

Potential Compounds for the Inhibition of TMPRSS2

Muhammad Sohaib Roomi¹, Yaser Daanial Khan¹
F2018279051@umt.edu.pk, Yaser.khan@umt.edu.pk

¹Department of Computer Science, School of Systems and Technology, University of Management and Technology, Lahore, Pakistan

Abstract:

The ongoing search to contain and control the spread of COVID-19 disease focuses on discovering drugs or vaccines that can play an essential role in treating this contagious disease. This paper focuses on natural compounds that can play a vital role in the treatment of Covid-19. The study spans over the chemicals that have the potential to bind with the key residues of type II Transmembrane Protease Serine (TMPRSS2). TMPRSS2 can be termed as the catalyst that cleaves the spike glycoproteins of Sars-Cov-2, which causes the replication and spread of virus inside the human body by facilitating virus-cell fusions. Drugs like Camostat Mesylate, Aprotinin, and Rimantadine have been proposed as potential inhibitors of TMPRSS2. After screening large sets of phytochemicals and flavonoids extracted from plants, potential compounds have been tested, and a set of most effective and suitable compounds are chosen for further studies. These selected compounds are further analyzed in terms of binding with key residues as well as high binding affinity with TMPRSS2. The in silico analysis of possible chemical compounds is carried out by using docking, screening analysis, Molecular Dynamics, and Electrostatic Potential Simulations. Chemicals extracted from different plants are comparatively analyzed with drugs like Aprotinin, Camostat Mesylate, and Rimantadine.

Introduction:

COVID-19 – emerged from Wuhan, China, and spread across the world in a short period. It is caused by a virus named Sars-Cov-2. It transfers from person to person through close contact and small droplets that are produced due to coughing and sneezing. The main symptoms are fever, illness, Flu, sore throat, and fatigue. It has spread across 200 countries, with 10.5 million confirmed infected cases and 600,000 deaths. It is still spreading because there is no vaccine available for the treatment of the virus yet. The worrisome sign is that an infected person may not realize about infection and cause infection to other people that are close to him. To understand the structure and working of this disease, researchers and medical units are working day and night.

In the replication of disease, human protease TMPRSS2 plays a vital role because it helps the Spike to speed up the replication of viruses after Spike gains entry into cells through Human Ace2. This duplication increases the dominance of viruses in human cells. The structure of TMPRSS2 can play an essential role in discovering suitable compounds for the treatment. That's where antiviral drugs can play a significant role. These drugs link with the structural residues of type II Transmembrane Protease Serine, which mediate

by cleaving Spike protein and instigating membrane fusion. Natural plants contain substances that are used for different treatment purposes, and these plants have been used for this purpose for ages. Plants containing phytochemicals can have antiviral properties. A range of chemicals such as flavones, flavonoids, Kambo, Ginkgo, alkaloids, chlorophyllins, Ayurveda is known to have medicinal properties that have been used for a variety of viruses. Many publicly accessible databases like PubChem, ChemSpider, and ChEMBL have thousands of compounds' structures. Proper research and clinical trial can cause millions of dollars and result in a lot of wasted time to know the failure of the drug. In this study, we used accurate and efficient methods to check the appropriateness and effects of drugs by getting fast and robust results concerning the suitability of these compounds. This study gives an insight into the pathogens and potential inhibitors to help understand the compounds that can play a significant role in the discovery of drugs for the diseases.

Structure of TMPRSS2:

Type II Transmembrane Protease Serine found in human chromosome 21q22.3, 112, and it encodes a protein of 492 amino acids. It is a multi-domain transmembrane serine protease that contains two chains: a non-catalytic amino acid chain that is formed by amino acids 1-255 and a catalytic chain which contains amino acids 256-492. Usually, serine protease contains three active binding sites of the catalytic triad, which are: HIS 296, ASP 345, and SER 441. In humans, it is present in the prostate, lungs, and other tissues. Mostly it is shown that TMPRSS2 is associated with prostate cancer neglecting viral infections. However, it is a proven fact that duplication of Coronavirus is occurred due to the binding of viral proteins with the Human Receptors and further cleavage of the Spike glycoproteins by human cell proteases. The Structure of TMPRSS2 is shown in figure 1.

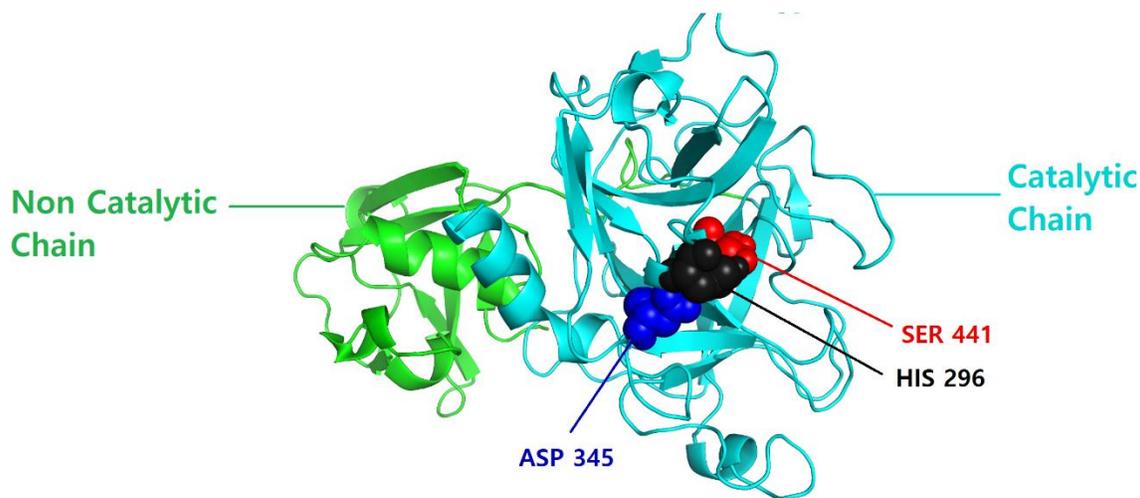


Figure 1 Structure of TMPRSS2

Catalysis Mechanism of TMPRSS2:

Type II Transmembrane Protease Serine performs the catalysis process using Serine, which is an active residue of the catalytic side chain. In addition to Serine, there are Histidine and Aspartate residues also which assist in the catalysis[1]. This protease is a cleave peptide bond and can cut specific molecules that it binds. Figure 2 shows the polypeptide chain of protein that the serine protease will cut between Carbon and Nitrogen. And this is the location for the peptide bond.

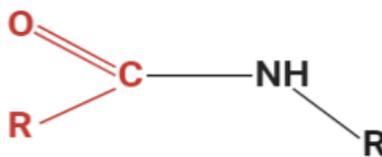


Figure 2 Substrate of Enzyme

It is imperative to study the structure, catalytic sites, and key residues of TMPRSS2 so that we can find suitable inhibitors to stop the splitting of Spike Glycoprotein inside the human body. Figure 2 illustrates the three site chains of Serine (SER), Histidine (HIS), and Aspartic Acid (ASP). It is essential to mention here that the key residues that are very significant in the catalytic process are at a distance from each other in a sequence, they brought closer to each other by the folding of the enzyme to make them physically close to each other. The closeness of these is essential to start, but more importantly, the flexibility of the protein with these side chains is vital to the catalytic function that is going to occur. Specifically, the green dotted region indicates the place where spike protein binds, and the cleavage process occurs after it interacts with this active catalytic triad of SER, HIS, and ASP[2]. The binding of a substrate (which is Spike protein in the case of Coronavirus) from Figure 1 will occur in the S1 pocket of an enzyme. S1 pocket is shown in a semicircle that is holding on the part of the protein. As we can see, that protein is going to be cut at the active site. It is vital to note here that after the protein binding with an enzyme, the Negative charge of oxygen in Aspartate is slightly moved closer to the ring of Histidine. As a result, the electronic configuration of the HIS ring is changed; results from the nitrogen of HIS are now attracting Hydrogen, which is in SER. So this binding of the enzyme is starting the process by which reaction is going to initiate. S1 pocket uses specific protein for binding[3]. It means that protein with a particular protein would bind to this pocket. Nitrogen from HIS abstracts a proton from SER side-chain, creating alkoxide ion, which makes the nucleophilic attack on the carbonyl carbon of peptide bond eventually releasing half of the polypeptide from enzyme and another half of polypeptide covalently links with SER. And this action will be repeated for the next cycle.

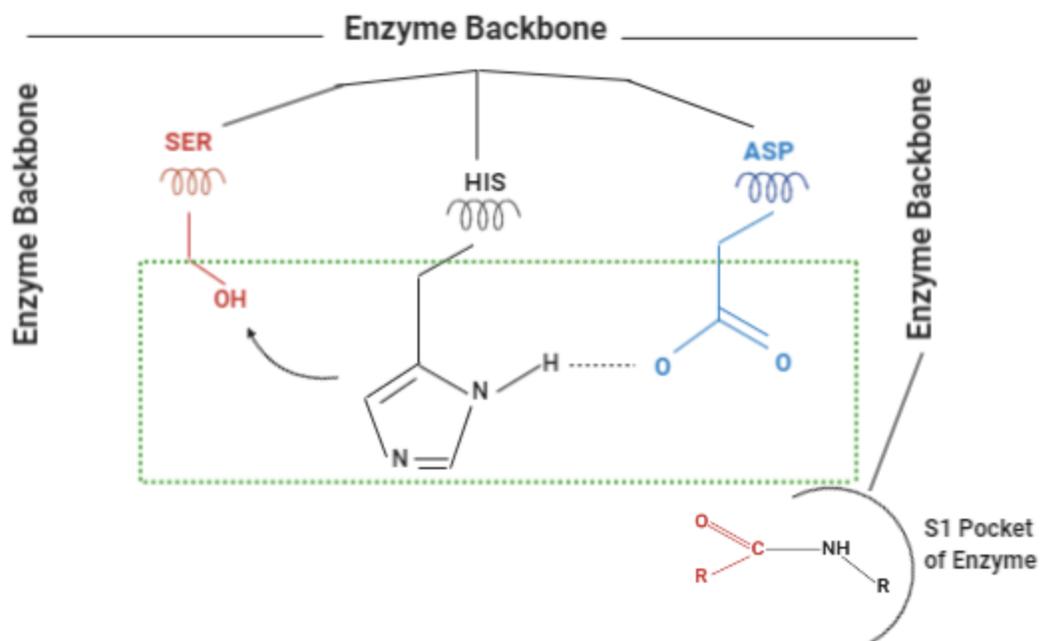


Figure 3 Catalysis Mechanism of TMPRSS2

From the above study, it is well established that the active binding sites of catalytic triad must be targeted to inhibit the process of TMPRSS2. The key residues that are catalytic in the cleavage process are HIS 296, ASP 345, and SER 441. Targeting these residues will inhibit the functioning of TMPRSS2 and hence will stop the splitting of Spike of Covid-19. One of the potential inhibitors of serine protease is Camostat Mesylate, which is currently under trial. This compound binds to key sites HIS 296 and SER 441 along with other key sites like GLY 464, GLY 439, and SER 460[4]. Binding details of Camostat Mesylate may help to propose promising compound that binds to active side residues along with another residue that is missed by Camostat Mesylate like ASP 345 and good binding affinity than Camostat Mesylate.

Phytochemicals and flavonoids consist of natural compounds and herbs that have been used since ages for treatment purposes due to their antiviral properties[5]. For example, Azadirachta, which is commonly known as neem, contains antiviral and antibacterial substances that are used for diabetes, skin diseases, and stomach upset, etc. In the same way, compounds of aconitum heterophyllum contain naturally occurring antibiotic and antiviral compounds that are used for the treatment of cough, diarrhea, and fever. Artemisia compounds are useful for the inhibition of tumor growth and can be used as anti-cancer substances. There are thousands of phytochemicals, flavonoids, flavones, and other naturally occurring substances that are readily available in databases like PubChem, ChemSpider, and ChEMBL. Proper clinical trials of these substances for a particular illness can devour enormous exertion as far as time and funds are concerned, and yet output results may not accomplish the desired goals. In-silico analysis] can extensively limit the number of significant substances through precise and careful demonstration. These strategies can give an insight into the structural behavior and anticipate the effectiveness against viruses.

Related Work:

Various computational methods such as docking, Virtual screening, Molecular Dynamics, binding analysis is used to analyze compounds for the discovery of potential inhibitors. 3D Structure of TMPRSS2 is developed by the SWISS Model. Ammar et al. performed the analysis to evaluate the potential of ligands interacting with human enzymes using Computer-Aided Design (CADD). To study the potential inhibitors, ADMET profiling is done by developing the structure of a protein using Homology Modeling. Then the broad set of phytochemicals is virtually screened by Autodock Vina[6]. T. Joshi performed the Virtual screening of 316 chemicals for the identification of potential drug for ACE2, analyzed the interaction with PyMol and Ligplot+ tool is used to analyze the different conformations of ligand on binding sites[7]. Claudia et al. used the Schrodinger server for the docking of protein and ligand. Grid is designed by a tool named Grid Generator. To further study the correlation of TMPRSS2 and ACE2 with Spike Protein, path enrichment analysis technique is followed[8]. Manoj followed the method of Molecular Dynamics for protein-ligand analysis; similar ligands are screened using the DrugMint tool. Pockets in the protein are identified by a web-based tool named CastP is used. Followed by these methods, Autodock is used for the interactions, and binding affinity of ligands is strictly observed[9]. NAMD tool is used to evaluate the stability of protein, which interacts with ligands, and solvation, minimization of protein in complex with the ligand is analyzed for RMSD and RMSF calculation[10]. R. Chen, in his study, has shown Naringen as a potential candidate as a therapeutic compound using Virtual screening, docking analysis, homology modeling, and protein-ligand interaction[11]. Many researchers have also applied machine learning methods for the study of large datasets[12]–[16].

Our study used the same virtual screening, Docking, Molecular Dynamics, and structure analysis method for the identification of TMPRSS2 inhibitor. Less useful and irrelevant compounds are discarded, and chosen compounds are further observed by docking and binding analysis, and most effective and suitable compounds are selected for further investigation.

Material and Methods:

Preparation of TMPRSS2 and Phytochemicals:

The structure of TMPRSS2 that is used in this study is obtained from UniProt (PDB ID: O15393). The 3D model is generated from SWISS-Model[17]. Molecules have two chains with a length of 492 amino acids. MGL tool is used to remove hydrogen atoms and water molecules from TMPRSS2

Phytochemicals are selected by investigating the properties of various plants. At that point, the 3D compound structure of 4217 phytochemicals is selected from chemical and organic databases like PubChem[18], IMMPAT[19], and ChemSpider[20]. After downloading these structures, these Structure Data file (SDF) format compounds are converted into protein data bank (PDB) format. In this way, the phytochemical library is set up to perform further operations.

Virtual Screening of Compounds:

RPBS is an online server that is used for virtual screening of phytochemicals. These compounds are uploaded on the server for screening. This server has built-in package of AutoDock Vina, which it uses for the screening of suitable compounds[21]. Table 1 shows the Grid Coordinates and search space that is used

in the screening of compounds. The server performed virtual screening of listed structures and discarded all those compounds which have a higher binding affinity than -7.

Table 1 Grid Coordinates and Spacing for Virtual Screening

Grid Coordinates	X=9.8, Y=-4.2, Z=17.6
Grid Spacing	X=45, Y=45, Z=45

Molecular Docking:

Molecular docking is the process of docking ligand in the potential pocket of the protein. The screened compounds are further listed for the docking process[22]. PyRx software is used for docking of screened compounds and Tmprss2. For the minimized energy of selected compounds, PyRx used universal force field analysis. Table 2 illustrates the Grid Coordinates and Grid Dimensions.

Table 2 Grid Coordinates and Spacing for Autodock Vina

Grid Coordinates	X=10, Y=-5, Z=18
Grid Dimensions (Angstrom)	X=35, Y=35, Z=35

BFGS Algorithm:

For optimization and solving the search problems of the conformer, the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm is deployed in Autodock Vina. This algorithm generates different poses of each compound. One of the optimization strategies is the Gradients scoring function, which uses the method of derivation with defined arguments. In the initial state, it follows metropolis criteria to decide on whether to continue with the original pose or update the new one. Regardless of the second-order derivative, that can increase complexity, BFGS uses gradient evaluation for approximating the rank one improvement. Ultimately structure is evaluated by metropolis criteria for selection, and selected structure is used as an initial point in the next cycle.

Visualization:

Docked compounds are analyzed by PyMol, which is an efficient and widely used package that provides a complete simulation of the 2D and 3D structure of the protein, ligand, or interaction of both. Binding sites of protein-ligand can be analyzed with the distance measurement of binding interactions.

Molecular Dynamic Analysis:

Molecular dynamic (MD) is an in silico method that is used to examine the structure and interaction of protein-ligand complex. For this purpose, Nanoscale Atomic Dynamic (NAMD) is a tool that analyzes the structural stability of Tmprss2 and ligand[23]. The temperature was set to 310K, and CHARMM site is used to create configuration files. Parameter files protein and ligand were generated from CHARMM General Power Field (CGenFF)[24]. A combination of protein-ligand is solvated using water atoms. Table 3 illustrates the details of temperature, stages, and dielectric, which is set in this process.

Table 3 shows the values set for the simulation

Temperature	310K
Stages	2000
Dielectric	1.0

Electrostatic Potential Calculation:

PyMol, which is implemented on the Poisson-Boltzmann method, is used to perform Electrostatic potential charge calculation and analyze the charge distribution on the protein surface using a technique of cubic spline charge discretization. It divides the positive (blue) and negative (red) charge all over the protein. Table 3 shows numerical details of the Grid box that is used to calculate the electrostatic charge on TMRSS2.

Table 4 Shows Grid box Details set for the electrostatic charge calculation

Temperature	310K
Solute Dielectric	2.0
Solvent Dielectric	78.0
Radius Range	1.400 Å

Results:

Table 5 shows the details of selected compounds after molecular docking. Binding affinity and binding residues are also listed in this table. The threshold of binding affinity was set less than -7. These Potential compounds bind to key sites of TMRSS2. Most of these compounds are found in plants and herbs and have shown antiviral properties to be considered for the inhibitors of the targeted protein.

Table 5 shows the list of Potential compounds that show Good binding interactions with TMRSS2

Ligand	Binding Sites	Binding Affinity (kcal/mol)
6-Deacetylrimbinene	ASP 359	-13.4
	SER 441	
2',3'-Dehydrosalannol	GLY 464	-12.9
	HIS 296	
Deacetylsalannin	SER441	-11.9
	GLY 464	
	SER 460	
Salannin	GLU 376	-11.2
	GLN 374	
	HIS 296	

Salannol acetate	SER 441	-10.7
	ASP 340	
Nimbolin A	GLN 374	-10.6
	HIS 296	
17-epi-17-Hydroxyzadiradione	HIS 296	-10
	ASP 340	
	SER 436	
17-Hydroxyzadiradione	GLU 376	-9.7
	GLN 377	
	SER 441	
Nobiletin	GLY 282	-9.5
	HIS 296	
5,6,7,8,3',4',5'- Heptamethoxyflavone	HIS 279	-9.5
	HIS 296	
Pinostrobin	SER 441	-9.3
	GLY 443	
Sakuranetin	SER 436	-9.3
	HIS 296	
Homoeriodictyol	GLY 282	-9.2
	SER 436	
	SER 441	
Umuhengerin	GLN 374	-9.1
	HIS 296	
Eucalyptin	SER 460	-9.1
	SER 441	
4'-Methoxyflavone	GLY 443	-9.0
	SER 436	
	HIS 296	
4',5,7-Trimethoxyflavone	GLU 376	-9.0
	SER 441	
3-Hydroxy-4'-methoxyflavone	HIS 296	-9.0
Nobiletin	ASP 340	-8.9
Pinostrobin	SER 460	-8.9
	SER 441	
Lutonarin	HIS 296	-8.8
	ASP 345	
	SER 441	
	HIS 279	
	GLY 474	
Sakuranetin	SER 460	-8.5
	SER 441	
Homoeriodictyol	SER 436	-8.4
	HIS 296	
Umuhengerin	GLN 374	-8.4
	HIS 296	

Eucalyptin	SER 441	-7.9
	GLY 443	
	SER 436	
4'-Methoxyflavone	ASP 359	-7.9
	SER 441	
4',5,7-Trimethoxyflavone	HIS 296	-7.8
	HIS 279	
3-Hydroxy-4'-methoxyflavone	ASP 345	-7.8
	GLY 443	
	SER 436	
3-Hydroxy-3',4',5,7-tetramethoxyflavone	HIS 296	-7.8
	GLU 376	
3-Hydroxy-3',4'-dimethoxyflavone	HIS 279	-7.8
	HIS 296	
5,6,7,8,3',4',5'-Heptamethoxyflavone	SER 441	-7.8
	SER 441	
	GLY 443	
Pinostrobin	SER 441	-7.8
	HIS 279	
	GLY 474	
Sakuranetin	SER 460	-7.8
	SER 441	
Homoeriodictyol	SER 441	-7.7
	ASP 340	
Umuhengerin	GLN 377	-7.7
	SER 441	
	GLY 474	
Glabone	SER 460	-7.6
	SER 441	
5,6,5'-Trihydroxy-3,7,2',4'-tetramethoxyflavone	HIS 296	-7.6
	GLN 374	
	HIS 279	
4H-1-Benzopyran-4-one, 2-(1,3-benzodioxol-5-yl)-5,7-dimethoxy-	HIS 296	-7.5
	SER 441	
	GLY 464	
Isothymonin	GLN 374	7.5
	HIS 279	
3-Acetyltricin	GLY 474	-7.5
	SER 460	
3,5-Dihydroxy-4',7'-dimethoxyflavanone	GLY 464	-7.4
	GLN 374	
2-(3-Hydroxy-4-methoxyphenyl)-3,7-dimethoxy-5,6-dimethylchromen-4-one	HIS 296	-7.4
	SER 441	
	GLN 374	
	GLY 464	-7.4

8-Butyl-5,7-dimethoxy-2-(3,4,5-trimethoxyphenyl)-4h-chromen-4-one	ASP 345	
Camostat Mesylate	HIS 296	-7.4
	SER 441	
	GLY 464	
	HIS 279	
7,3',4',5'- Tetramethoxyflavanone	GLY 464	-7.3
	HIS 279	
Aprotinin	HIS 296	-7.2
	SER 441	
	GLY 464	
	SER 460	
3-Methyl-4'-methoxy flavone	HIS 296	-7.2
	ASP 345	
	GLY 464	
2-(3-Hydroxy-4-oxo-2,5-cyclohexadienylidene)-5,7-dihydroxy-2H-1-benzopyran-3,4-dione	GLN 374	-7.2
	HIS 279	
	GLY 464	
8-Chloro-5,7-dimethoxy-2-(4-methoxyphenyl)-2,3-dihydrochromen-4-one	HIS 296	-7.1
	ASP 345	
	GLY 464	

From Table 5, only those compounds are selected for further analysis that shows good binding affinity, interaction with key residues, and are naturally occurring substances. Some off the compounds listed in the above table is also under trial. In the next step, we tried to select only those compounds which have an excellent binding affinity and bind to maximum key residues, and then we compared the proposed compound with the proven antiviral compounds for the TMPRSS2.

In Table 6, we shortlisted the best compounds from Table 5 and highlighted the key sites and distance of residues binding with the ligand, inhibition constant, and comparison with Camostat Mesylate, and Aprotinin is also shown.

Table 6 Shows Comparison of the proposed compound with proven Antiviral Drugs

Ligand	Binding Sites with TMPRSS2	Residue Distance	Binding Affinity (kcal/mol)	Inhibition Constant (nM)
Lutonarin (Proposed Compound)	HIS 296	2.3	-8.8	348nM
	ASP 345	2.7		
	SER 441	2.3		
	ARG 470	2.6		
	GLY 474	2.8		

	TYR	2.7		
Camostat Mesylate (Proven Antiviral Drug)	HIS 296	2.4	-7.4	3713nM
	SER 441	2.5		
	GLY 464	2.2		
	HIS 279	2.1		
Aprotinin (Proven Antiviral Drug)	HIS 296	2.7	-7.3	4396nM
	SER 441	2.9		
	GLY 464	2.3		
	SER 460	2.5		

Below details of proposed compounds have been described with uses, structural properties, toxicity, and visualization of binding interaction to verify its potency with TMPRSS2.

Lutonarin/Isooreintin 7-glucoside (PubChem ID: 44559810):

Lutonarin is a natural compound that is found in barley seeding, which is also a source of food. It belongs to the class of organic compounds containing flavonoids and showed antiviral and antioxidant activities. It is used for the treatment of cough, metabolic syndrome, and diabetes with very low toxicity and minimal side effects. The structure contains O-Glycosidically combined with carbohydrate moiety at the C7 position. Figure 4 illustrates the structure of Lutonarin.

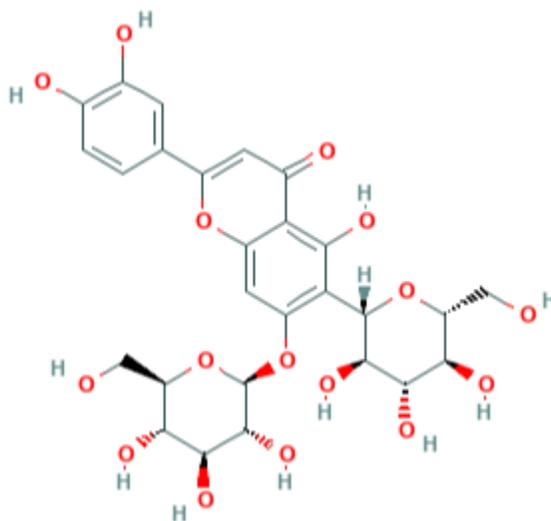


Figure 4 Chemical Structure of Lutonarin[18]

Lutonarin can be used to inhibit the working of TMPRSS2, which is used to split the Spike glycoproteins of Coronavirus. Now we shall see the interaction of the proposed compound with TMPRSS2.

Lutonarin:

The interaction of Lutonarin-TMRSS2 is shown in figure 5. The proposed compound not only binds to all three key residues but other critical residues such as HIS 279, GLY 464 also.

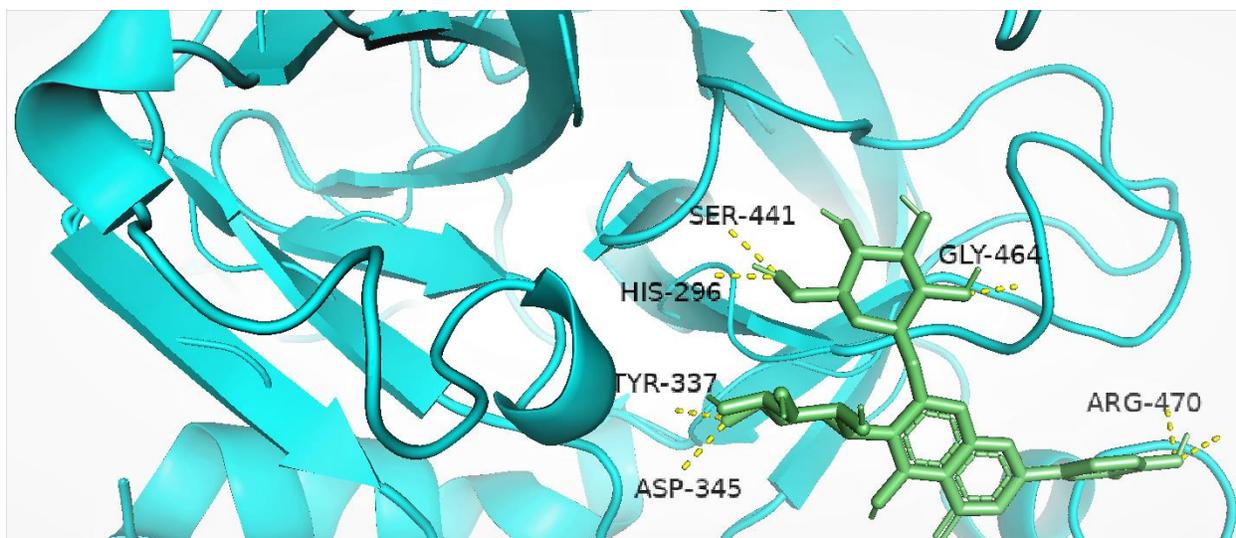


Figure 5 Binding Interaction of Lutonarin-TMPRSS2

Table 7 illustrates the interaction of Lutonarin with TMPRSS2 that binds to all three key sites with an affinity of -8.8 and an inhibition constant of 348nM.

Table 7 shows the details of Binding Residues, Binding Affinity, and Inhibition Constant

Ligand	Color	No. of Key Binding Sites	Binding Residues	Binding Affinity (kcal/mol)	Inhibition Constant (nM)
Lutonarin	Green	3	HIS 296, ASP 345, SER 441, GLY 464, ARG 470, TYR 337	-8.8	348nM

Electrostatic Potential Calculation:

Electrostatic potential is used to calculate the potential charges distribution on the surface of the protein. We used PyMol to map the charges on the target protein. Figure 6 shows the distribution of charges on the surface of TMPRSS2 during the interaction of Lutonarin.

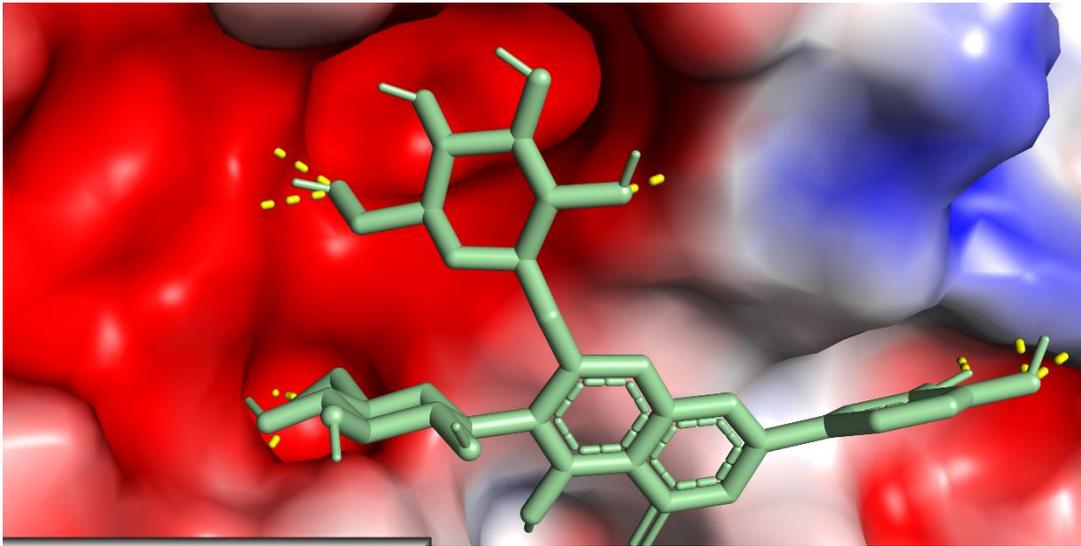


Figure 6 Shows the Inner Cavity of TMPRSS2 during interaction with Lutonarin

From figure 6, Blue and Red color indicate positive and negative charges respectively, and we can visualize that inner cavity of the binding pocket of TMPRSS2 is covered with a negative charge (Red), which indicates that this internal cavity is where key residues are, and cause cleavage of Spike. Binding this cavity will inhibit the function of TMPRSS2; thus, our proposed ligand (Lutonarin) targets this cavity and binds to all key residues, which make it more suitable.

Molecular Dynamic Simulation (MDS) Analysis:

MDS is used to analyze the stability of the combination of both protein and ligand. Root Mean Square Deviation (RMSD) is used to calculate the interval between the backbone molecules of the superimposed atom. Figure 7 illustrates the RMSD analysis of free TMPRSS2 and the complex of TMPRSS2 and Lutonarin. RMSD of free TMPRSS2 arose until 14ns at 1.6 Å, where it stabilized till 25ns at 1.5 Å, again after slight fluctuations, it became consistent at 43ns at 1.52 Å. For TMPRSS2_Lutonarin, RMSD ascended until 13ns at 1.66 Å and stabilized until 23ns at 1.5 Å.

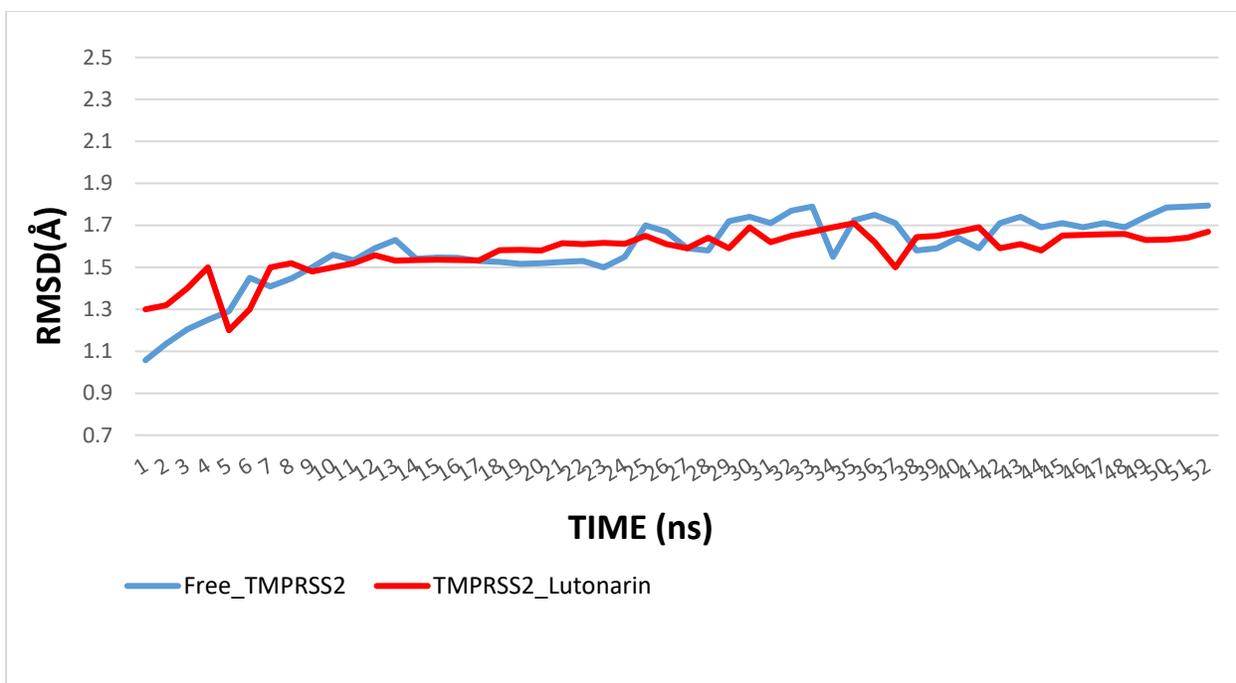


Figure 7 RMSD illustration of Free_TMPRSS2 and complex of TMPRSS2 with Lutonarin

Figure 8 shows a brief analysis of Root Mean fluctuation of residues of TMPRSS2 and its complex. As it can be observed, free TMPRSS2 and its complex showed variation from 1.22 Å to 1.78 Å, which is the clear indication that TMPRSS2 aligned with the proposed compound and sustained close contact with the proposed inhibitor.

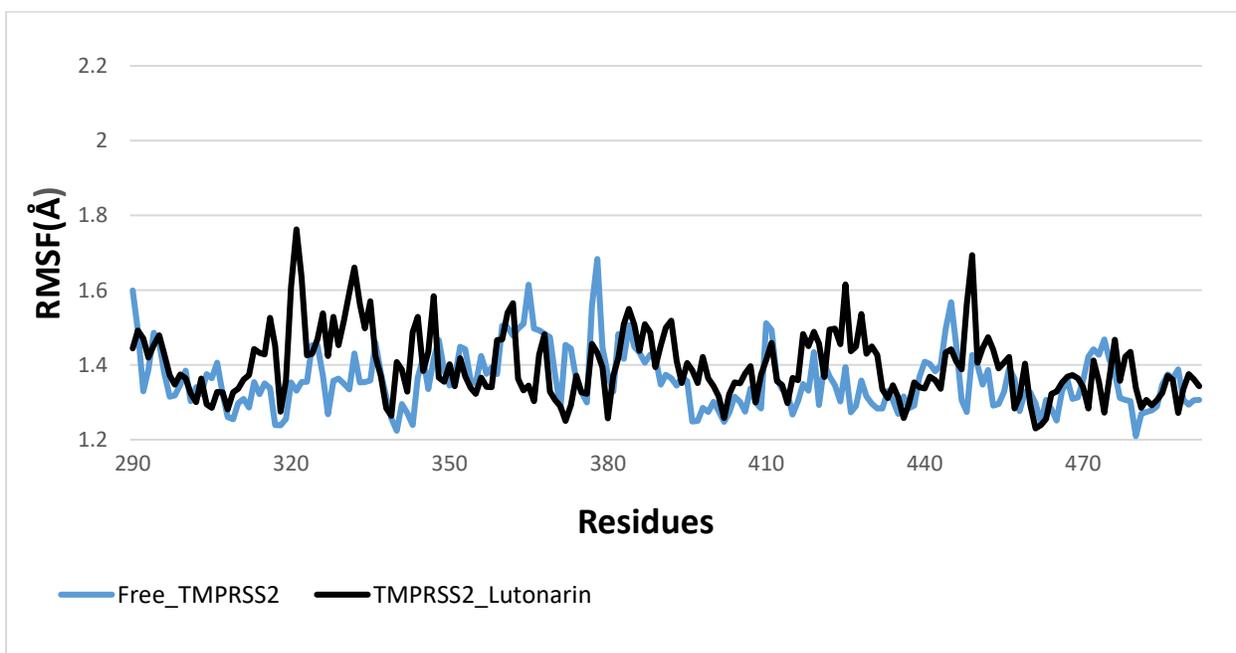


Figure 8 RMSF illustration of Free_TMPRSS2 and complex of TMPRSS2 with Lutonarin

Conclusion:

Covid-19 affected millions of people all over the world and is still on the rise. It is inevitable to design a vaccine that can help to cure the viral infection. Researchers are working hard to understand the structure and devise a way to come out of this plague. Phytochemicals are naturally occurring substances with fewer side effects and natural antiviral properties. In this paper, we have proposed the potential inhibitor to stop the replication of the virus inside the human body by targeting type II Transmembrane Protease Serine, which contains catalytic side chain that causes the splitting of Spike protein of Coronavirus. Targeting key residues of this catalytic chain would inhibit the functioning of the target protein, and as a result, virus replication can be avoided. We used the in-silico method and machine learning approaches to propose a compound. The proposed chemical compound is analyzed based on the binding affinity, key binding residues, Molecular Dynamic Simulation, RMSD, RMSF, and Electrostatic potential Distribution. After qualifying all the mentioned approaches, we proposed a natural compound that has a good binding affinity and binds to all key residues of the Transmembrane Protease Serine, which gives a strong indication that the proposed compound can act as an inhibitor of TMPRSS2 and mitigate the ferocity of the infection.

Future Work:

All results of this study have been obtained from computational methods. Proper medicinal research and thorough analysis will reveal important features of the compound. If this compound is proven stable under medical tests and yield good results, then this would open new ways of treating this virus.

References

- [1] J. Jankun, "COVID-19 pandemic; transmembrane protease serine 2 (TMPRSS2) inhibitors as potential drugs.," *Transl. Univ. Toledo J. Med. Sci.*, vol. 7, pp. 1–5, 2020.
- [2] D. Huggins, "Structural Analysis of Experimental Drugs Binding to the COVID-19 Target TMPRSS2," 2020.
- [3] T. Hirano and M. Murakami, "COVID-19: A new virus, but a familiar receptor and cytokine release syndrome," *Immunity*, 2020.
- [4] Y.-J. Dai, W.-N. Zhang, W.-D. Wang, S.-Y. He, C.-C. Liang, and D.-W. Wang, "Two potential novel SARS-CoV-2 entries, TMPRSS2 and IFITM3, in healthy individuals and cancer patients," *Heal. Individ. Cancer Patients*, 2020.
- [5] Y. D. Khan and M. S. Roomi, "Promising Compounds for Treatment of Covid-19," *VAWKUM Trans. Comput. Sci.*, vol. 17, no. 1, pp. 1–8, 2020.
- [6] A. D. Elmezayen, A. Al-Obaidi, A. T. Şahin, and K. Yelekçi, "Drug repurposing for coronavirus (COVID-19): in silico screening of known drugs against coronavirus 3CL hydrolase and protease enzymes," *J. Biomol. Struct. Dyn.*, no. just-accepted, pp. 1–12, 2020.

- [7] T. Joshi *et al.*, "In silico screening of natural compounds against COVID-19 by targeting Mpro and ACE2 using molecular docking," *Eur. Rev. Med. Pharmacol. Sci.*, vol. 24, pp. 4529–4536, 2020.
- [8] C. Cava, G. Bertoli, and I. Castiglioni, "In Silico Discovery of Candidate Drugs against Covid-19," *Viruses*, vol. 12, no. 4, p. 404, 2020.
- [9] M. K. Gupta, S. Vemula, R. Donde, G. Gouda, L. Behera, and R. Vadde, "In-silico approaches to detect inhibitors of the human severe acute respiratory syndrome coronavirus envelope protein ion channel," *J. Biomol. Struct. Dyn.*, no. just-accepted, pp. 1–17, 2020.
- [10] R. C. Bernardi, L. F. Milles, and H. E. Gaub, "NAMD as a Tool for In Silico Force Spectroscopy," *Biophys. J.*, vol. 118, no. 3, p. 144a, 2020.
- [11] R. Chen, Q.-L. Qi, M.-T. Wang, and Q.-Y. Li, "Therapeutic potential of naringin: an overview," *Pharm. Biol.*, vol. 54, no. 12, pp. 3203–3210, 2016.
- [12] A. W. Ghauri, Y. D. Khan, N. Rasool, S. A. Khan, and K.-C. Chou, "pNitro-Tyr-PseAAC: predict nitrotyrosine sites in proteins by incorporating five features into Chou's general PseAAC," *Curr. Pharm. Des.*, vol. 24, no. 34, pp. 4034–4043, 2018.
- [13] S. A. Khan, Y. D. Khan, S. Ahmad, and K. H. Allehaibi, "N-MyristoylG-PseAAC: sequence-based prediction of N-myristoyl glycine sites in proteins by integration of PseAAC and statistical moments," *Lett. Org. Chem.*, vol. 16, no. 3, pp. 226–234, 2019.
- [14] Y. D. Khan, N. Rasool, W. Hussain, S. A. Khan, and K.-C. Chou, "iPhosT-PseAAC: Identify phosphothreonine sites by incorporating sequence statistical moments into PseAAC," *Anal. Biochem.*, vol. 550, pp. 109–116, 2018.
- [15] W. Hussain, Y. D. Khan, N. Rasool, S. A. Khan, and K.-C. Chou, "SPalmitoylC-PseAAC: A sequence-based model developed via Chou's 5-steps rule and general PseAAC for identifying S-palmitoylation sites in proteins," *Anal. Biochem.*, vol. 568, pp. 14–23, 2019.
- [16] Y. D. Khan, F. Ahmed, and S. A. Khan, "Situation recognition using image moments and recurrent neural networks," *Neural Comput. Appl.*, vol. 24, no. 7–8, pp. 1519–1529, 2014.
- [17] T. Schwede, J. Kopp, N. Guex, and M. C. Peitsch, "SWISS-MODEL: an automated protein homology-modeling server," *Nucleic Acids Res.*, vol. 31, no. 13, pp. 3381–3385, 2003.
- [18] S. Kim *et al.*, "PubChem substance and compound databases," *Nucleic Acids Res.*, vol. 44, no. D1, pp. D1202–D1213, 2016.
- [19] K. Mohanraj *et al.*, "IMPPAT: A curated database of Indian Medicinal Plants, Phytochemistry and Therapeutics," *Sci. Rep.*, vol. 8, no. 1, pp. 1–17, 2018.
- [20] H. E. Pence and A. Williams, "ChemSpider: an online chemical information resource." ACS Publications, 2010.
- [21] C. M. Labbé *et al.*, "MTiOpenScreen: a web server for structure-based virtual screening," *Nucleic Acids Res.*, vol. 43, no. W1, pp. W448–W454, 2015.
- [22] O. Trott and A. J. Olson, "AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading," *J. Comput. Chem.*, vol. 31, no. 2, pp. 455–461, 2010.

- [23] M. Aljofan and A. Gaipov, "COVID-19 Treatment: The Race Against Time. *Electron J Gen Med.* 2020; 17 (6): em227." 2020.
- [24] K. Vanommeslaeghe *et al.*, "CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields," *J. Comput. Chem.*, vol. 31, no. 4, pp. 671–690, 2010.
- [25] M. Awais, W. Hussain, Y. D. Khan, N. Rasool, S. A. Khan, and K.-C. Chou, "iPhosH-PseAAC: Identify phosphohistidine sites in proteins by blending statistical moments and position relative features according to the Chou's 5-step rule and general pseudo amino acid composition," *IEEE/ACM Trans. Comput. Biol. Bioinforma.*, 2019.
- [26] W. Hussain, Y. D. Khan, N. Rasool, S. A. Khan, and K.-C. Chou, "SPrenylC-PseAAC: A sequence-based model developed via Chou's 5-steps rule and general PseAAC for identifying S-prenylation sites in proteins," *J. Theor. Biol.*, vol. 468, pp. 1–11, 2019.
- [27] O. Barukab, Y. D. Khan, S. A. Khan, and K.-C. Chou, "iSulfoTyr-PseAAC: Identify Tyrosine Sulfation Sites by Incorporating Statistical Moments via Chou's 5-steps Rule and Pseudo Components," *Curr. Genomics*, vol. 20, no. 4, pp. 306–320, 2019.
- [28] Y. D. Khan, A. Batool, N. Rasool, S. A. Khan, and K.-C. Chou, "Prediction of nitrosocysteine sites using position and composition variant features," *Lett. Org. Chem.*, vol. 16, no. 4, pp. 283–293, 2019.