempferol	2.17	111.12	21	286.24	6	4	1	232.07	0
Luteolin	1.97	111.12	21	286.24	6	4	1	232.07	0
Fisetin	1.97	111.12	21	286.24	6	4	1	232.07	0
Alpha-terpinyl acetate	3.30	26.30	14	196.29	2	0	3	207.16	0
1,8-cineole	2.72	9.23	11	154.25	1	0	0	166.66	0
Alpha-terpineol	2.60	20.23	11	154.25	1	1	1	170.65	0
Linalyl acetate	3.92	26.30	14	196.29	2	0	6	212.10	0
Terpinen-4-ol	2.60	20.23	11	154.25	1	1	1	170.65	0
Beta-pinen	3.33	0.00	10	136.24	0	0	0	152.37	0
Beta-selinene	5.02	0.00	15	204.36	0	0	1	230.51	1

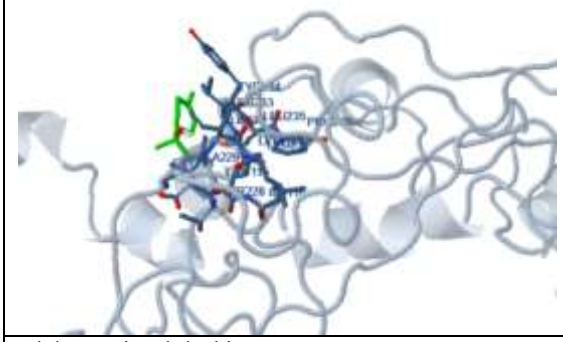
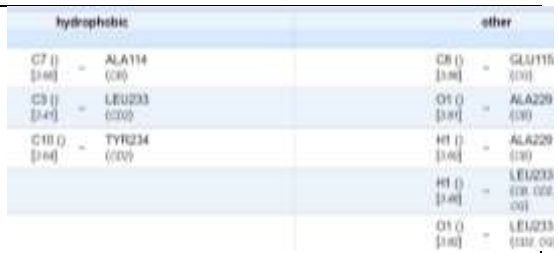
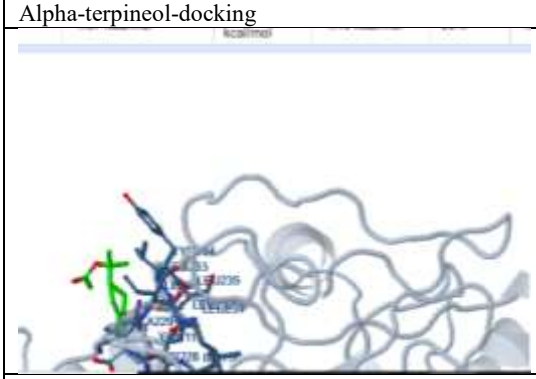
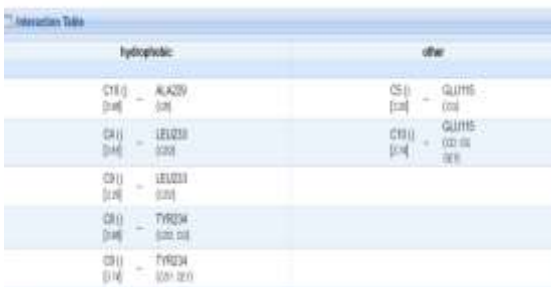
The phycompounds having nviolations 0 are considered for further docking studies (Bikadi and Hazai, 2009) (Table 5). The phytocompounds are docked with best receptor models selected above in Table 2 and 3.

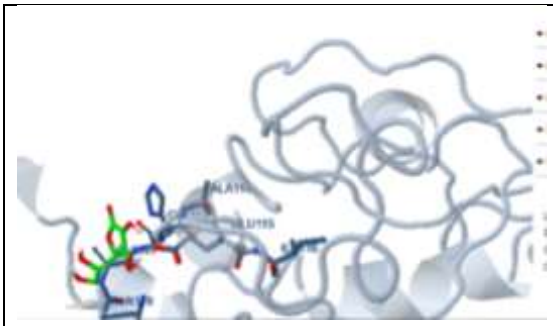
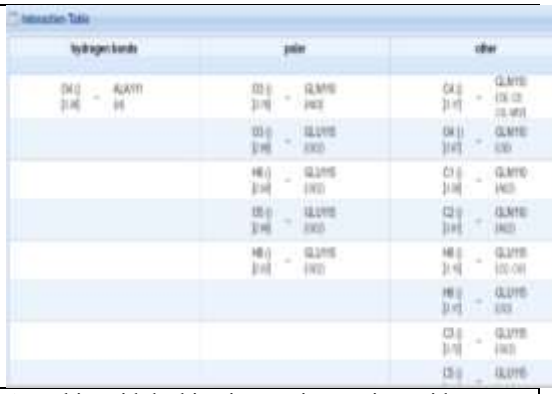
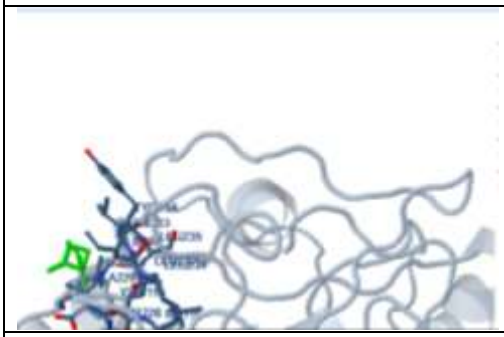
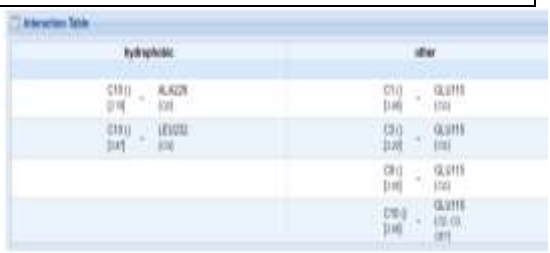
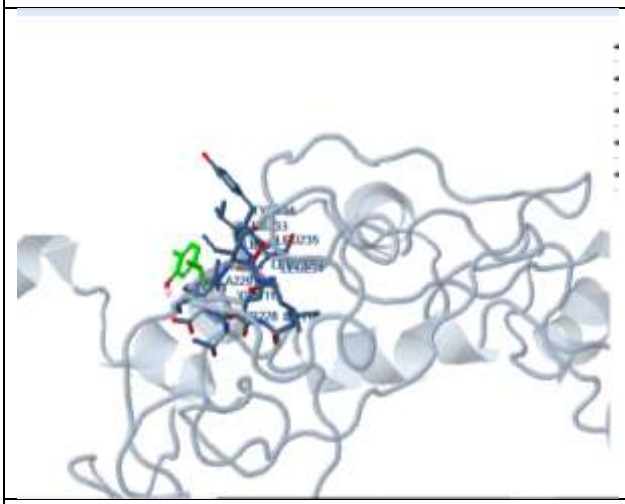
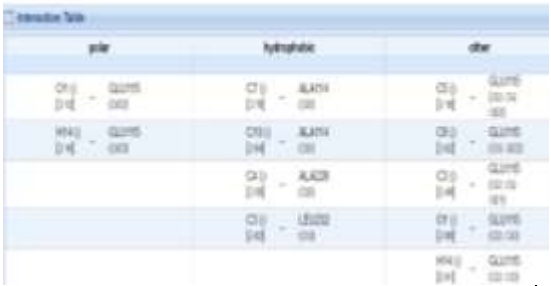
Table 5(a): Docking studies of CCL2.D99990004

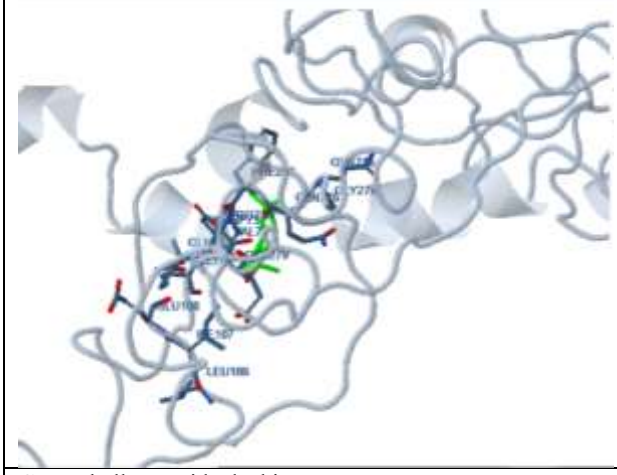
Phyto-compound	Docking score	No. of interactions	Docking (yes/no)
Alpha-terpineol	-4.95kcal/mol	8	Yes
Alpha-terpinyl acetate	-4.10kcal/mol	7	Yes
Ascorbic acid	-2.87kcal/mol	14	Yes
Beta-pinen	-4.07kcal/mol	6	Yes
Carvacrol	-4.50kcal/mol	11	Yes
Caryophyllene oxide	-4.82kcal/mol	13	Yes

Eugenol	-3.91kcal/mol	11	Yes
Fisetin	-4.84kcal/mol	16	Yes
Kaempferol	-4.65kcal/mol	19	Yes
Linalool	-3.57kcal/mol	14	Yes
Linalyl acetate	-4.03kcal/mol	14	Yes
Luteolin	-4.67kcal/mol	20	Yes
Methylchavicol	-3.93kcal/mol	10	Yes
Methyleugenol	-4.59kcal/mol	10	Yes
Piperin	-5.04kcal/mol	15	Yes
Terpinen-4-ol	-4.11kcal/mol	7	Yes
Vasicine	-3.57kcal/mol	11	Yes

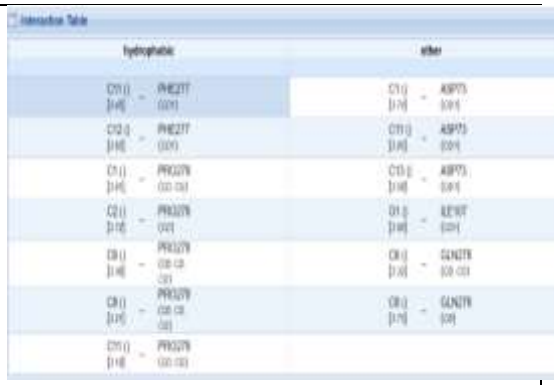
Table 5(b): Docking images and their interacted amino acid residues.

	
Alpha-terpineol-docking	Alpha-terpineol docking interacting amino acids
	
Alpha-terpinyl acetate docking	Alpha-terpinyl acetate docking interacting amino acids

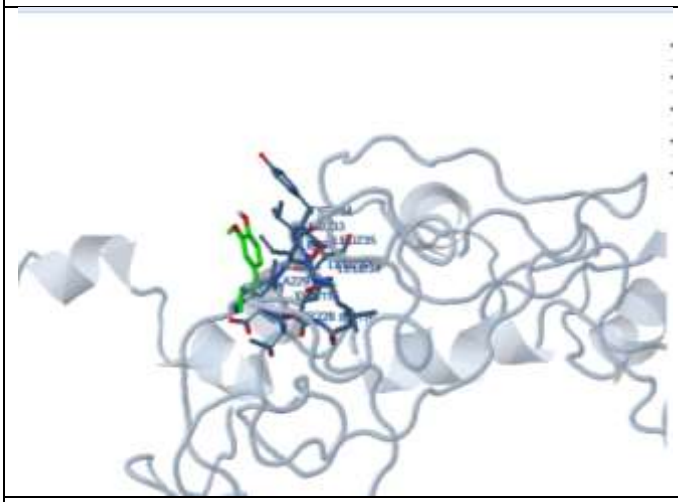
	
Ascorbic acid docking	Ascorbic acid docking interacting amino acids
	
Beta-pinen docking	Beta-pinen docking interacting amino acids
	
Carvacrol docking	Carvacrol docking interacting amino acids



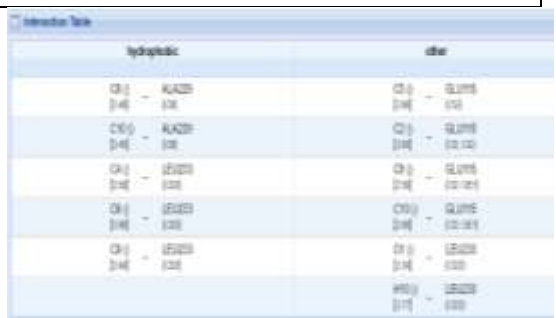
Caryophyllene oxide docking



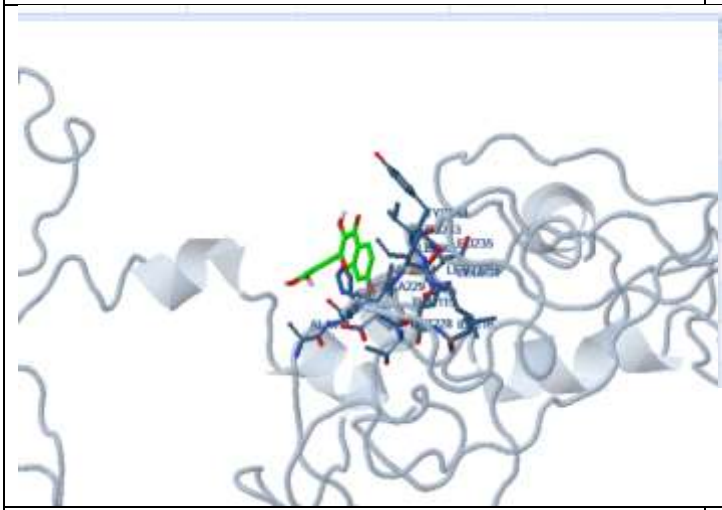
Caryophyllene oxide docking interacting amino acids



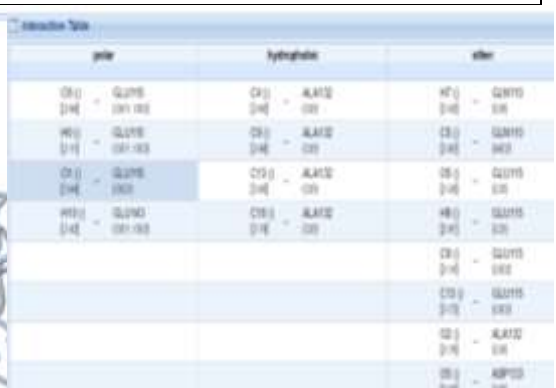
Eugenol docking



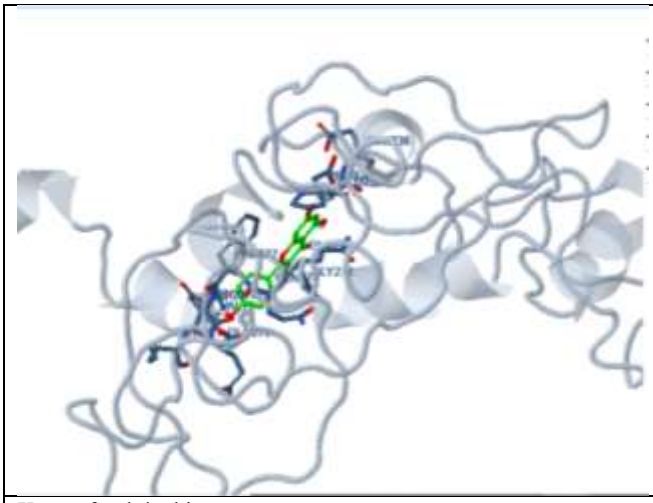
Eugenol docking interacting amino acids



Fisetin docking

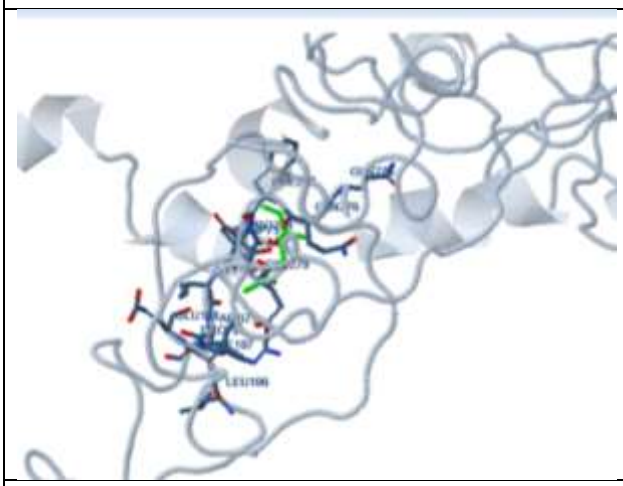


Fisetin docking interacting amino acids

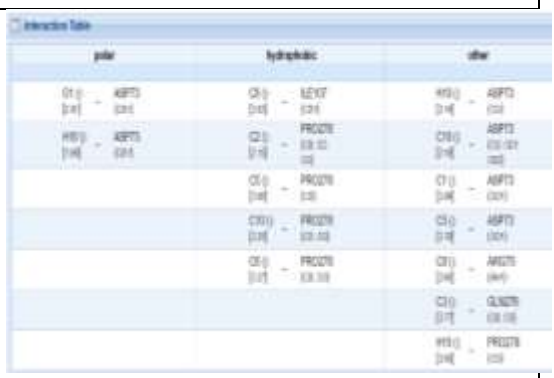


Kaempferol docking

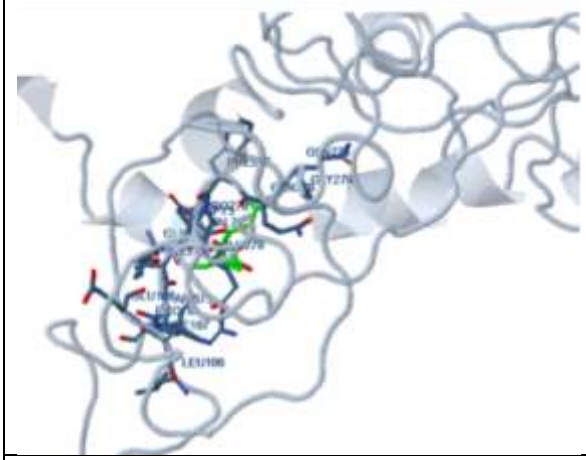
Kaempferol docking interacting amino acids



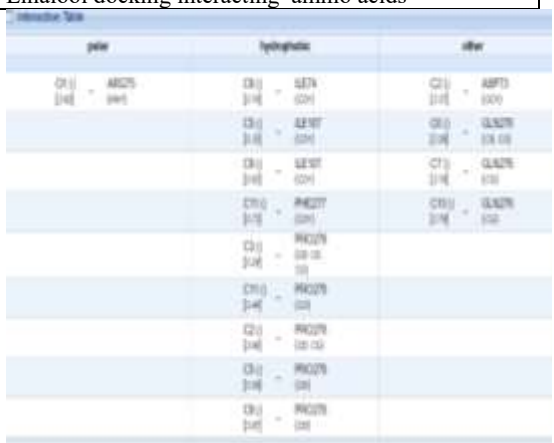
Linalool docking



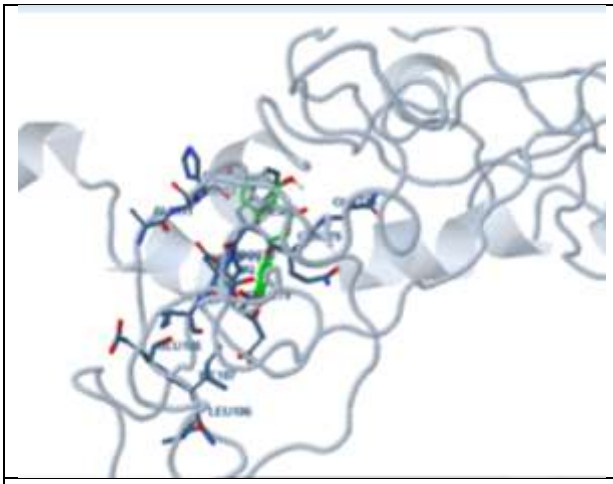
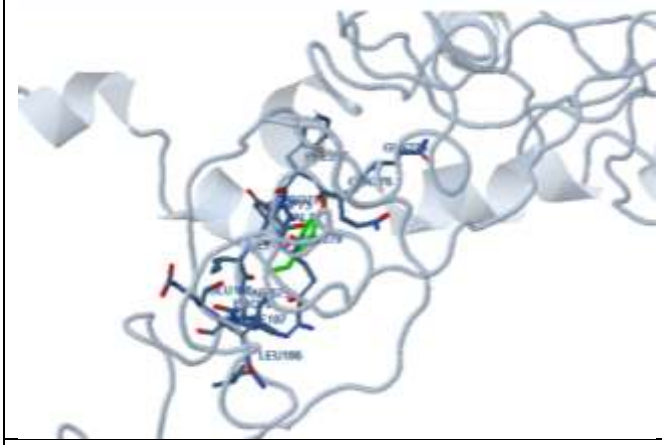
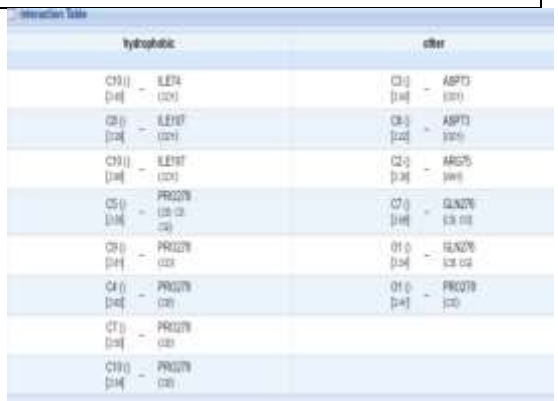
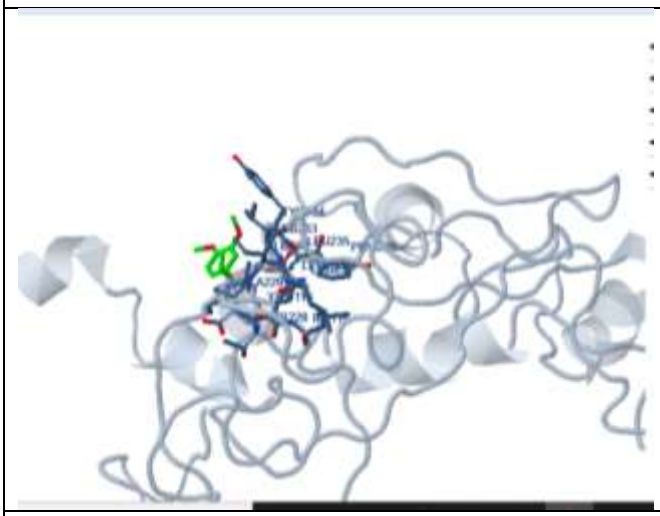
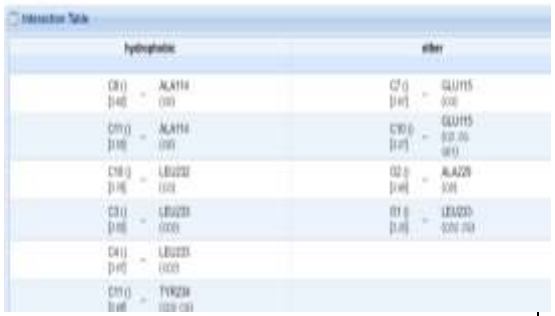
Linalool docking interacting amino acids



Linalyl acetate docking



Linalyl acetate docking interacting amino acids

	
Luteolin docking	Luteolin docking interacting amino acids
	
Methylchavicol docking	Methylchavicol docking interacting amino acids
	
Methyleugenol docking	Methyleugenol docking interacting amino acids

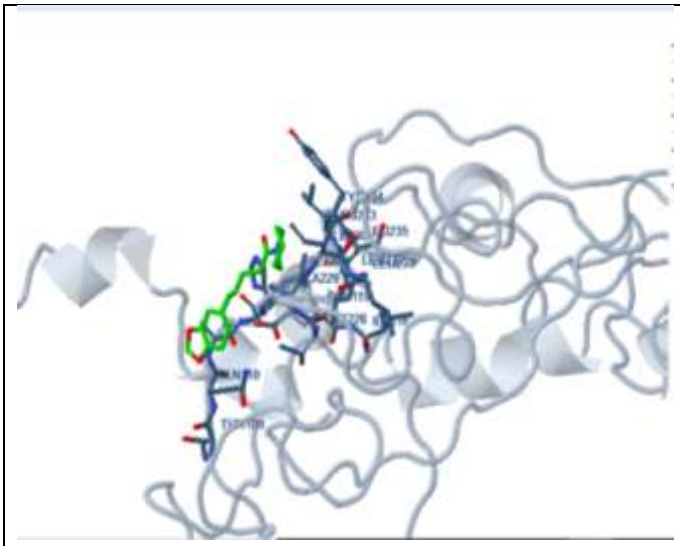
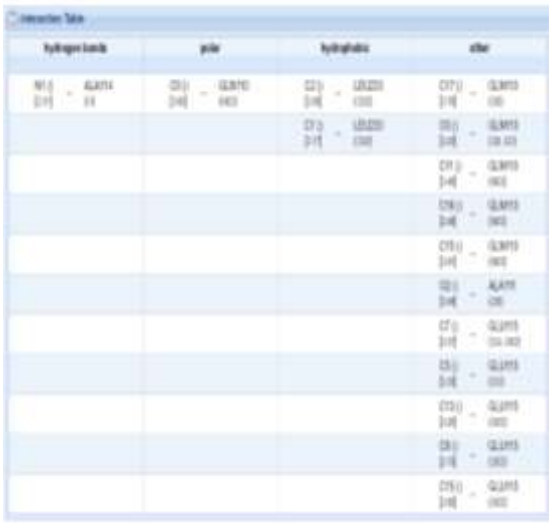
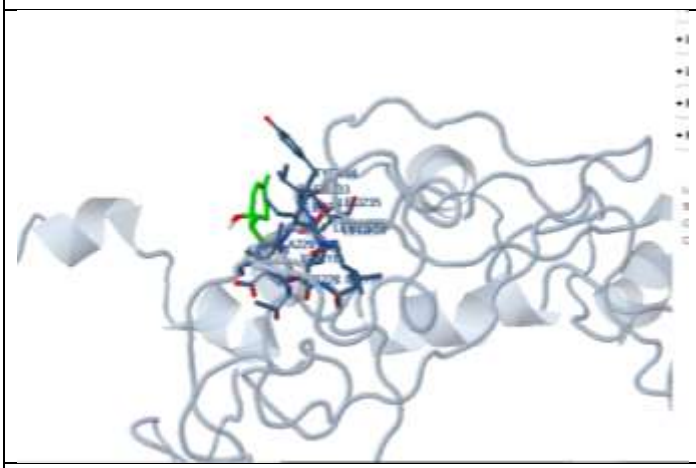
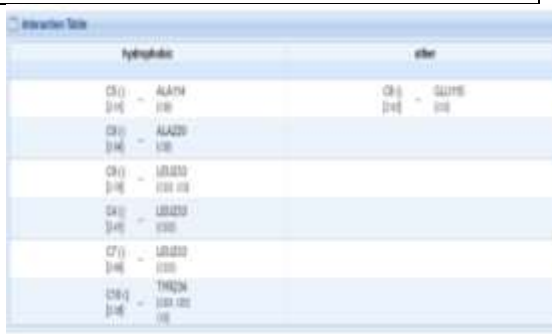
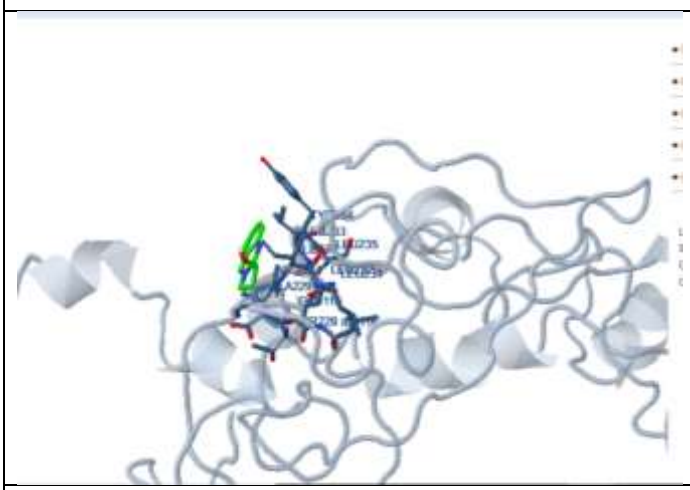
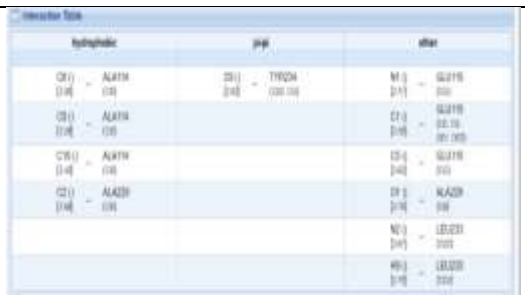
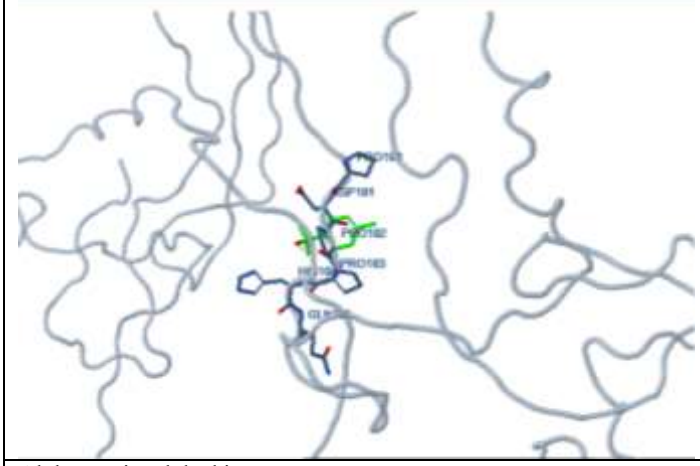
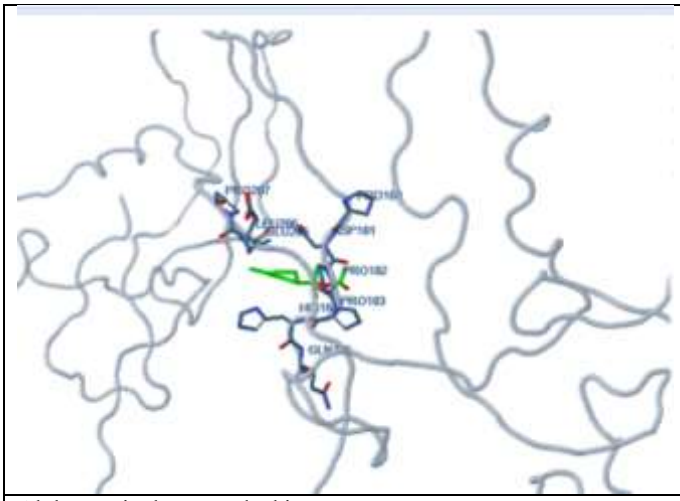
	
Piperin docking	Piperin docking interacting amino acids
	
Terpinen-4-ol docking	Terpinen-4-ol docking interacting of amino acids
	
Vasicine docking	Vasicine docking interacting of amino acids

Table 6(a): Docking studies of NRAMP1

Phyto-compound	Docking score	No. of interactions	Docking (yes/no)
Alpha-terpineol	-3.64kcal/mol	9	Yes
Alpha-terpinyl acetate	-3.25kcal/mol	11	Yes
Ascorbic acid	-1.51kcal/mol	9	Yes
Bita-pinen	-3.01kcal/mol	15	Yes
Carvacrol	-3.12kcal/mol	16	Yes
Caryophyllene oxide	-3.83kcal/mol	10	Yes
Eugenol	-2.96kcal/mol	9	Yes
Fisetin	-3.73kcal/mol	21	Yes
Kaempferol	-3.64kcal/mol	16	Yes
Linalool	-2.98kcal/mol	13	Yes
Luteolin	-2.77kcal/mol	12	Yes
Linalyl acetate	-3.67kcal/mol	21	Yes
Methylchavicol	-3.21kcal/mol	10	Yes
Methyleugenol	-3.52kcal/mol	9	Yes
Piperine	-4.16kcal/mol	17	Yes
Terpinen-4-ol	-3.09kcal/mol	13	Yes
Vasicine	-3.07kcal/mol	10	Yes

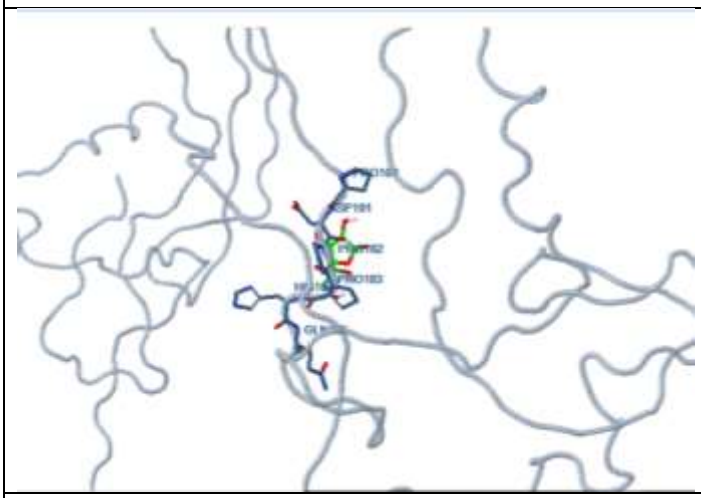
Table 6(b): Docking images and their interacted amino acid residues.

	
Alpha-terpineol docking	Alpha-terpineol docking interacting of amino acids



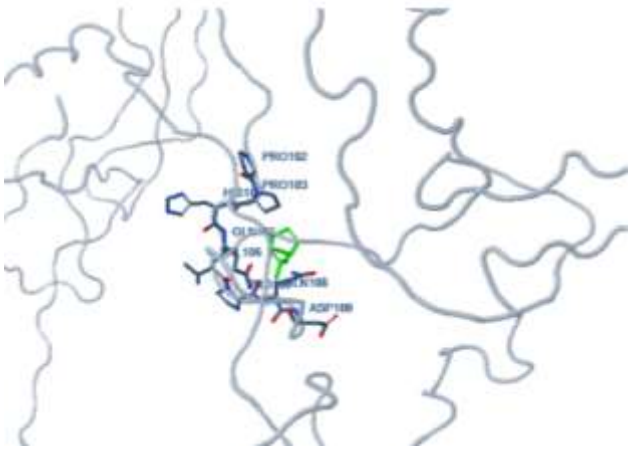
Alpha-terpinyl acetate docking

Alpha-terpinyl acetate docking interacting of amino acids



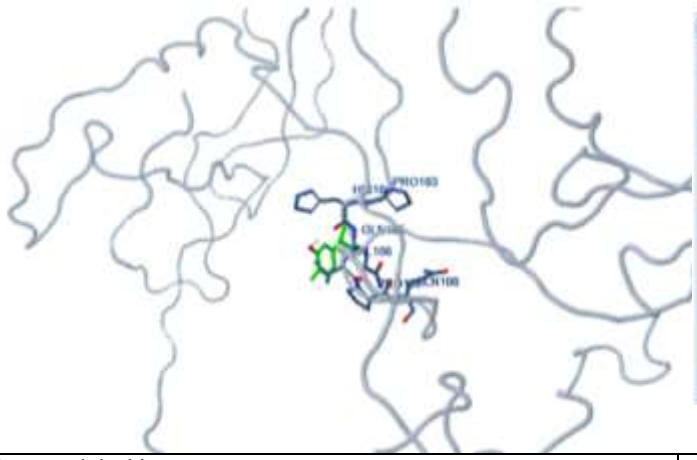
Ascorbic acid docking

Ascorbic acid docking interact with amino acids

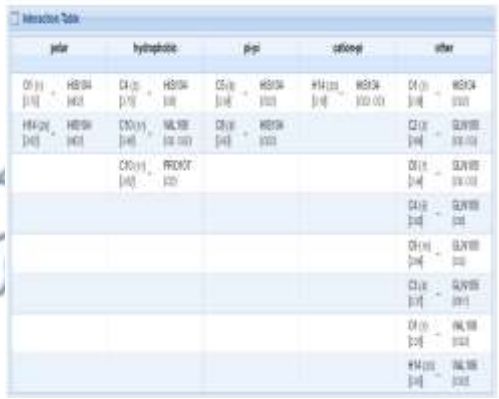


Bita-pinen docking

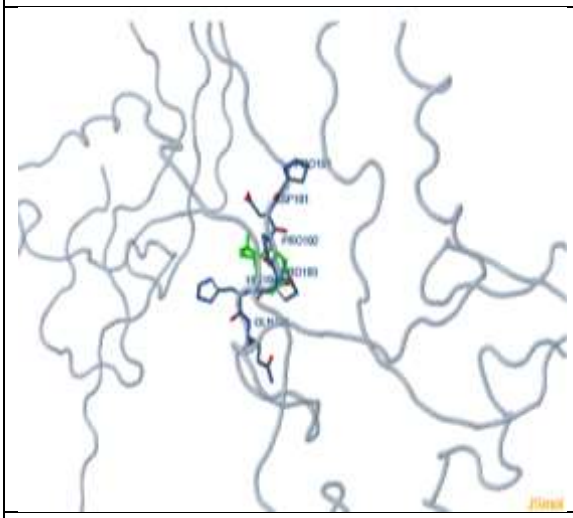
Bita-pinen docking interacting of amino acids



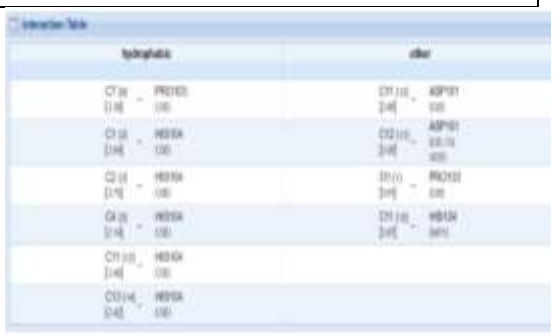
Carvacrol docking



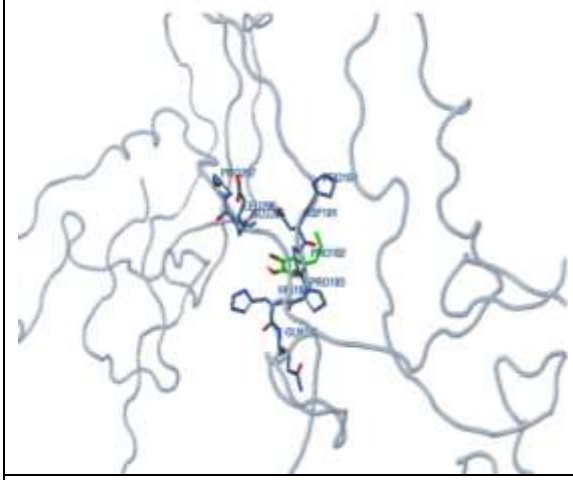
Carvacrol docking interacting of amino acids



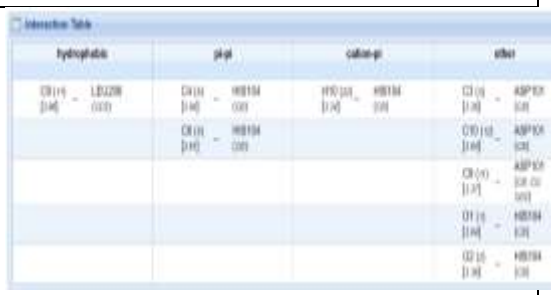
Caryophyllene oxide docking



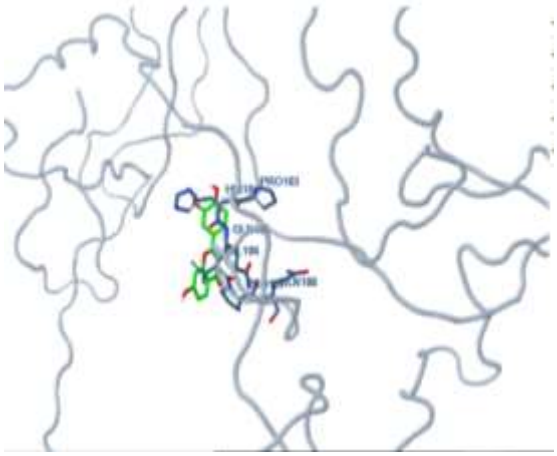
Caryophyllene oxide docking interacting of amino acids



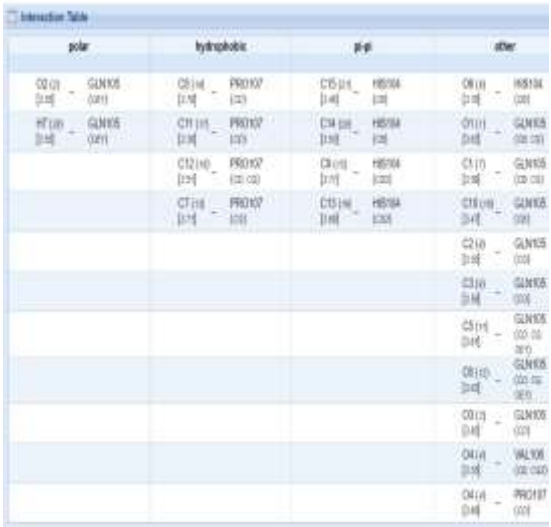
Eugenol docking



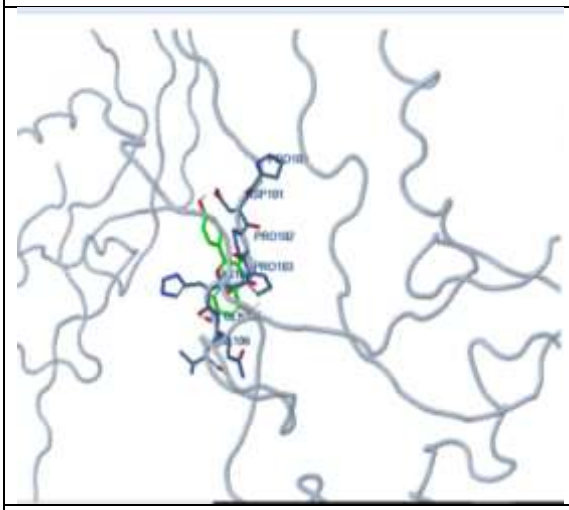
Eugenol docking interacting of amino acids



Fisetin docking

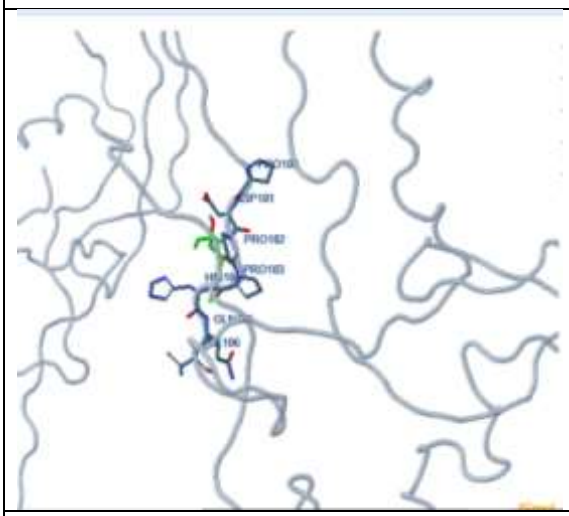


Fisetin docking interacting of amino acids

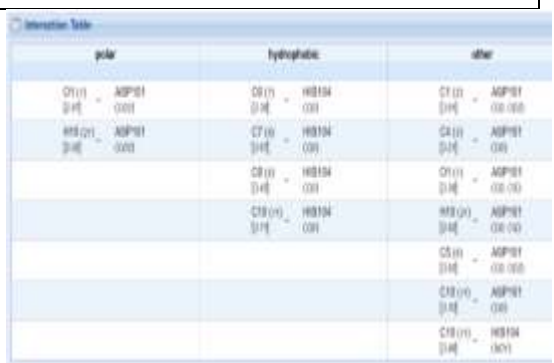


Kaempferol docking

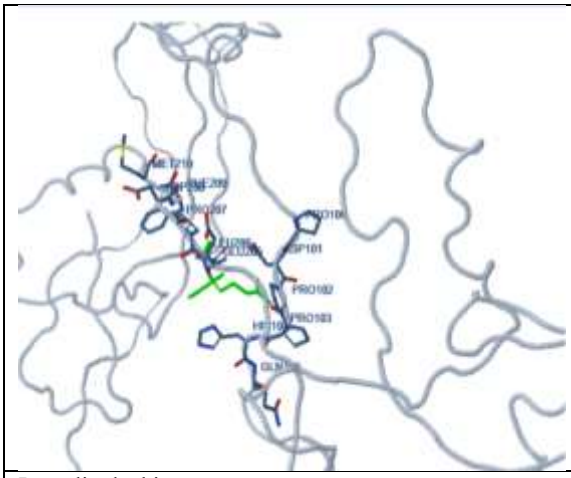
Kaempferol docking interacting of amino acids



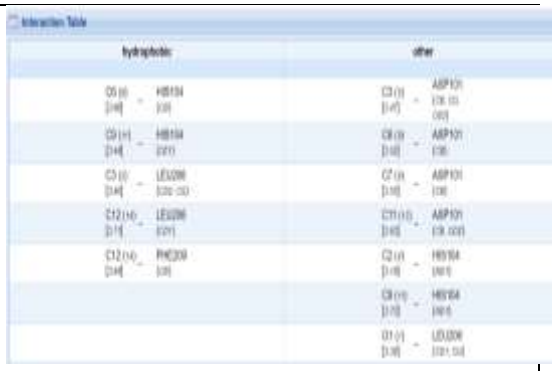
Linalool docking



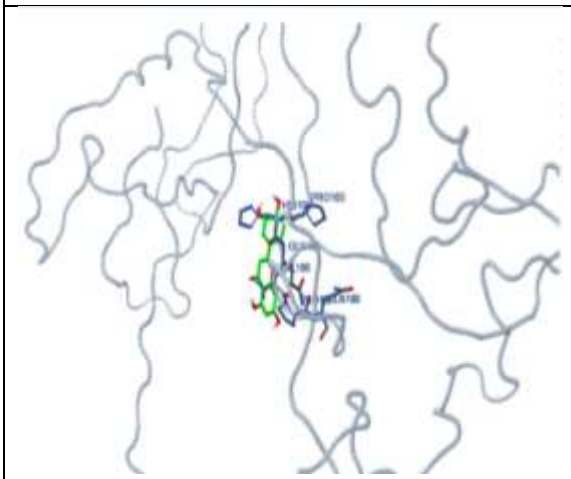
Linalool docking interacting of amino acids



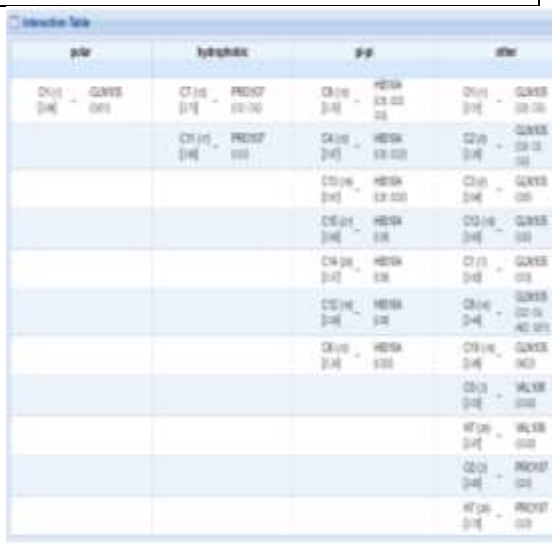
Luteolin docking



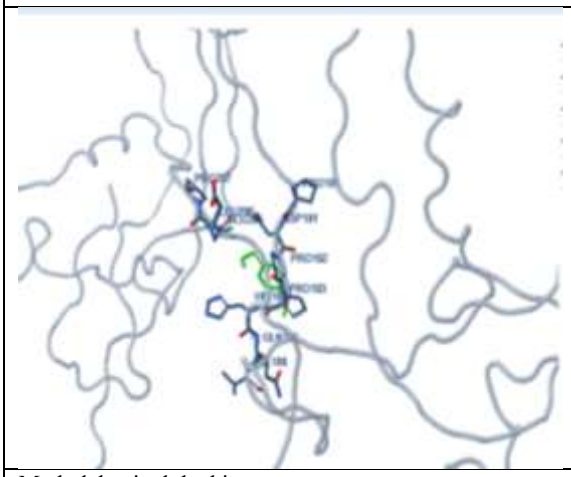
Luteolin docking interacting of amino acids



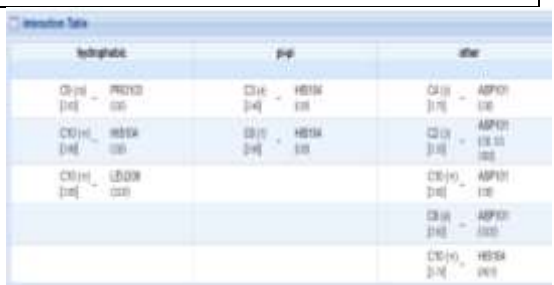
Linalyl acetate docking



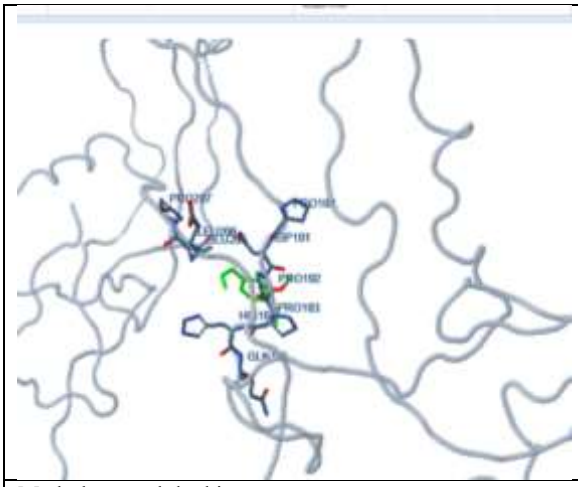
Linalyl acetate docking interacting of amino acids



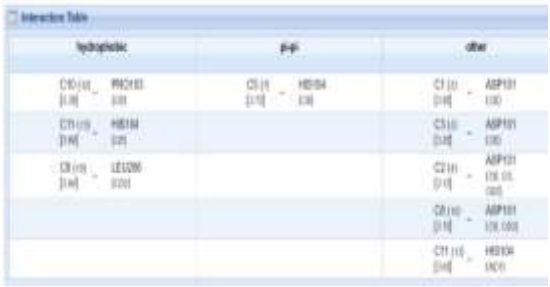
Methylchavicol docking



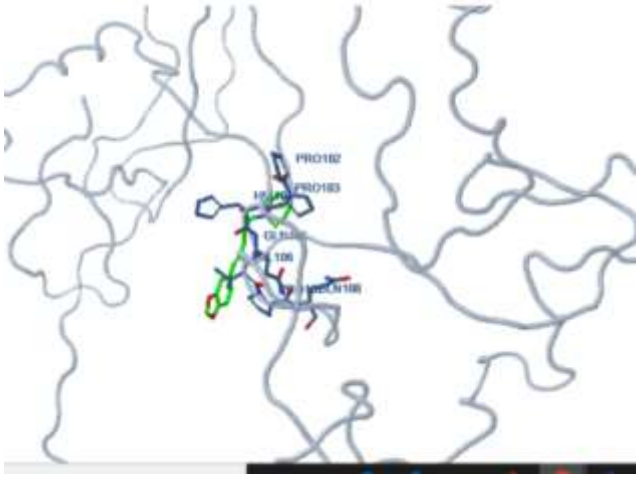
Methylchavicol docking interacting of amino acids



Methyleugenol docking

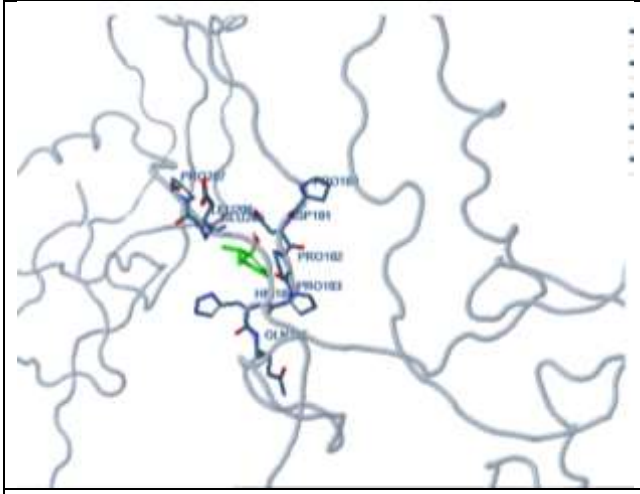


Methyleugenol docking interacting of amino acids

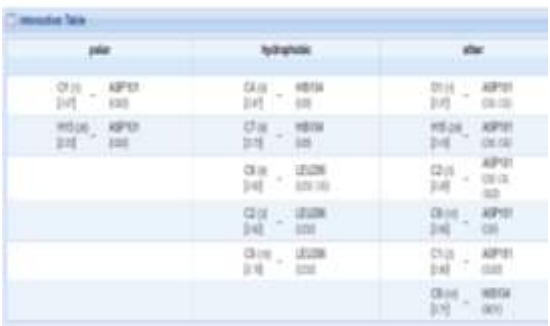


Piperine docking

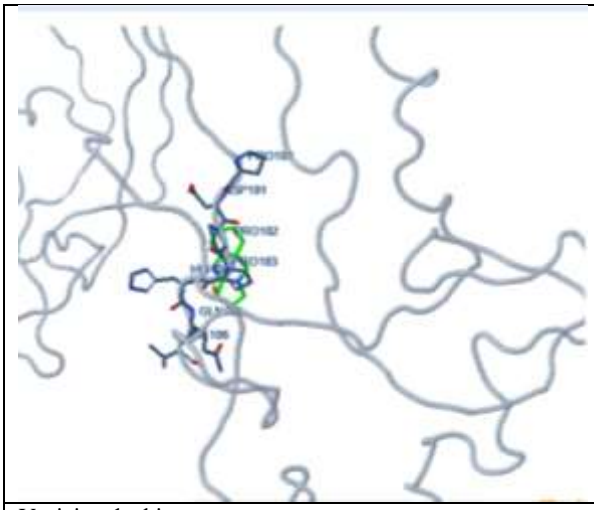
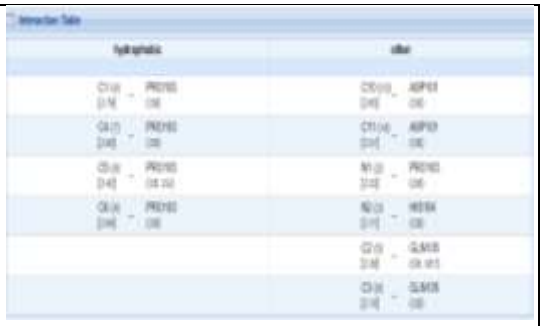
Piperine docking interacting of amino acids



Terpinen-4-ol docking



Terpinen-4-ol docking interacting of amino acids

	
Vasicine docking	Vasicine docking interacting of amino acids

4. Discussion

As per the docking results it is seen that Kaempferol docks with CCL2 with docking score of 4.67 with 19 interactions and NRAMP1 docks with Kaempferol with docking score of -3.64 with 16 interactions. CCL2 docks with Fisetin with docking score of -4.84 with 16 interactions and NRAMP1 docks with Fisetin with docking score of -3.73 with 21 interactions. CCL2 docks with Piperin with docking score of -5.04 with 16 interactions and NRAMP1 docks with Piperin with docking score of -4.16 with 17 interactions.

5. Conclusion

As per the results it is seen that phytocompounds Kaempferol, Fisetin and Piperin with docking score of -4.84 with 16 interactions. NRAMP1 docks with Fisetin with docking score of -3.73 with 21 interactions.CCL2 docks with Kaempferol with docking score -4.67 with 19 interactions. NRAMP1 docks with Kaempferol with docking score of -3.64 with 16 interactions.

Hence the above compound can be used as ligands for treating TUBERCULOSIS.

References

1. Bloom BR, Atun R, Cohen T, Dye C, Fraser H, Gomez GB, Knight G, Murray M, Nardell E, Rubin E, Salomon J, Vassall A, Volchenkov G, White R, Wilson D, Yadav P. Tuberculosis. In: Holmes KK, Bertozzi S, Bloom BR, Jha P, editors. Major Infectious Diseases. 3rd ed. Washington (DC): The International Bank for Reconstruction and Development / The World Bank; 2017 Nov 3. Chapter 11. PMID: 30212088.
2. Taylor GM, Stewart GR, Cooke M, Chaplin S, Ladva S, Kirkup J, Palmer S, Young DB. Koch's bacillus - a look at the first isolate of Mycobacterium tuberculosis from a modern perspective. Microbiology (Reading). 2003 Nov;149(Pt 11):3213-3220. doi: 10.1099/mic.0.26654-0. PMID: 14600233.
3. Delogu G, Sali M, Fadda G. The biology of mycobacterium tuberculosis infection. Mediterr J Hematol Infect Dis. 2013 Nov 16;5(1):e2013070. doi: 10.4084/MJHID.2013.070. PMID: 24363885; PMCID: PMC3867229.

4. Blackwell JM, Goswami T, Evans CA, Sibthorpe D, Papo N, White JK, Searle S, Miller EN, Peacock CS, Mohammed H, Ibrahim M. SLC11A1 (formerly NRAMP1) and disease resistance. Cell Microbiol. 2001 Dec;3(12):773-84. doi: 10.1046/j.1462-5822.2001.00150.x. PMID: 11736990; PMCID: PMC3025745.
5. Biswas SK, Mittal M, Sinha E, Singh V, Arela N, Bajaj B, Tiwari PK, Katoch VM, Mohanty KK. Exploring the Role of C-C Motif Chemokine Ligand-2 Single Nucleotide Polymorphism in Pulmonary Tuberculosis: A Genetic Association Study from North India. J Immunol Res. 2020 Dec 16;2020:1019639. doi: 10.1155/2020/1019639. PMID: 33381602; PMCID: PMC7759415.
6. 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.
7. 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.
8. Yahyazadeh R, Ghasemzadeh Rahbardar M, Razavi BM, Karimi G, Hosseinzadeh H. The effect of *Elettaria cardamomum* (cardamom) on the metabolic syndrome: Narrative review. Iran J Basic Med Sci. 2021 Nov;24(11):1462-1469. doi: 10.22038/IJBMS.2021.54417.12228. PMID: 35317114; PMCID: PM
9. Delogu G, Sali M, Fadda G. The biology of mycobacterium tuberculosis infection. Mediterr J Hematol Infect Dis. 2013 Nov 16;5(1):e2013070. doi: 10.4084/MJHID.2013.070. PMID: 24363885; PMCID: PMC3867229.
10. Feng L, Zhai YY, Xu J, Yao WF, Cao YD, Cheng FF, Bao BH, Zhang L. A review on traditional uses, phytochemistry and pharmacology of Eclipta prostrata (L.) L. J Ethnopharmacol. 2019 Dec 5;245:112109. doi: 10.1016/j.jep.2019.112109. Epub 2019 Aug 5. PMID: 31395303.
11. https://www.molinspiration.com/
12. Saha S, Ghosh S. Tinospora cordifolia: One plant, many roles. Anc Sci Life. 2012 Apr;31(4):151-9. doi: 10.4103/0257-7941.107344. PMID: 23661861; PMCID: PMC36447