Pak. J. Bot., 56(5): 2037-2044, 2024. DOI: http://dx.doi.org/10.30848/PJB2024-5(9)

IN SILICO PREDICTION OF DRUG BASED POTENTIAL INHIBITORS AGAINST COAT PROTEIN OF TOBACCO RINGSPOT VIRUS

MUHAMMAD NAVEED SHAHID[1*], SAIF ULLAH[1], ADIL JAMAL[2], NAILA MUKHTAR[3] AND SANA KHALID[4]

[1]*Department of Botany, Division of Science and Technology, University of Education, Lahore, Pakistan*
[2]*Sciences and Research, College of Nursing, Umm Al Qura University, Makkah, 715, Saudi Arabia*
[3]*Department of Botany, University of Okara, Okara, Pakistan*
[4]*Department of Botany, Lahore College for Women University, Lahore, Pakistan*
**Corresponding author's email: naveed.shahid@ue.edu.pk*

Abstract

Tobacco Ring Spot Virus (TRSV) is a highly damaging virus that affects a variety of crops. The viral pathogen is difficult to control since it is spread by a variety of vectors Viz., aphids, whiteflies, treehoppers and grasshoppers etc. The infectious nature of TRSV is due to its coat protein (CP), which helps to protect the virus genome against chemicals. As a result, it is necessary to break its CP to control the viral disease. The CP is conserved and plays a very significant role in the formation of the viral capsid. Moreover, vector specificity is also determined by CP. Virtual screening is the most accurate method to predict inhibitors for active site residues in CP. For active site prediction on CP, we used many servers, including SOMPA, ConSurf, ConCavity, SWISS-MODEL, PHYRE2, 3-D refinement, MTI Open Screen, PyRx, and ProSA COACH. Furthermore drugs that are suitable for these active sites were identified after active site prediction. Twenty drugs, each with a unique ZINC ID number, were identified. After assessing their ADMET characteristics and carcinogenicity, suitable drugs were chosen. One of the twenty most favorable drugs for TRSV is ZINC ID-597691, which has the best ADMET characteristics and carcinogenic value. In this study, we identified a drug with ZINC ID-597691 as the most effective treatment for TRSV-related infections that has risen.

Key words: Coat proteins, Tobacco ringspot virus, Virtual analysis, Inhibitors, Drug design.

Introduction

Tobacco ringspot virus (TRSV) belongs to the genus Nepovirus, subgroup A and family Secoviridae, and subfamily Comovirinae (Sanfacon *et al*., 2015; Anon., 2017; Lefkowitz *et al*., 2018) is one of the most prevalent and dangerous plant viruses (Almedia, 1980). It emanated first from the Midwest and East regions of the United States (Anon., 1997). This virus has now spread around the world, infecting a broad variety of woody and non-woody hosts (Abdalla *et al*., 2012; Kundu *et al*., 2015) and causing major infections in tobacco, soybeans, and tomatoes as well as in cucurbits, blueberries, and grapevines (Sinclair & Backman, 1989). It has the ability to affect more than three hundred plant species as well as several economically important crops. This viral infection has presented a serious danger to the productivity and quality of various crops, resulting in considerable yield losses (Hill & Whitham, 2014). It causes necrotic ringspots and systemic chlorotic in tobacco, mottled leaves in cucurbits (Abdalla *et al*., 2012), bud blight disease in soybeans (Demski & Kuhn, 1989), chlorotic spots, rings, and red blotches with deformed leaves in berries (Mitra *et al.,* 2021).

TRSV is an isometric, non-enveloped and 25 to 30 nm in diameter and exhibits icosahedral symmetry (Murant *et al*., 1996). The genome is made up of single-stranded polyadenylated RNA and is bipartite in the positive sense. TRSV consists of two RNA molecules with distinct lengths: the first RNA has a length of 7514 nucleotides, while the second has a length of 3929 nucleotides (Dunez & Gall, 2011; Sanfacon *et al*., 2012; Thompson *et al*., 2017). However, a distinct virion layer (coat protein/capsid protein) exists that surrounds each RNA molecule of identical diameter.

Antiviral treatment has made significant progress in enhancing the quality of life of infected plants during the last few decades. However, because of the long-term use of the same medicine and its associated adverse effects, it is necessary to establish new pharmaceuticals with unique mechanisms of action. The increasing of TRSV infection in various plants is posing an increasing threat throughout the world. The infection leads to spreads of various type of diseases in different crops which result to the loss of productivity. The TRSV itself does not pose a direct threat to human health, it can have significant economic and agricultural impacts, which in turn can affect global food security and trade. For the first time, infection with this virus was found in the tobacco (*Nicotiana tabacum*) plant by Fromme and his coworker in 1927. It is a significant virulent virus that is widely spread, causes serious infections/ diseases in different economic crops and urgent/ immediate need to be control. Its infection is difficult to manage because it is spread by a variety of vectors. The host range of TRSV encompasses both plants and insects due to its unique ability of host shifting (Li *et al*., 2014). Virtual screening is usually the most accurate method for predicting the residue of the active site using inhibitors. The entire CP of the examined molecule encodes approximately 513 amino acids having a molecular mass of 56.99 KDa (Geourjon & Deleage, 1995). The CP has a highly conserved sequence and is critical to the virus's life cycle. Because of their highly conserved sequence across successive generations, capsid proteins can be altered or damaged by a variety of small compounds or medicines. Drugs can have an effect both in the beginning and late stages of viral infection and replication cycles. As a result of these effects, several small compounds that can attach to CP have been discovered, with a few of them showing a vital role in preclinical and therapeutic research (Fu *et al*., 2019).

The primary goal of obtaining biological responses in medication development is to identify chemical species that are more capable of binding to proteins. Virtual screening offers an accelerated method to detect an effective drug. A diversified and developed chemical database is utilized to assess distinct protein groups and complex proteins and receptors for excellent performance (Ghosh *et al*., 2006). For the prediction of residual sites (ligand binding sites) in CP of TRSV to bind with ligand (inhibitor) and for docking study, several kinds of servers have been used (ConSurf, ConCavity, etc.). The goal of this research is to find possible inhibitors using various tools that can target the specific active sites of the virus coat protein and subsequently impede viral replication in affected plants and denying their role to act as reservoir for the vectors.

Material and Methods

Retrieval of protein sequence and secondary structure prediction: The amino acids sequences of chain A, a coat protein with the identification number 1A6C, were retrieved from the protein data bank in pdb format. These sequences correspond to the coat protein (CP) of the tobacco ringspot virus (TRSV) strain NC-87. The CP with chain A was selected throughout several coat proteins of TRSV due to its large size. The coat protein with ID 1A6C has been uploaded to the SOPMA server (http:// www.ibcp.fr/predict.html) to predict the secondary structure of CP. The sequence with the name was pasted on the SOPMA server for output (Eisenhaber *et al*., 1995).

Tertiary structure prediction, homology modeling, and structure refinement: The tertiary structure was predicted by using SWISS-MODEL (https://swissmodel. expasy.org/) and PHRE2 web server (http: //www.sbg. bio.ic.ac.uk/phyre2). For refinement, the sequence of CP was uploaded in FASTA format in a 3D refine server (http://sysbio.rnet.missouri.edu/3Drefine/). The server generated five refined TRSV protein structures. The fifth structure was the most refined model with the lowest potential energy. Hence model number five was used for subsequent study.

Error recognition and protein consistency analysis: For error recognition, refined structure of the TRSV protein was uploaded to ProSAweb server (https://prosa.services. came.sbg.ac.at/prosa.php). Then the PDB file of coat protein was uploaded to Procheck to determine protein consistency analysis (https://servicesn. mbi.ucla.edu/ PROCHECK/). The final model was selected on the basis of G-score, several residues in the core, allowed regions, favored regions, and outlier regions.

Active site prediction and calculation of grid center coordinates: For prediction of active sites, the sequence of CP was uploaded on COACH server (https:// zhanglab.ccmb.med.umich.edu/COACH/). The PDB file of CP was uploaded to the Prank Web for the calculation of grid center coordinates (https://prankweb.cz/).

Virtual screening and lazy structure-activity: To find optimal ligands, protein and ligand files were submitted to MTI AutoDock web server in mol2 and pdb formats, respectively (https://bioserv.rpbs.univ-paris-diderot.fr/ services/MTiOpenScreen/). The chemical structures or SMILE strings were submitted to LAZAR (httpr://laza.in-silico.ch) for identification of lazy structure-activity relationship and determination of carcinogenicity of the structural molecules.

Toxicity identification and auto docking: For identification of toxicity, the chemical structures of twenty target drugs were analyzed by using ADMETSAE web tool (http://lmmd.ecust.edu.cn/admetsar1/). For autodocking purposes, ligands and CP sequences were uploaded in pdbqt and PDB formats, respectively, within PyRx software (https://pyrx.sourceforge.io/). The results of PyRx software were visualized through Drug Discovery Studio (https://discover.3ds.com/discovery-studio-visualizer-download).

Results and Discussion

Chain A, accession ID 1A6C of CP, consists of 513 amino acids with a 56.99 KDa mass. The SOPMA server was used to generate the protein secondary structure (Eisenhaber *et al*., 1995). In this research work, we predicted different inhibitors against the chain A coat protein of TRSV strain NC-87 which was the most common and virulent. The Chain A coat protein with the accession ID 1A6C consisted of 513 amino acids having a 56.99 KDa mass. The results showed that the CP was consisted of 10.53% alpha-helix, 34.21% beta-strand, and 43.86% coils. The understanding of protein functions requires the knowledge about secondary structure. Its prediction is an important initial step toward predicting tertiary structure, along with providing data on protein functions, interactions, and activities (Ma *et al*., 2018). The 3D model of CP was predicted by PHYRE2 (Kelley *et al*., 2015) and SWISS-MODEL (Waterhouse *et al*., 2018), which predicted the tertiary structure (Biasini *et al*., 2014) (Fig. 1). Tertiary structure accuracy in resulting model was assessed by the GMQE score. As a consequence, the best three-dimensional model with 100% sequence accuracy was found**.** An automated server performed each step, and when the model was built, three-dimensional structure was downloaded in a PDB file (Guex *et al*., 2009). It is often required to establish three-dimensional structure of proteins to recognize their functions at a molecular level (Yang *et al*., 2018).

The refining of CP was done with the aid of a 3D refine server, which resulted in the 5 refined TRSV protein structures (Bhattacharya *et al*., 2016). All models are listed in order of potential energy base (Table 1). The best CP has a lower potential energy that is closer to its original structure. The structure that came in 5th place was an extremely refined model having the lowest potential energy. The 3D models of CP generated by online prediction servers were passed to an automated server (MODREFINER) for structure refinement and energy minimization (Xu & Zhang, 2011). The acquired results demonstrated the comparability between the modeling studies of the provided molecule (Rollinger *et al*., 2008). As a result, a 5th model was used for the next study. For the Error identification in coat protein, the ProSA-web (The protein structural analysis program (ProSA) is the most effective tool

ever for determining the validity of proteins as well as their structural prediction and modeling (Wiederstein & Sipple, 2007) gives three graphs: the first is a "Z score" that is used for model quality indication, the second is a "plot of residue score" that is used for local model quality, and the third is an "interactive molecular viewer" that is used for checking the target protein's 3D structure (Fig. 2A, B). The model quality was indicated by the Z score, which assesses the change in total energy of the structural form of the energy distribution that excretes and is generated from the mechanism. The CP had a Z score of -7.96. The results of Ramachandran plot analysis and sequence position plots revealed that the SWISS-MODEL server was developed as an excellent structural model for CP. The Ramachandran plot produced by the Pro-check server (Nageswara *et al*., 2019) showed that 70.8% of residues of SWISS-MODEL (Biasini *et al*., 2014) created a model of CP, which was located in the favored region of the Ramachandran plot as well as with the minimum percentage (additional allowed region, 23.5%; general allowed region, 3.6%) of residue inside the allowed region and (3.0%) in the outlier regions (Fig. 3). The Ramachandran plot analysis provides a score for the amino acid residues located within the favored region. The higher the proportion of amino acid residues within these favored regions, the greater the stability of the predicted structure. On the other hand, Ramachandran outliers represent amino acid residues that are less conducive to structural stability; in such cases, a smaller proportion is preferred (Chauhan *et al*., 2023).

The identification of particular ligands is challenging for a variety of reasons, including the high cost of experimental ligand identification. The structural information and ligand–protein binding details for 40% of proteins in the Protein Data Bank are missing (Yang *et al*., 2013). As a result, the prediction of ligand and protein binding information is critical for medicinal and biological research. To predict protein–ligand binding sites *in silico*, a variety of techniques have been used. Among them, the COACH (Yang *et al*., 2013) server plays a crucial function in the prediction of these active sites. COACH is a web server approach that predicts active sites for ligands that are more suitable. It incorporates predictions from several approaches, such as TM-SITE, COFACTOR (Roy *et al*., 2012), S-SITE, Concavity (Capra *et al*., 2009), and FINDSITE (Brylinski & Skolnick, 2008). All server results were compiled, and the foremost active sites were used to identify the ligand-binding sites (Table 2). The prediction of active sites was necessary after a model's native state was confirmed. Consequently, ligand uniqueness may be identified. This server produced forty-three active sites ((GLU)115, (TRP)116, (GLN)117, (THR)188, (ALA)190, (ASN)336, (THR)337, (ILE)33, (ILE)37, (THR)57, (THR)59, (CYS)76, (THR)77, (PHE)78, (GLU)100, (ASN)119, (LEU)128, (CYS)129, (TRP)131, (GLN)137, (LEU)140, (HIS)163, (THR)178, (ILE)244, (GLY)246, (SER)247, (VAL)248, (PHE)260, (ILE)262, (MET)284, (GLY)286, (PHE)290, (ILE)292, (LEU)311, (ILE)313, (PHE)333, (VAL)333, (HIS)334, (ILE)335, (ASP)39, (GLN)46, (ASN)130, (PRO)177)) for CP of TRSV which were more favorable to dock with ligands. The prediction of active sites in a protein requires the prediction of the 3-dimensional structure of that protein (Goh & Foster, 2000). The active site of a protein is made up of a catalytic (to increase the reaction) and binding site (for attachment).

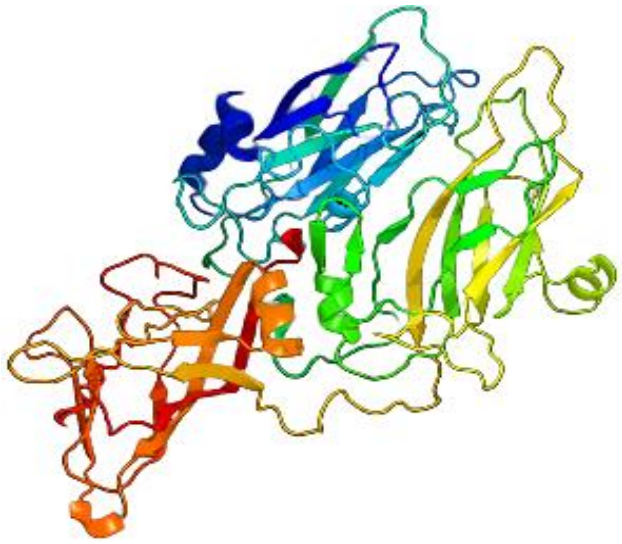


Fig. 1. Tertiary Structure of Coat Protein Chain A predicted by PHRE2 web server and Swiss model.

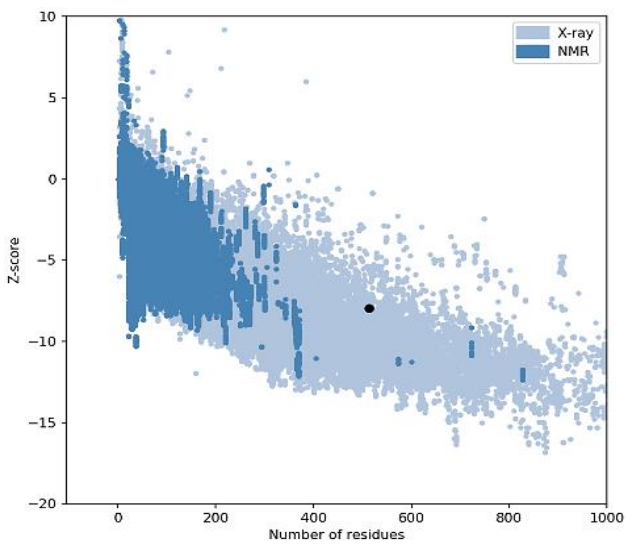


Fig. 2A. Represents the value of Z- score in the residues of the coat protein.

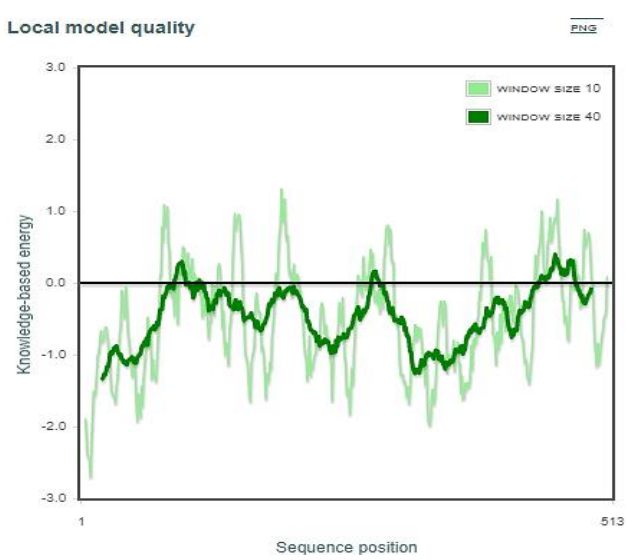


Fig. 2B. ProSAweb calculated energy profile of CP.

Table 1. 3D refine server predicted refinement models of coat protein with different energy scores.

Model #	$3D^{refine}$ Score	GDT-TS	GDT-HA	RMSD (Å)	Mol Probity	RW Plus
5	22626.7	0.9995	0.9859	0.308	3.036	-106097.891799
4	22944.3	1.0000	0.9898	0.290	3.035	-105865.040395
3	23471.4	1.0000	0.9932	0.265	3.058	-105679.886359
2	24738.6	1.0000	0.9976	0.229	3.021	-105454.633027
1	30456.9	1.0000	1.0000	0.169	3.085	-105201.116728

Table 2. List of 43 ligand-binding active sites in coat protein predicted by several servers.

Server	Rank	C-Score	Predicted binding site residues
COACH	1	0.06	(GLU)115, (TRP)116, (GLN)117, (THR)188, (ALA)190, (ASN)336, (THR)337
TM-SITE	1	0.19	(ILE)33, (ILE)37, (THR)57, (THR)59, (CYS)76, (THR)77, (PHE)78, (GLU)100, (GLN)117, (ASN)119, (LEU)128, (CYS)129, (TRP)131, (GLN)137, (LEU)140, (HIS)163
S-SITE	1	0.11	(GLU)115, (TRP)116, (GLN)117, (THR)178, (THR)188, (ALA)190, (ASN)336, (THR)337
ConCavity	1	0.34	(ILE)244, (GLY)246, (SER)247, (VAL)248, (PHE)260, (ILE)262, (MET)284, (GLY)286, (PHE)290, (ILE)292, (LEU)311, (ILE)313, (PHE)331, (VAL)333, (HIS)334, (ILE)335
COFACTOR	1	0.01	(ASP)39, (GLN)46, (ASN)130
FINDSITE	1	0.05	(THR)57, (THR)59, (GLU)115, (TRP)116, (GLN)117, (PRO)177, (THR)188, (ALA)190, (ASN)336, (THR)337

Table 3. ADMET characteristics of different ligands in order to examine their toxic effects.

Ligand	Blood-brain barrier	Human-intestinal absorption	Caco-2-permeability	P-glycoprotein substrates	Toxicity (LD50 in kgmol−1)
1	0.9875	0.9771	0.7545	0.5000	3.391
2	0.9788	0.9775	0.8405	0.7717	2.395
3	0.9795	0.8734	0.8786	0.6598	2.875
4	0.9795	0.8734	0.8786	0.6598	2.875
5	0.9638	0.9915	0.8037	0.7085	2.346
6	0.9875	0.9771	0.7545	0.5000	3.391
7	0.9840	0.9919	0.8682	0.6103	3.291
8	0.9795	0.8734	0.8786	0.6598	2.875
9	0.9840	0.9919	0.8682	0.6103	3.291
10	0.9735	0.9619	0.8521	0.8261	3.741
11	0.9782	0.9619	0.8309	0.7531	4.009
12	0.9737	0.9266	0.8562	0.8371	3.435
13	0.9834	0.9411	0.8251	0.7774	3.013
14	1.0000	0.9863	0.6396	0.5369	2.668
15	0.9814	0.9907	0.8346	0.7830	3.176
16	0.9351	0.9651	0.7936	0.7953	3.36
17	0.9834	0.9411	0.8251	0.7774	3.013
18	0.9930	0.6804	0.8138	0.9343	2.539
19	0.9546	0.9834	0.8548	0.8347	2.54
20	0.9829	0.9632	0.8626	0.5547	2.751

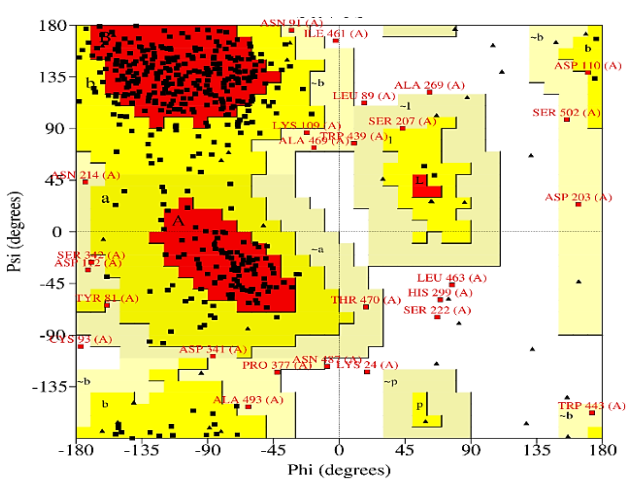
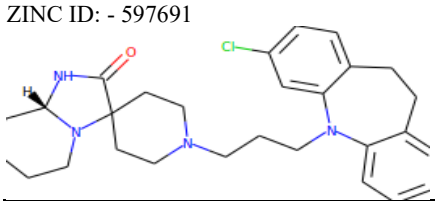


Fig. 3. Represents the position of coat protein residues which present 98% in most favorably allowed regions (bright and light yellow and red color) while 2% residues (lightest yellow) lie in disallowed regions in the Ramachandran plot.

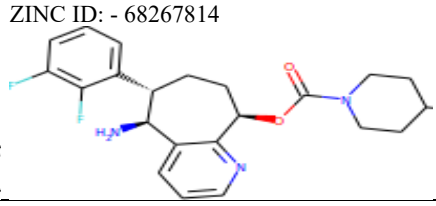
In the binding sites, residues form hydrogen bonds or hydrophobic linkages or transient covalent bonds with the substrate to form a complex of protein and substrate (Chtterjee *et al*., 2017). Catalysis can commence after the binding is complete and aligned in the active site. Some residues of the catalytic site occur naturally near the binding site, and a few of them can play both functions in binding and catalysis (Chatterjee *et al*., 2011). Nearly 1501 medicines were identified against the binding site of CP using the ZINC database. For docking purposes, only a small number of top-ranked compounds are chosen (Li & Shah, 2017).

As a consequence, 20 of the most suitable medicines were identified as CP inhibitors. Structures such as (chemical), ZINC identification scores, and numbers are given in Fig. 4. The best docking scores were used to show all of the selected compounds that have a high binding affinity for various coat proteins. All compounds are chosen for the creation of efficient antiviral medicines based on their unique binding affinity (Joshi *et al*., 2021).

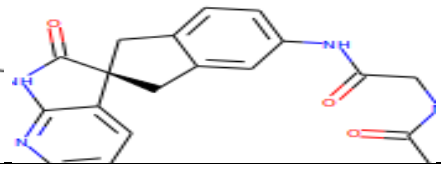
Ligand 1 (Mosapramine)
ID Score (-10.90000)
ZINC ID: - 597691



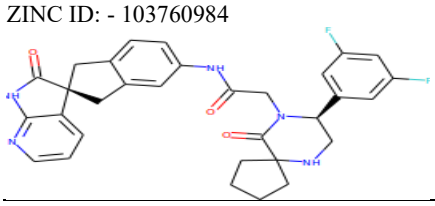
Ligand 2 (Rimegepant)
ID Score (-10.90000)
ZINC ID: - 68267814



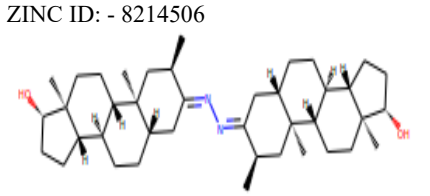
Ligand 3 (Mk3207)
ID Score (-10.80000)
ZINC ID: - 103760978



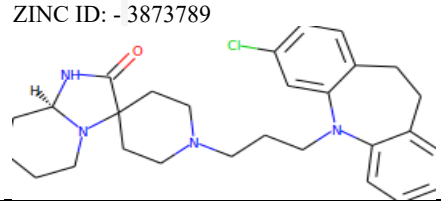
Ligand 4 (Mk3207)
ID Score (-10.70000)
ZINC ID: - 103760984



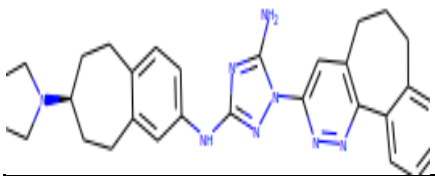
Ligand 5 (Bolazine)
ID Score (-10.60000)
ZINC ID: - 8214506



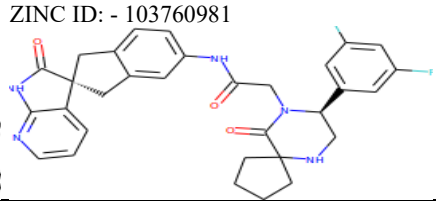
Ligand 6 (Mosapramine)
ID Score (-10.60000)
ZINC ID: - 3873789



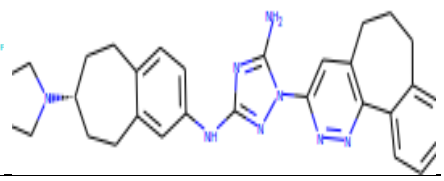
Ligand 7 (R428)
ID Score (-10.40000)
ZINC ID: - 51951669



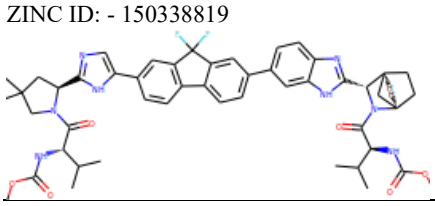
Ligand 8 (Mk3207)
ID Score (-10.40000)
ZINC ID: - 103760981



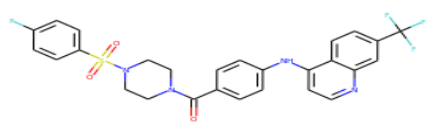
Ligand 9 (R428)
ID Score (-10.40000)
ZINC ID: - 51951668



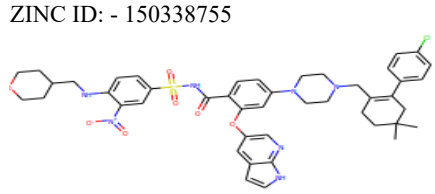
Ligand 10 (Ledipasvir)
ID Score (-10.30000)
ZINC ID: - 150338819



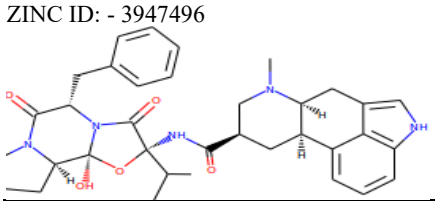
Ligand 11 (Losulazine)
ID Score (-10.30000)
ZINC ID: - 4216779



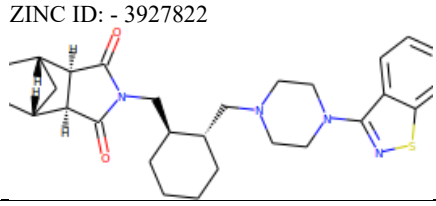
Ligand 12 (Venetoclax)
ID Score (-10.20000)
ZINC ID: - 150338755



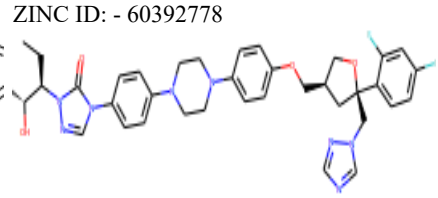
Ligand 13 (Dihydroergocristine)
ID Score (-10.20000)
ZINC ID: - 3947496



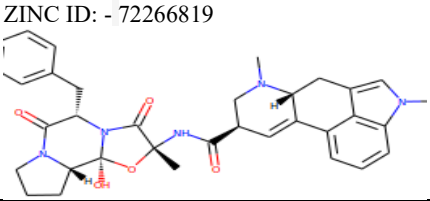
Ligand 14 (Lurasidone)
ID Score (-10.20000)
ZINC ID: - 3927822



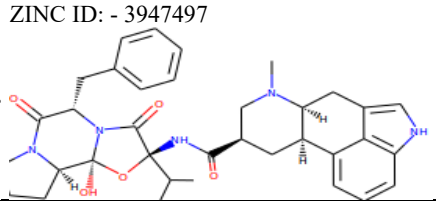
Ligand 15 (Posaconazole)
ID Score (-10.20000)
ZINC ID: - 60392778



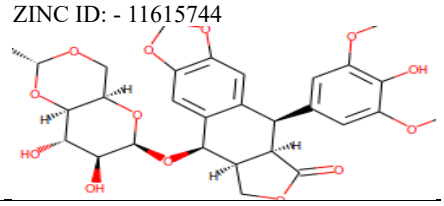
Ligand 16 (Metergotamine)
ID Score (-10.20000)
ZINC ID: - 72266819



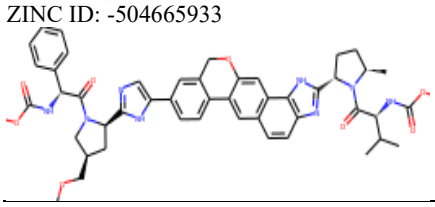
Ligand 17 (Dihydroergocristine)
ID Score (-10.10000)
ZINC ID: - 3947497



Ligand 18 (Etoposide)
ID Score (-10.10000)
ZINC ID: - 11615744



Ligand 19 (Velpatasvir)
ID Score (-10.10000)
ZINC ID: -504665933



Ligand 20 (Conivaptan)
ID Score (-10.10000)
ZINC ID: - 12503187

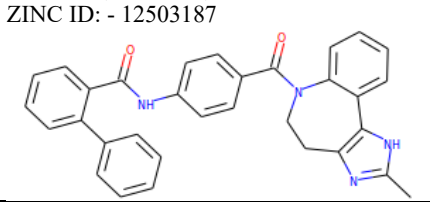


Fig. 4. This is the top twenty most appropriate ligands 2D structures with their docking scores selected for coat protein. All ligands were chosen based on their ability to attach to active sites. Their binding affinity increase as the binding score decrease.

Table 4. ADMET properties and Carcinogenicity profile for selected top twenty ligands.

Ligand	Blood-brain barrier	Human-intestinal absorption	Caco-2 permeability	P-glycoprotein substrates	Toxicity (LD50 in molkg^{-1})	Carcino-genicity
1	0.9875	0.9771	0.7545	0.5000	3.391	NC
2	0.9788	0.9775	0.8405	0.7717	2.395	NC
3	0.9795	0.8734	0.8786	0.6598	2.875	NC
4	0.9795	0.8734	0.8786	0.6598	2.875	NC
5	0.9638	0.9915	0.8037	0.7085	2.346	NC
6	0.9875	0.9771	0.7545	0.5000	3.391	NC
7	0.9840	0.9919	0.8682	0.6103	3.291	NC
8	0.9795	0.8734	0.8786	0.6598	2.875	NC
9	0.9840	0.9919	0.8682	0.6103	3.291	NC
10	0.9735	0.9619	0.8521	0.8261	3.741	NC
11	0.9782	0.9619	0.8309	0.7531	4.009	NC
12	0.9737	0.9266	0.8562	0.8371	3.435	NC
13	0.9834	0.9411	0.8251	0.7774	3.013	NC
14	1.0000	0.9863	0.6396	0.5369	2.668	NC
15	0.9814	0.9907	0.8346	0.7830	3.176	NC
16	0.9351	0.9651	0.7936	0.7953	3.36	NC
17	0.9834	0.9411	0.8251	0.7774	3.013	NC
18	0.9930	0.6804	0.8138	0.9343	2.539	NC
19	0.9546	0.9834	0.8548	0.8347	2.54	NC
20	0.9829	0.9632	0.8626	0.5547	2.751	NC

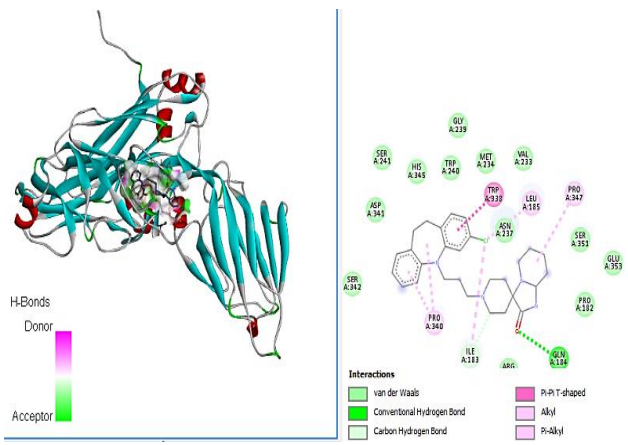


Fig. 5. Shows the 3D binding complex in the left side part while in the right-side part 2D binding complex of coat protein along with the best ligand. Both complexes represent the binding interaction of the best ligand with coat protein.

Through the AdmetSAR server (Cheng *et al*., 2012), top 20 compounds were more examined for toxic effects and drug ability, and the results were obtained (Table 3). The ligands underwent an evaluation of their absorption, metabolism, distribution, excretion, and toxicity (ADMET) profile, which assessed the drugs overall ADMET effectiveness. The ligands showing promise in terms of favorable pharmacokinetic and pharmacodynamics characteristics are more likely candidates for future drug development (Yousuf *et al*., 2017).

Based on their minimum toxicity and high ADMET (Merlot, 2010) properties, Zn ID-597691 had exhibited top docking result (-10.90 kcal/mol) out of twenty tested compounds. This target CP can cross the blood–brain barrier, which may be used as a criterion for selecting the optimal ligand. After docking ligands with CP using a laser (https://lazar.in-silico.ch/predict), the best ligand (Zn ID-597691) was shown to have the minimum toxicity and carcinogenic effects (Table 4) and was chosen as the final best ligand to bind with CP. ADMET screening plays a crucial role in drug discovery and development, as it facilitates the achievement of the desired balance of properties that ensure safety and efficacy during lead/hit discovery. The compounds obtained after virtual screening were selected based on their drug-like characteristics (Chauhan *et al*., 2023). By using PyRx software (Dallakyan & Olson, 2015), the final ligand was redocked against the CP of TRSV. The complexes of the final ligand and CP were viewed through Drug Discovery Studio (Jejurikar & Rohane, 2021) (Fig. 5). The interpretation of molecular docking plays a vital role in the determination of results. Docking is a widely employed technique for investigating bimolecular interactions, primarily focusing on protein-ligand interactions and the exploration of molecular mechanisms (Baroroh *et al*., 2023). The top and most effective ligand for the CP of TRSV was found to be the first ligand with Zn ID-597691.

The CP is characterized by docking and binding site details obtained from PyRx, which compares their lowest toxic effects, absorption, comparative docking, and binding energy by utilizing the grid resolution (3.50 Å) and Auto Docking. Docking was set moldable docking as well as excellent precision. This interaction mechanism was studied by describing the size of the ligand's binding site and the docking box, Z = 121.2615, X = -18.7121, Y = 31.9344, and 404040. In the present investigation, our goal was to perform a computational-based virtual analysis and we have found the best binding affinity molecule (Zn ID-597691) with -10.40000 kcal/mol residual energy.

Conclusion

In this study, we screened the chain A coat protein of tobacco ringspot virus strain NC-87 for the prediction of potential inhibitors. The secondary structure was predicted, which aided in the prediction of tertiary structure. The

refinement of the CP consequences resulted in five refined models, with the fifth model having the lowest energy and being selected for further study. By using the coach server, we predicted 43 active sites in CP that might act as binding sites for ligands. Through virtual screening, we obtained 1501 compounds, 20 of which were selected on the basis of docking scores for CP inhibitors. These 20 compounds were further examined for toxic effects and drug ability based on their minimum toxicity and high ADMET properties. Zn ID-597691 had exhibited the top docking result (-10.90 kcal/mol) out of these 20 tested compounds and was selected as the final best ligand with -10.40000 kcal/mol binding energy to make the protein complex.

Acknowledgement

The authors are highly grateful to the University of Education, Lahore for providing a platform to conduct this research.

References

Abdalla, O.A., B.D. Bruton, W.W. Fish and A. Ali. 2012. First confirmed report of tobacco ring spot virus in cucurbits crops in Oklahoma. *Plant Dis*., 96: 1705.

Almedia, A.M.R. 1980. Survey of soybean common mosaic and bud blight viruses in different regions of Parana State. *Fitopatol. Bras*, 5(1): 125-128.

Anonymous. 1997. EPPO/CABI, Tobacco ringspot nepovirus. In: (Eds.): Smith, I.M., D.G. McNamara, P.R. Scott & M. Holderness. Quarantine Pests for Europe, 2nd edn, pp. 1357-1362. CAB International, Wallingford (GB) (1997).

Anonymous. 2017. PM 7/2 (2). Tobacco ringspot virus. *EPPO Bulletin*, 47: 135-145.

Baroroh, U., M. Biotek, Z. S. Muscifa, W. Destiarani, F.G. Rohmatullah and M. Yusuf. 2023. Molecular interaction analysis and visualization of protein-ligand docking using Biovia Discovery Studio Visualizer. *Indon. J. Comp. Biol.*, 2(1): 22-30.

Bhattacharya, D., J. Nowotny, R. Cao and J. Cheng. 2016. 3Drefine: An interactive web server for efficient protein structure refinement. *Nucl. Acids Res*., 44(1): 406-409.

Biasini, M, S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, F. Kiefer, T.G. Cassarino, M. Bertoni and L. Bordali and T. Schewede. 2014. Schwede. SWISS-MODEL: Modeling protein tertiary and quaternary structure using evolutionary information. *Nucl. Acids Res.*, 42: 252-258.

Brylinski, M. and J. Skolnick. 2008. A threading-based method (FINDSITE) for ligand binding site prediction and functional annotation. Proceedings of the National Academy of Sciences USA, 105: 129-134.

Capra, J.A., R.A. Laskowski, J.M. Thornton, M. Singh and T.A. Funkhouser. 2009. Predicting protein-ligand binding sites by combining evolutionary sequence conservation and 3D structure. *Plos Comp. Biol*., 5(12): 1-18.

Chatterjee, A., U.K. Roy and D. Halder. 2017. Protein active site structure prediction strategy and algorithm. *Int. J. Curr. Eng. Technol*., 7(3): 1092-1096.

Chatterjee, P., S. Basu, M. Kundu, M. Nasipuri and D. Plewczynski. 2011. PPI_SVM: Prediction of protein-protein Interactions using Machine Learning, Domain-domain Affinities and Frequency Tables. *Cell. Mol. Biol. Lett*., 16: 264-278.

Chauhan, A., N. Sangwan, J. Singh, A. Prakash, B. Medhi and P.K. Avti. 2023. Allosteric modulation of conserved motifs and helices in 5HT2BR: Advances drug discovery and therapeutic approach towards drug resistant epilepsy. *J. Biomol. Struct. Dyn.*, 1-14.

Cheng, F., W. Li, Y. Zhou, J. Shen, Z. Wu, G. Liu and Y. Tang. 2012. AdmetSAR: a comprehensive source and free tool for assessment of chemical ADMET properties. *J. Chem. Inf. Model*, 52: 3099-3105.

Dallakyan, S. and A.J. Olson. 2015. Small-molecule library screening by docking with PyRx. In: (Eds.): Hempel, J., C. Williams & C. Hong. Chemical Biology. Methods in Molecular Biology, vol. 1263. Humana Press, New York, NY.

Demski, J.W. and C.W. Kuhn. 1989. Tobacco ringspot virus. In: (Ed.): Hartman, G.L. Compendium of Soybean Diseases, 3rd ed. *Am. Phytopath. Soc., St. Paul*, MN, pp. 57-59.

Dunez, J. and O.L. Gall. 2011. Nepovirus. Comoviridae. In: The Springer Index of Viruses. Springer, New York, pp 361-369.

Eisenhaber, F., B. Persson and P. Argos. 1995. Protein structure prediction: recognition of primary, secondary and tertiary structural features from amino acid sequence. *Crit. Rev. Biochem. Mol*., 30: 1-94.

Fu, D.Y., Y.R. Xue, X. Yu and Y. Wu. 2019. Anti-virus reagents targeting the capsid protein assembly. *J. Mater. Chem*., B 7(21): 3331-3340.

Geourjon, C. and G. Deleage. 1995. ANTHEPORT 2.0; a three dimensional module fully coupled with protein sequence analysis methods. *J. Mol. Graph*, 13: 209-212.

Ghosh, S., A. Nie, J. An and Z. Huang. 2006. Structure-based virtual screening of chemical libraries for drug discovery. *Curr. Opin. Chem. Biol.*, 10(3): 194-202.

Goh, G. and J.A. Foster. 2000. Evolving molecules for drug design using genetic algorithm, *Proc. Int. Conf. on Genetic & Evol*., Computing, Morgan Kaufmann, 27-33.

Guex, N., M.C. Peitsch and T. Schwede. 2009. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis*, 30(1): 162-173.

Hill, J.H. and S.A. Whitham. 2014. Control of virus diseases in soybeans. In: (Eds.): Loebenstein, G. & N. Katis. Advances in virus research (pp. 355-390). Salt Lake City: Academic Press.

Jejurikar, B.L. and S.H. Rohane. 2021. Drug Designing in Discovery Studio. *Asian J. Res. Chem.*, 14(2): 135-138.

Joshi, T., P. Sharma, T. Joshi, H. Pundir, S. Mathpal and S. Chandra. 2021. Structure-based screening of novel lichen compounds against SARS Coronavirus main protease (Mpro) as potentials inhibitors of COVID-19. *Mol. Divers.*, 25(3): 1665-1677.

Kelley, L.A., S. Mezulis, C.M. Yates, M.N. Wass and M.J. Sternberg. 2015. The Phyre2 web portal for protein modeling, prediction, and analysis. *Nat. Protoc.*, 10(6): 845-858.

Kundu, J.K., S. Gadiou, G. Schlesingerova, M. Dziakova and V. Cermak. 2015. Emergence of quarantine Tobacco ringspot virus in Impatiens walleriana in the Czech Republic. *Plant Prot. Sci.*, 51: 115-123.

Lefkowitz, E.J., D. M. Dempsey, R.C. Hendrickson, R.J. Orton, S.G. Siddell and D.B. Smith. 2018. Virus taxonomy: The database of the international committee on taxonomy of viruses (ICTV). *Nucl. Acids Res.*, 46: D708-D717.

Li, J. L., R.S. Cornman, J.D., Evans, J.S., Pettis, Y., Zhao, C., Murphy and Y.P. Chen. 2014. Systemic spread and propagation of a plant-pathogenic virus in European honeybees, Apis mellifera. *mBio.*, 5(1): 10-1128.

Li, Q. and S. Shah. 2017. Structure-based virtual screening. In Protein Bioinformatics. Humana Press, New York, NY, 111-124.

Ma, Y., Y. Liu and J. Cheng. 2018. Protein secondary structure prediction based on data partition and semi-random subspace method. *Sci. Rep.*, 8(1): 1-10.

Merlot, C. 2010. Computational toxicology–a tool for early safety evaluation. *Drug Discov. Today*, 15: 16-22.

Mitra, A., S. Jarugula, G.A. Hoheisel and A.R. Naidu, 2021. First report of Tobacco ringspot virus in high bush blueberry in Washington State. *Plant Dis.*, 105(9): 2739.

Murant, A.F., A.T. Jones, G.P. Martelli and R. Stace-Smith. 1996. Nepovirus: General properties of the disease, and virus identification. In: (Eds.): Harrison, B.D. & A.F. Murant. *The Plant Viruses*, 5: 99-137. New York: Plenum.

Nageswara, S., G. Guntuku and B.L. Yakkali. 2019. Purification, characterization, and structural elucidation of serralysin-like alkaline metalloprotease from a novel source. *J. Genet. Eng. Biotech.*, 17(1): 1-15.

Rollinger, J.M., T.M. Steindl, D. Schuster, J. Kirchmair, K. Anrain, E.P. Ellmerer and M. Schmidtke. 2008. Structure-based virtual screening for the discovery of natural inhibitors for human rhinovirus coat protein. *J. Med. Chem.*, 51(4): 842-851.

Roy, A., J. Yang and Y. Zhang. 2012. Cofactor: An accurate comparative algorithm for structure-based protein function annotation. *Nucl. Acids Res.*, 40: 471-477.

Sanfacon, H. 2015. Secoviridae: a family of plant picorna-like viruses with monopartite or bipartite genomes. *Encyclopedia of Life Sciences,* 1-14.

Sanfacon, H., T. Iwanami, A.V. Karasev, R. Van der Vlugt, J. Wellink and T. Wetzel. 2012. Family Secoviridae. In: Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses (Eds.): King, A.M.Q., M.J. Adams, E.B. Carstens & E.J. Lefkowitz, pp. 881-899. Elsevier Academic Press, San Diego, CA (US).

Sinclair, J.B. and P.A. Backman. 1989. *Compendium of soybean diseases* (No. BOOK). American Phytopathological Society.

Thompson, J.R., I. Dasgupta, M. Fuchs, T. Iwanami, A.V. Karasev, K. Petrzik, H. Sanfaçon, I.E. Tzanetakis, R. van der Vlugt, T. Wetzel and N. Yoshikawa. 2017. ICTV Report Consortium ICTV virus taxonomy profile: Secoviridae. *J. Gen. Virol.*, 98: 529-531.

Waterhouse, A., M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F.T. Heer, T.C. A P de Beer, L. Rempfer, R. Bordoli, Lepore and T. Schwede. 2018. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucl. Acids Res.*, 46: W297-W303.

Wiederstein, M., and M.J. Sippl. 2007. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucl. Acid. Res.*, 35(Web Server issue), 407-410.

Xu, D. and Y. Zhang. 2011. Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophys. J.*, 101(10): 2525-2534.

Yang, J., A. Roy and Y. Zhang. 2013. Protein–ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. *Bioinformatics,* 29(20): 2588-2595.

Yang, Y., J. Gao, J. Wang, R. Heffernan, J. Hanson, K. Paliwal and Y. Zhou. 2018. Sixty-five years of the long march in protein secondary structure prediction: the final stretch? *Brief Bioinform*., 19(3): 482-494.

Yousuf, Z., K. Iman, N. Iftikhar and M.U. Mirza 2017. Structure-based virtual screening and molecular docking for the identification of potential multi-targeted inhibitors against breast cancer. *Breast Cancer: Targ. Therap.*, 447-459.

(Received for publication 08 March 2023)