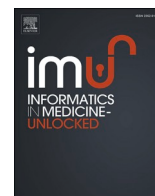




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Identification of novel TMPRSS2 inhibitors for COVID-19 using e-pharmacophore modelling, molecular docking, molecular dynamics and quantum mechanics studies

Abdulrahim A. Alzain, PhD^{1,*}, Fatima A. Elbadwi¹

Department of Pharmaceutical Chemistry, Faculty of Pharmacy, University of Gezira, Gezira, Sudan

ARTICLE INFO

Keywords:

SARS-CoV-2
TMPRSS2
Homology modeling
e-pharmacophore mapping and screening
Docking
Molecular dynamics

ABSTRACT

SARS coronavirus 2 (SARS-CoV-2) has spread rapidly around the world and continues to have a massive global health effect, contributing to an infectious respiratory illness called coronavirus infection-19 (COVID-19). TMPRSS2 is an emerging molecular target that plays a role in the early stages of SARS-CoV-2 infection; hence, inhibiting its activity might be a target for COVID-19. This study aims to use many computational approaches to provide compounds that could be optimized into clinical candidates. As there is no experimentally derived protein information, initially we develop the TMPRSS2 model. Then, we generate a pharmacophore model from TMPRSS2 active site consequently, and the developed models were employed for the screening of one million molecules from the Enamine database. The created model was then screened using e-pharmacophore-based screening, molecular docking, free energy estimation and molecular dynamic simulation. Also, ADMET prediction and density functional theory calculations were performed. Three potential molecules (Z126202570, Z46489368, and Z422255982) exhibited promising stable binding interactions with the target. In conclusion, these findings empower further *in vitro* and clinical assessment for these compounds as novel anti-COVID19 agents.

1. Introduction

The current COVID-19 coronavirus disease pandemic is a worldwide health concern associated with SARS-CoV-2 illness [1,2]. As of August 7th, 2021, the pandemic has grown to over 200 million illnesses and over 4 million deaths [3]. This epidemic has impacted negatively on national healthcare systems and rained down the global economy, necessitating the development of efficient therapies to halt its spread and control its severity [4].

Coronaviruses can affect both animals and humans, producing a wide range of illnesses [5]. MERS coronavirus, SARS coronavirus and the 2019 coronavirus that originated from Wuhan in Dec 2019 are the three contagious viral human coronaviruses (hCoVs) that are being identified to date [6,7]. However, from WHO statistics, the incidence of SAR-CoV-2 transmission in people has far outpaced that of SARS-coronavirus and MERS- coronavirus [8].

Transmembrane spike glycoproteins (S proteins) extending from the viral surface to give the coronavirus its crown-like ("corona")

appearance [9]. The alpha, beta, gamma, and delta are by far the four classes of coronaviruses [10]. Bats are thought to be the source of SARS-CoV-2, a beta coronavirus that can be transmitted without causing illness. The infectious agent of is a virus that has 4 glycoproteins: spike, envelope, membrane, and nucleocapsid [11,12].

The virus's interaction with the host influences the incidence and progression of COVID-19 illness [13]. For viral proliferation, all corona viruses have unique genes that encode protein in downstream ORF1 regions. Spike glycoprotein controls the virus adhesion and entry [14]. The virus's transmembrane spike glycoprotein attaches to the ACE2 protein to accelerate virus penetration into the host cell [15–17].

Various viruses, including SARS coronavirus, Ebola virus, influenza virus and MERS coronavirus, utilize host cell proteases to activate their envelope glycoproteins, according to prior [18–20]. Transmembrane protease/serine subfamily member 2 cleaves and activates the SARS-CoV-2 spike protein, which is critical for membrane fusion and host cell penetration [21–23]. It's worth mentioning that when the corona spike protein binds to the ACE2 receptor, the host cell's

* Corresponding author. 21111 Barakat Street, Medani, Gezira, Sudan.

E-mail addresses: abdulrahim.altoam@uofg.edu.sd, awh3134@hotmail.com (A.A. Alzain).

¹ Authors contributed equally.

<https://doi.org/10.1016/j.imu.2021.100758>

Received 27 August 2021; Received in revised form 22 September 2021; Accepted 11 October 2021

Available online 14 October 2021

2352-9148/© 2021 The Author(s).

Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

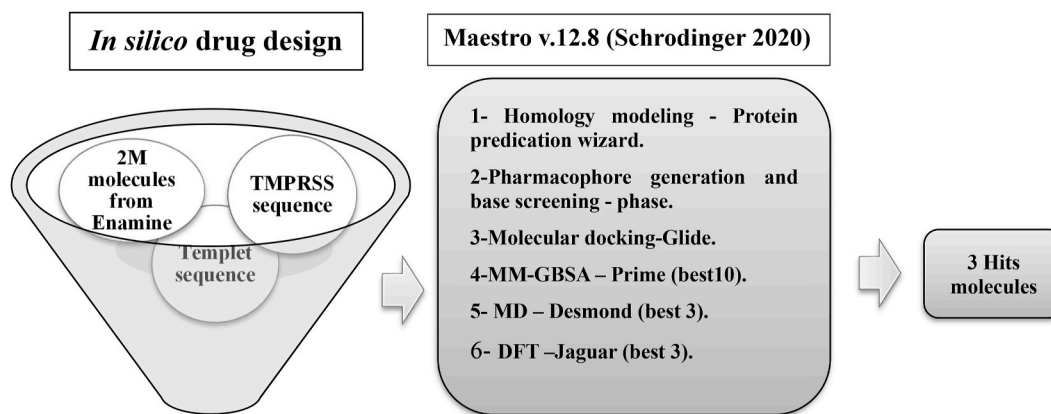


Fig. 1. Study workflow.

surrounding serine protease TMPRSS-2 receptor facilitates in the virion's entry via priming the S-2 domain of the spike protein [10,22, 24–26]. Since TMPRSS2 mRNA is abundant in a variety of tissues including the prostate, kidney, small intestine, lung, and stomach and others, it is linked to the progression and severity of the illness [1,18,27, 28]. As a result, TMPRSS2 processing is amongst the most important stages in activating the SARS-CoV-2 S protein's membrane activity [29]. Because of the extreme role of TMPRSS2, it might be a promising therapeutic target in the struggle against COVID-19 infection [30,31].

Intensive attempts by several groups globally are in full turn to find novel agents, but this may take a considerable time since current drugs have to progress through long preclinical and clinical safety testing [32]. Structure-based drug discovery is one of the most and newest used techniques to investigating a medication's action [33]. When experimental approaches fail or there is insufficient data about the target, homology modeling is a successful *in silico* strategy for developing a tertiary structure of the target protein to estimate structural information for docking studies [34]. The number of compounds that a medicinal chemist can synthesize can be substantially decreased by using pharmacophore model creation. As a result, drug research and development time and cost can be decreased [35]. It offers the estimation of activity in a wide range of structurally varied compounds and could also assist in the discovery of novel molecules with enhanced activity [36,37]. The pharmacophore model shows the key chemical characteristics that are responsible for activity in a logical manner. Pharmacophore modeling has been one of the most fundamental and promising approaches for novel drug development in recent years [38].

In present study, we aim to find novel TMPRSS2 inhibitors through using *in silico* computational approaches namely homology modeling to construct TMPRSS2 protein model, pharmacophore mapping and screening, docking to calculate their binding affinities and their free binding energies. Molecular dynamics and density functional theory calculations were also performed to validate the results.

2. Method

All the *in silico* studies were carried out by Maestro v 12.8 (Schrodinger 2020) as depicted in Fig. 1.

2.1. Homology modeling search and model preparation

TMPRSS2 sequence was downloaded from UniProt (accession number: O15393) [39]. A similarity search was done using BLAST server [40] in order to find the best template for modeling of TMPRSS2. Human plasma kallikrein was the best template for homology modeling based on sequence identity of 42.21% and sequence similarity of 47%. Accordingly, human plasma kallikrein 3D structure (PDB ID: 5TJX) [41] was obtained from protein databank and used as a template in the Prime

module of Schrodinger suite [42]. Alignment was performed using single template alignment (STA) and knowledge-based methods of the protein predication wizard were used to build a model using Prime of Schrodinger. Then, using the protein preparation wizard tool [43], the protein structure model was preprocessed to correct the order of all bonds in the structure and through Epik, proper ligands ionization and tautomeric state were confirmed in the specified pH range. Hydrogen atoms were added to the structure; bonds were broken to metals and correct the formal charge on the metal and the neighboring atoms to treat the bonds as ionic, which is required for use with force fields, which treat metal compounds as ionic rather than covalent. Then add zero-order bonds between the metal and its ligands, so that it is still considered part of the same molecule. It also finds sulfur atoms that are within 3.2 Å of each other and adds bonds between them. Also, the model was refined by optimizing hydrogen bond order using PROPKA in PH = 7 and restrained minimization was performed in OPLS3e force field to decrease side-chain energies.

The Ramachandran plot obtained by Prime module of Schrodinger was used for further verification in order to estimate model validity for the optimized homology model [44].

2.2. Active site prediction

The SiteMap tool was used to predict receptor sites for TMPRSS2 protein. This software uses grid points on the surface of the protein that could be suitable for ligands to bind to the receptor to provide data regarding binding site qualities. The final part of the process is the evaluation, which includes rating each site by computing different characteristics: two characteristics that provide distinct measurements of the site's availability to the solvent are exposure and enclosure, Donor/acceptor character, H donor/acceptor areas; contact, SiteScore and Ability of the site to absorb and bind to small molecules (druggability).

2.3. Ligand library preparation and grid generation

The chemical structures of million compounds from Enamine database [45] were downloaded and prepared using LigPrep [43]. LigPrep assures that compound structure is optimized, that structurally undesirable ligands are excluded, and that a low-energy 3-dimensional structure with optimal chiralities is developed for the input ligand. Briefly, the ionization states of ligands were defined at pH 7.0 ± 2.0 , and the salt was removed using Epik. For each ligand, a maximum of 32 conformations was allowed to be generated. The energies of each ligand were minimized by OPLS3e. The Grid Generation panel was used to create the grid. The grid was constructed at the active site's centroid point for consistency.

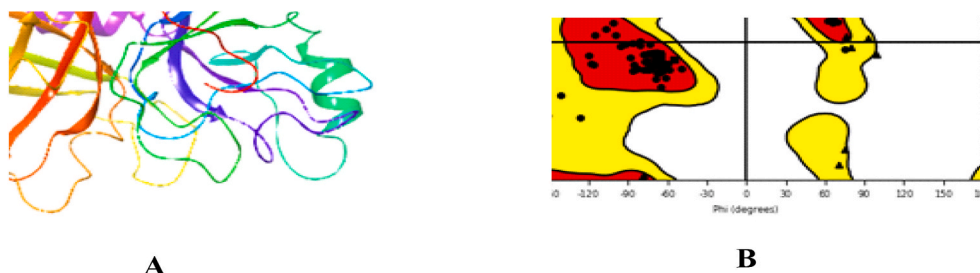


Fig. 2. A. Protein structures, B. Ramachandran plot.

2.4. E-pharmacophore generation

Using the option Generate pharmacophore from Receptor Cavity in the Phase module of Schrodinger [46], an energy-optimized pharmacophore hypothesis was developed using receptor centroid of the amino acids of the active site predicted by SiteMap. Before generating the e-pharmacophore, Phase first recruits the Glide XP tool to dock the already bound ligand inside the binding pocket of the protein under study and selects the top ranked pose for auto-pharmacophore generation, taking into account the following six chemical features: aromatic ring (R), hydrogen bond acceptor (A), hydrophobic group (H), hydrogen bond donor (D), negative ionic feature (N) and positive ionic feature (P).

2.5. E-pharmacophore based database screening

For the e-pharmacophore screening against TMPRSS2, the prepared conformations of the Enamine database were utilized. The compounds had to match a minimum of five features on the created e-pharmacophore model in order to get TMPRSS2 inhibitors with the appropriate chemical characteristics.

2.6. Docking of the generated compounds and MM-GBSA

The molecular docking makes it possible to predict the strength and affinity of the generated compounds for the target and also to determine the compound's binding site residues in the target protein and their interactions. The resulted conformations were docked into TMPRSS2 active site via HTVS mode of glide [42], and then the top-ranked poses were subjected to SP mode of glide and further to XP mode. The results were compared against approved TMPRSS2 inhibitors ambroxol acefyllinate and nafamostat mesylate.

By using the default settings of the Prime-MMGBSA function [42], MM-GBSA estimations have also been done for the top-ranked XP docked poses and compared with known TMPRSS2 inhibitors ambroxol acefyllinate and nafamostat mesylate.

2.7. Molecular dynamics

The stability of binding of the top compounds in complex with TMPRSS2 was studied using by molecular dynamics (MD). It helps assess patterns of interaction between proteins and ligands and time-dependent protein-ligand interaction associations and conformational dynamics of complex systems. Using Academic Desmond v6.5 by D.E. Shawn Research [47], the protein-ligand complexes were minimized

and relaxed under NPT using Desmond's default method. Simulation time was 50 nanosecond, the temperature (300 K) and pressure (1.01325 bar) were kept constant throughout the process by Nose-Hoover thermostat and Martyna-Tobias-Klein barostat. The OPLS3e force field was used in MD simulations. 5208 frames for each system during the simulations were collected. Lastly, the interaction analysis was performed.

2.8. Density functional theory

Using Jaguar [47] the top molecules were subjected to density functional theory calculations using optimization panel. After structures were included in the workspace, the theory was adjusted to be DFT which allows employing a variety of functional to describe exchange and correlation for either open or closed-shell systems. Utilizing a basis set of 6-31G** level, complete geometry optimization was done by using functional B3LYP. The geometries were optimized 100 times. The accuracy for SCF calculations is quick, which Use mixed pseudo-spectral grids with loose cutoffs (iacc = 3). The Poisson Boltzmann solver was used to calculate energy in a water body to model physiological circumstances. Molecular electrostatic properties were analyzed.

3. Results

3.1. Homology modeling

The TMPRSS2-spike structure is currently unavailable, its 3-dimensional structure with 234 amino acids was modeled utilizing human plasma kallikrein as a template to better explain the substrate binding mechanism to the TMPRSS2 receptor. With the TMPRSS2 sequence, the plasma kallikrein template has the highest level of similarity and identity. TMPRSS2 model was subjected to model refinement and energy minimization (Fig. 2A). The Ramachandran plot (Fig. 2B) shows the quality of the projected models, since most residues are in favorable and permitted areas of the plot.

3.2. Active site prediction

Characterization of the binding site and its physicochemical characteristics is essential not only in the creation of new compounds, but also in the identification and optimization of molecules with known binding profiles. The template structures derived from the PDB database did not contain any co-crystallized ligands. As a result, the binding sites were obtained using SiteMap (Table 1 and Fig. 3). Site1 was selected as

Table 1

Proprieties for the five predicted active sites of the TMPRSS2 protein.

Title	SiteScore	size	Dscore	volume	exposure	contact	phobic	philic	don/acc
Site 1	1.016	110	1.067	262.738	0.622	0.885	0.73	0.829	1.705
Site 2	0.951	27	0.955	43.904	0.46	1.383	4.946	0.294	0.99
Site 3	0.875	71	0.86	169.785	0.587	0.954	0.658	1.048	1.375
Site 4	0.866	66	0.857	282.289	0.707	0.811	0.495	0.973	1.179
Site 5	0.727	41	0.518	95.354	0.667	1.022	0.092	1.517	0.362

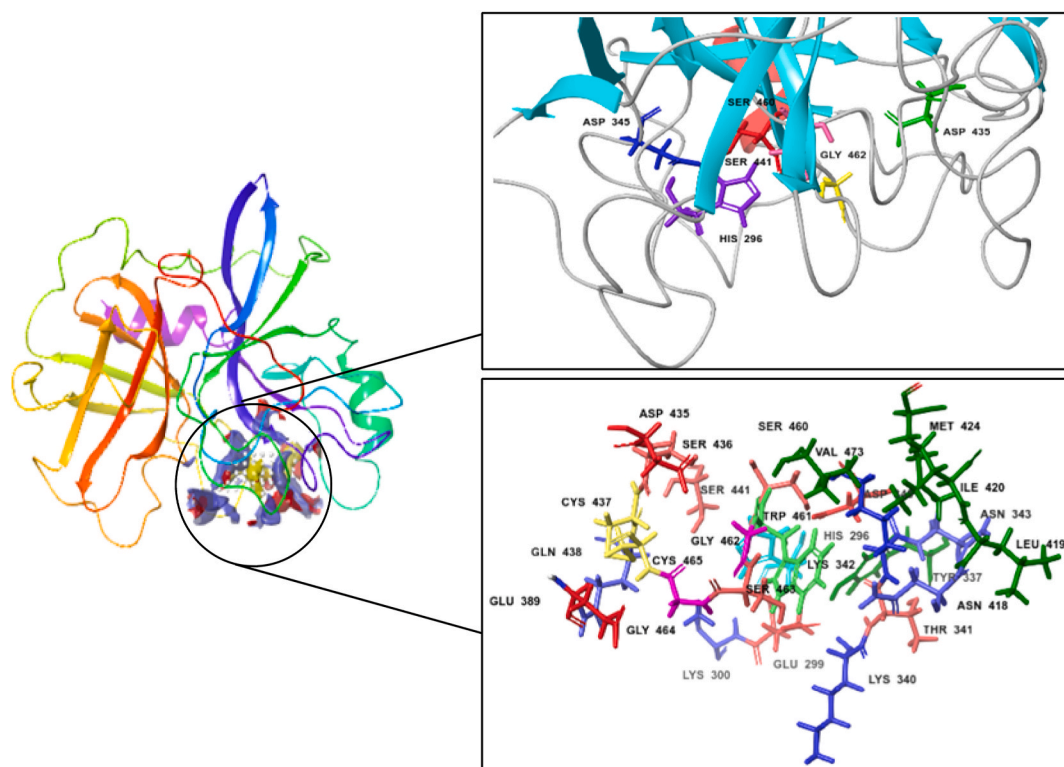


Fig. 3. The selected active site of TMPRSS2.

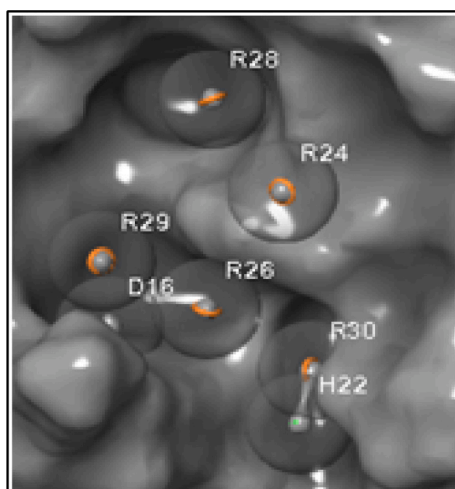


Fig. 4. The generated e-pharmacophore model of TMPRSS2 active site.

binding site because of high site score and Dscore and include the residues HIS296, GLU299, LYS300, TYR337, LYS340, THR341, LYS342, ASN343, ASP345, GLU389, ASN418, LEU419, ILE420, MET424, ASP435, CYS437, GLN438, SER441, SER460, TRP461, GLY462, SER463, GLY464, CYS465, VAL473. The catalytic residues HIS 296 ASP 345 and SER 441 have also been located in the tmprss2 model's terminal serine domain which validated the binding site [48].

3.3. E-pharmacophore model and pharmacophore-based database screening

Based on amino acid residues of the selected active site, a seven-featured pharmacophore hypothesis consisting of one donor (D), one hydrophobic region (H) and five aromatic rings (R) were obtained

(Fig. 4).

Validated e-pharmacophore models of TMPRSS2 were used to screen over 4 million prepared molecules from the enamine database. From the total numbers of compounds and conformers present in database. 31340 compounds match the screened hypothesis. Following that, these hits went through a three-tier Glide docking method for docking-based ligand screening.

3.4. Docking of the generated compounds and MM-GBSA

In computational chemistry, docking has now become a common method for estimating the affinity of molecules to their targets. The 31059 compounds that obtained from phase screen were narrowed down to only 10 top-ranked molecules with docking score -9.68 to -8.017 kcal/mol, through successive docking by Glide HTVS, SP and XP in order to determine their binding interactions and affinity to TMPRSS2 compared to docking score of -5.42 and -6.46 kcal/mol for references nafamostat and ambroxol as shown in Table 2. In addition, the top 10 compounds were investigated for MMGBSA binding free energy calculations. They showed binding energies better than the reference compounds, ranging from -63.55 to -89.75 kcal/mol.

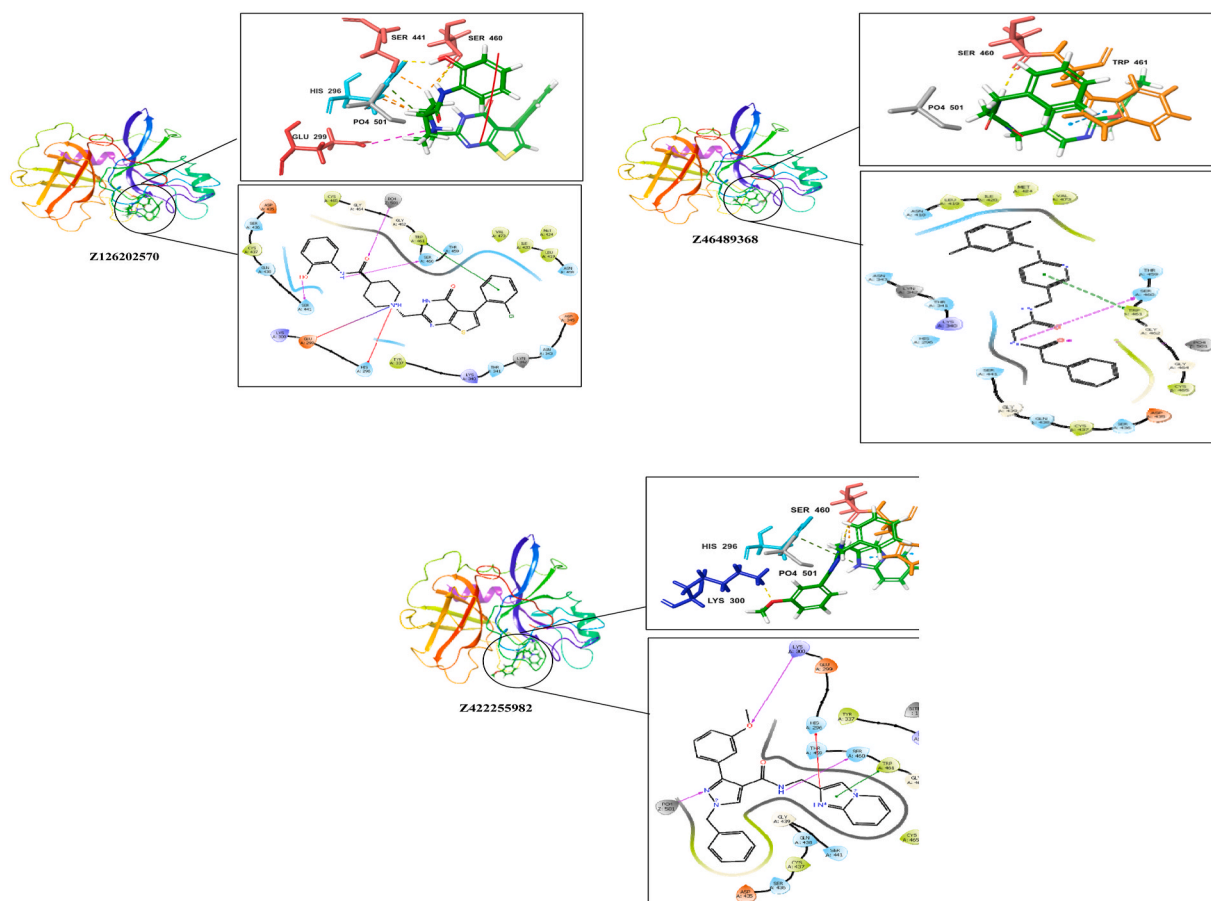
The interactions of top 3 molecules were selected for further analysis (Fig. 5). Z126202570 molecule forms four hydrogen bonds, one pi-cation interaction, one pi-pi interaction and hydrophobic interactions with the active pocket of TMPRSS2. Amino acids involved in the hydrogen bond formation are GLU299, GLY464, ASP435 and PO4501. One pi-cation interaction was observed between molecule and HIS296 residue and the one pi-pi interaction with TRP461. The hydrophobic interactions were formed with CYS437, CYS465, TRP461, TYR337, VAL473, ILE420, LEU419, MET424 and TYR33 residues.

Binding affinity between active pocket amino acids of TMPRSS2 and Z46489368 was maintained with two hydrogen bonds, one pi-pi interaction and hydrophobic interactions. SER460 and PO4501 made the hydrogen bonding, the one pi-pi interaction with TRP461 and the hydrophobic interaction were observed with the CYS437, CYS465,

Table 2

Docking score, MM-GBSA results, interactions and fitness score of the best ten compounds with TMPRSS2.

Name	docking score	MM -GBSA ΔG Bind	Fitness	Pi-pi interactin	Pi-cation interaction	Hydrogen bonding interaction	Hydrophobic interaction
Nafamostat	-5.424	-59.76	-	-	LYS300	GLU299, GLY464, ASP435, PO4501	CYS437, TRP461, VAL473, CYS465, ALA466, PRO301
Ambroxol	-6.464	-55.48	-	-	HIS296	SER460, THR431	TYR337, TRP461, CYS437, CYS465
Z126202570	-9.68	-89.75	0.643	TRP461	HIS296	SER460, PO4501, SER441	CYS437, CYC465, TRP461, TYR337, VAL473, ILE420, LEU419, MET424, TYR33
Z46489368	-8.996	-77.83	1.039	TRP461	-	SER460, PO4501	CYS437, CYC465, TRP461, VAL473, ILE420, LEU419, MET424.
Z422255982	-8.781	-74.48	0.947	TRP461	HIS296	SER460, LYS300, PO4501.	CYS437, CYC465, TRP461, TYR337.
Z168796704	-8.548	-73.53	0.648	TRP461	HIS296	PO4501.	CYS437, CYC465, TRP461, TYR337.
Z94732345	-8.489	-63.55	1.298	TRP461	-	ASN418, LYN342, PO4501.	CYS437, CYC465, TRP461, LEU419, MET424.
Z67551341	-8.287	-83.75	1.165	TRP461	HIS296	SER460, SER436	CYS437, CYC465, TRP461
Z166819168	-8.054	-82.49	1.297	TRP461	HIS296	SER436, LYS300, GLY464	CYS437, CYC465, LEU419, MET424, ILE420, PRO301
Z1462300333	-8.043	-80.25	1.083	TRP461	-	-	CYS437, CYC465, TRP461, TYR337, VAL473, ILE420, LEU419, MET424
Z109718002	-8.024	-74.95	1.146	TRP461	-	-	CYS437, CYC465, TRP461, TYR337, VAL473, ILE420, LEU419, MET424
Z168797262	-8.017	-52.13	1.527	TRP461	TRP461	-	CYS437, CYC465, TRP461, TYR337, ILE420, LEU419, MET424

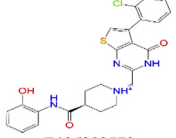
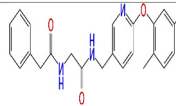
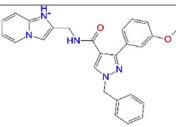
**Fig. 5.** The 2D and 3D interactions of the best three compounds with TMPRSS2.

TRP461, VAL473, ILE420, LEU419 and MET424 residues. Many interactions were formed between TMPRSS2 active pocket and Z422255982. Of these interactions, three hydrogen bonds with SER460, LYS300 and PO4501 residues, one is a pi-pi interaction with TRP461, one is a pi-cation interaction with HIS296 residue and the remaining were hydrophobic interactions with CYS437, CYC465, TRP461 and TYR337 residues.

3.5. Molecular dynamics

Docking methods are often quick and imprecise, but they lack protein flexibility, which might compromise the precision of the resulting ligand-protein complexes. As a result, more precise molecular dynamic modeling approaches may provide a better docking complement. Here, we examined the various molecular simulation results including RMSD,

Table 3
Average of RMSD and RMSF values for the three compounds.

Name	RMSD (Å)		RMSF (Å)	
	C α	ligand	C α	ligand
 Z126202570 Z126202570	2.35 \pm 0.25	2.32 \pm 0.42	1.04 \pm 0.66	1.18 \pm 0.48
	 Z46489368 Z46489368	2.35 \pm 0.25	1.60 \pm 0.45	1.04 \pm 0.67
 Z422255982 Z422255982		2.35 \pm 0.24	1.91 \pm 0.42	1.04 \pm 0.66

RMSF and PLC.

RMSD represents the dynamic stability and conformation changes occur in C α backbone of the TMRSS-inhibitor complexes during the simulation. The average RMSD values of the TMRSS2 protein and TMRSS2-inhibitor complexes are shown in Table 3 and Fig. 6. The C α backbone RMSD organization patterns of the TMRSS2 and its inhibitors complexes showed no significant changes, indicating their stability. The RMSF was used to evaluate the flexibility of the ligands and proteins. Throughout the simulation, the coordinates of the C α atom fluctuate from their typical location for example the average C α fluctuation in Z126202570 was from 0.66 to 1.04 Å. The average RMSF values of the TMRSS2 protein and ligands are shown in Table 3 and Fig. 7. The lower

RMSF and RMSD (≤ 2.35 Å) average for both the TMRSS2 and its inhibitors indicate how reliable our predicated model and the affinity of these promising inhibitors.

As it can be seen in Fig. 8A, Z126202570 interacts with GLY391 (23%), GLU299(7%), GLN438(5%), GLU389(5%), CYS297(3%) through H-bonds in less contact time during the whole time of stimulation. Water bridge was also formed with GLU389 (20%), LYS300 (15%), and CYS297 (15%), HIS296 (10%), GLY391 (10%). Hydrophobic interactions were formed with VAL280 (15%), LYS340 (10%) and VAL298 (5%). Finally, two ionic interactions were formed with GLU299 and GLU399 mainly with N14 of the 3H,4H-thieno[2,3-d]pyrimidin-4-one moiety.

Z46489368 formed H-bonds with SER460 (40%), SER441(20%), GLY462 (20%) and GLU389(5%) with more contact time during the whole time of stimulation (Fig. 8B). Water bridge bonding was also formed with HIS296 (30%), MET424 (15%), SER436 (15%), GLN438 (25%), and ASP440 (10%). Finally, hydrophobic interactions were formed with TRP461 (70%).

While in Z422255982-TMRSS2 complex (Fig. 8C), only two types of interactions were formed. GLU299 maintains strong hydrogen bond interaction for 80% of the stimulation time. Weak water bridge interactions were shown by HIS296, CYS297, and LYS340.

3.6. Density functional theory

The molecular orbitals provide details on electrical and optical characteristics, as well as quantum chemistry, and are utilized to understand the different types of conjugated system processes. The reactivity of compounds is also linked to molecular orbital energy and orbital coefficients according to molecular orbital theory. The B3LYP/6-311** basis set was used to calculate the graphical representations of MOs and the solvation energy for the top three compounds as shown in Table 4 and Fig. 9.

The HOMO orbitals in Z126202570 are located on the N8 that interacts with SER460, aromatic ring and O1 that interact with SER441 and O10 that interacts with PO4501. The LUMO orbitals are found on S19, N17, N32, C16, C18, C20, C18 of 3H,4H-thieno[2,3-d]pyrimidin-4-

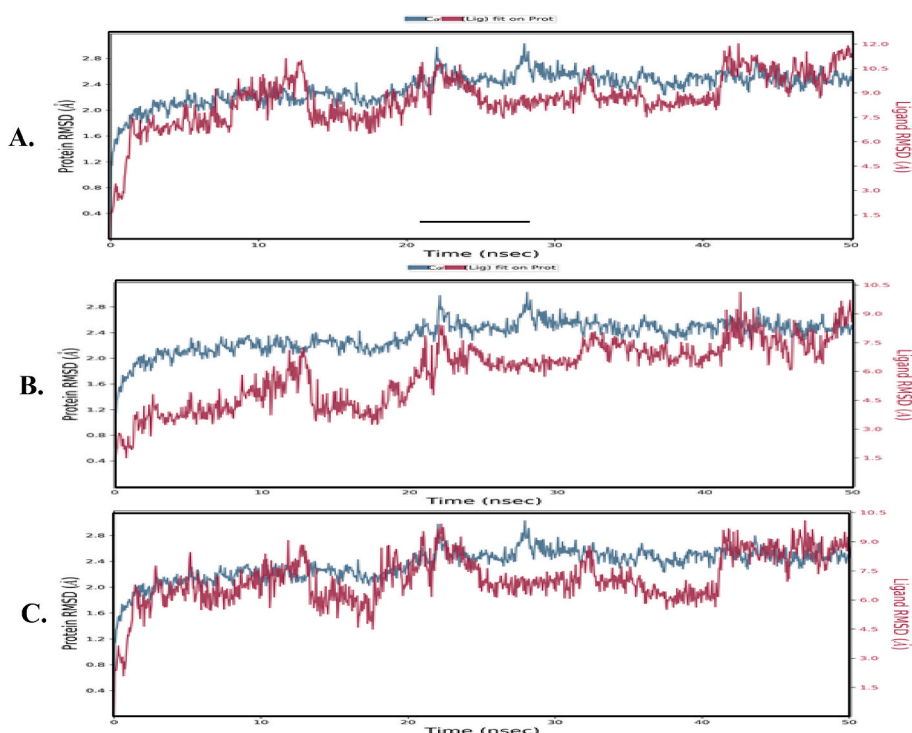


Fig. 6. PL-RMSD.

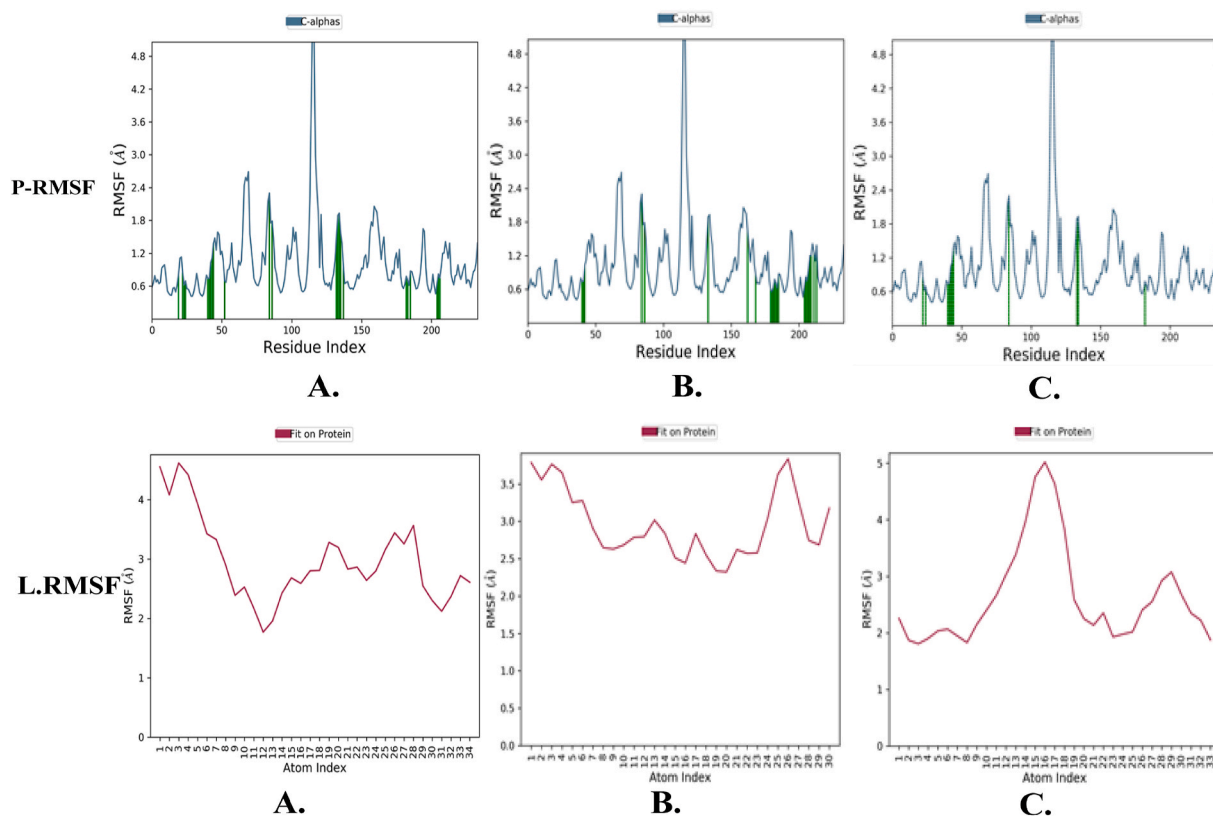


Fig. 7. Protein and ligand RMSF.

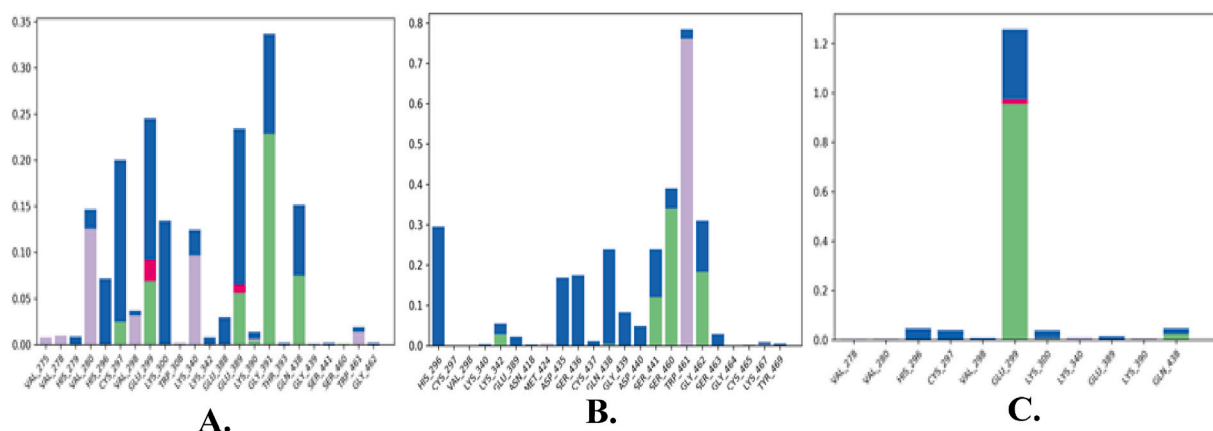


Fig. 8. Interaction diagram of the three compounds with TMPRS2.

Table 4
Quantum chemical proprieties of the references and the best three compounds.

Drug	HOMO kcal/mol	LUMO kcal/mol	HLG kcal/mol	Solvation Energy kcal/mol
Z126202570	-0.217	-0.075	-0.141	-55.86
Z46489368	-0.231	-0.035	-0.195	-20.73
Z422255982	-0.222	-0.068	-0.154	-50.45

one moiety and on N14 of the piperidine moiety that interacts with GLU299 and HIS296. In case of Z46489368 The HOMO orbitals are spread on 1,4-xylene moiety, 2,5-dimethylpyridine moiety that interacts with TRP461 and on O8 that link between these two moieties. The LUMO orbitals are located on 2,5-dimethylpyridine moiety. While in Z422255982 The HOMO orbitals are presented on O2 which interacts

with LYS300 and N10 that interacts with PO4501. The LUMO orbitals are spread over imidazo[1,2-a]pyridin-1-ium in which N33 interacts with HIS296 and TRP461.

The Molecular Electrostatic Potential (MEP) is a crucial way of understanding electrostatic allocation potential and visualizing the reactive site in chemical reactions involving both electrophilic and nucleophilic interactions. The ESP surface map provides a significant visual method to comprehend about reactive sites. MEP maps for the top three compounds were computed at the B3LYP/6-311** level to address potential sites for nucleophilic and electrophilic interactions and represented in Fig. 10.

The MEP map of Z126202570 range from -1.087 to 155.8 kcal/mol. The highest negative (red) regions were spread over O31 of 3H,4H-thieno[2,3-d]pyrimidin-4-one moiety and O10 of the propan-2-one moiety that interact with PO4501. While the highest positive (purple)

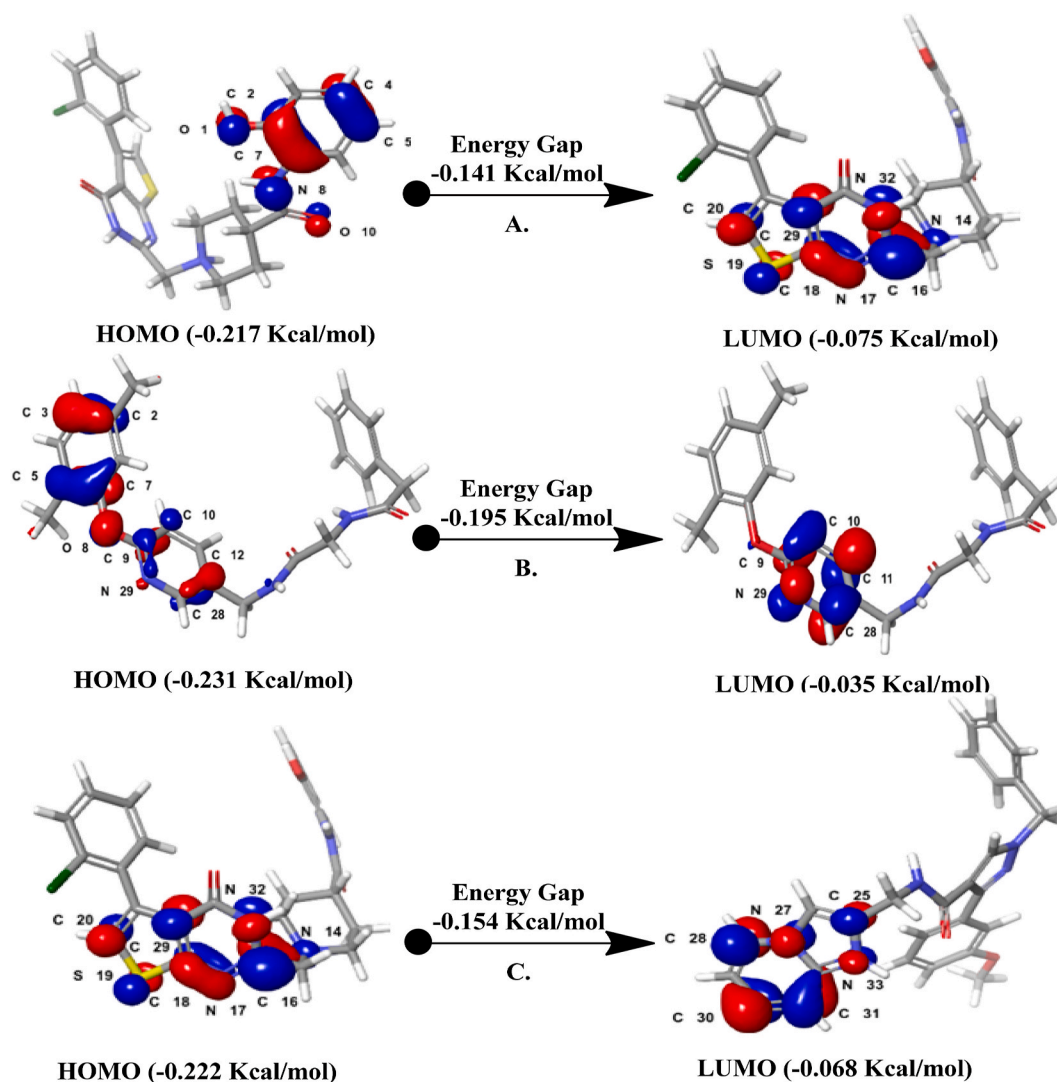


Fig. 9. HOMO and LUMO representation for the three compounds with TMRSS2.

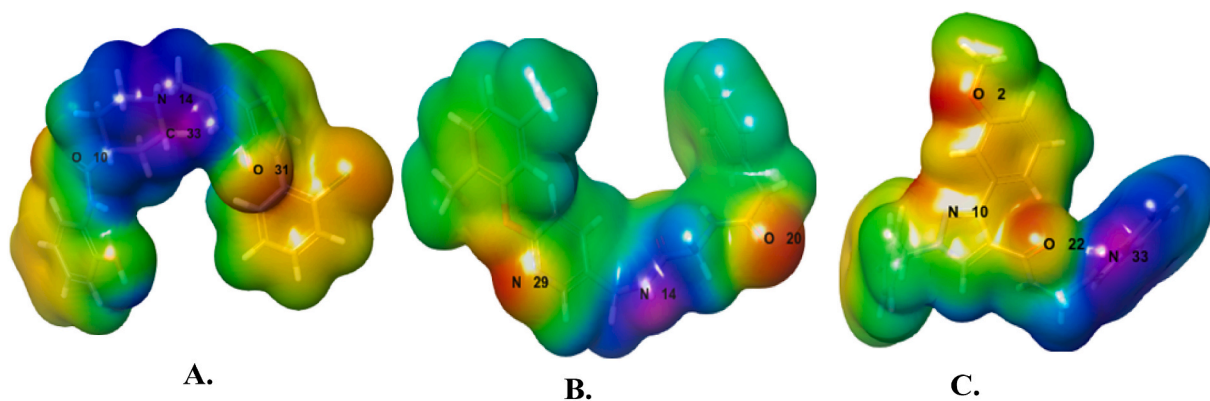


Fig. 10. MESP map for the three compounds with TMRSS2.

regions are located on N14 of the piperidine moiety that interacts with GLU299 and HIS296.

The MEP map of Z46489368 range from -53.9 to 64.8 kcal/mol. The red regions are spread over 2,5-dimethylpyridine moiety that interact with TRP461 and on O20 of N-(2-methylprop-2-en-1-yl)acetamide moiety. While the purple regions are presented on N14 of ethyl(methyl)

amine moiety.

The MEP map of Z422255982 range from -4.5 to 125.4 . The red regions are located on O2 of the benzenolate moiety that interacts with LYS300 and on N10 of the 1H-pyrazol-1-ide moiety. While the purple regions are found on N33 of the imidazo[1,2-a]pyridin-1-ium moiety that interacts with HIS296 and TRP461.

4. Conclusion

The TMPRSS2 protein is an important target of SARS-CoV-2 viral replication owing to its proteolytic activity. In the current study, a good homology model for TMPRSS2 has been built. E-pharmacophore model was developed and screened against one million compounds from enamine database, followed by molecular docking and MM-GBSA binding energy calculations to validate their affinity to the target. The top three molecules (Z126202570, Z46489368, and Z422255982) were also subjected to molecular dynamics simulations to confirm the stability of interactions obtained from docking analysis as well as density functional theory calculation. These results indicated that these compounds might be potential hits against TMPRSS2 and if experimentally confirmed may serve as future anti-covid agents.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

We acknowledge Mme Katia Dekimeche from Schrodinger for her support and help.

References

- Morens DM, Fauci AS. Emerging pandemic diseases: how we got to COVID-19. *Cell* 2020;182:1077–92. <https://doi.org/10.1016/j.cell.2020.08.021>.
- Wang W, Xu Y, Gao R, Lu R, Han K, Wu G, Tan W. Detection of SARS-CoV-2 in different types of clinical specimens. *JAMA, J Am Med Assoc* 2020;323:1843–4. <https://doi.org/10.1001/jama.2020.3786>.
- COVID-19 map - Johns Hopkins coronavirus resource center, (n.d.). <https://coronavirus.jhu.edu/map.html>. [Accessed 7 August 2021]. accessed.
- Mckee DL, Sternberg A, Stange U, Laufer S, Naujokat C. Candidate drugs against SARS-CoV-2 and COVID-19. *Pharmacol Res* 2020;157:104859. <https://doi.org/10.1016/j.phrs.2020.104859>.
- Su S, Wong G, Shi W, Liu J, Lai ACK, Zhou J, Liu W, Bi Y, Gao GF. Epidemiology, genetic recombination, and pathogenesis of coronaviruses. *Trends Microbiol* 2016; 24:490–502. <https://doi.org/10.1016/j.tim.2016.03.003>.
- Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, Zhao X, Huang B, Shi W, Lu R, Niu P, Zhan F, Ma X, Wang D, Xu W, Wu G, Gao GF, Tan W. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med* 2020;382:727–33. <https://doi.org/10.1056/nejmoa2001017>.
- Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus ADME, Fouchier RAM. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl J Med* 2012;367:1814–20. <https://doi.org/10.1056/nejmoa1211721>.
- Sheng WH. Coronavirus disease 2019 (covid-19). *J Intern Med Taiwan* 2020;31: 61–6. [https://doi.org/10.6314/JIMT.202004_31\(2\).01](https://doi.org/10.6314/JIMT.202004_31(2).01).
- Tortorici MA, Vesler D. Structural insights into coronavirus entry. *first ed. Elsevier Inc.*; 2019. <https://doi.org/10.1016/bs.aivir.2019.08.002>.
- Li C, Yang Y, Ren L. Genetic evolution analysis of 2019 novel coronavirus and coronavirus from other species. *Infect Genet Evol* 2020;82:1–3. <https://doi.org/10.1016/j.meegid.2020.104285>.
- De Wit E, Van Doremalen N, Falzarano D, Munster VJ. SARS and MERS: recent insights into emerging coronaviruses. *Nat Rev Microbiol* 2016;14:523–34. <https://doi.org/10.1038/nrmicro.2016.81>.
- Drosten C, Günther S, Preiser W, van der Werf S, Brodt H-R, Becker S, Rabenau H, Panning M, Kolesnikova L, Fouchier RAM, Berger A, Burguière A-M, Cinatl J, Eickmann M, Escriviou N, Grywna K, Kramme S, Manuguerra J-C, Müller S, Rickerts V, Stürmer M, Vieth S, Klenk H-D, Osterhaus ADME, Schmitz H, Doerr HW. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N Engl J Med* 2003;348. <https://doi.org/10.1056/nejmoa030747>. 1967–1976.
- V'kovski P, Kratzel A, Steiner S, Stalder H, Thiel V. Coronavirus biology and replication: implications for SARS-CoV-2. *Nat Rev Microbiol* 2021;19:155–70. <https://doi.org/10.1038/s41579-020-00468-6>.
- Parthasarathy P, Vivekanandan S. An extensive study on the COVID-19 pandemic, an emerging global crisis: risks, transmission, impacts and mitigation. *J Infect. Public Health* 2020;249–59. <https://doi.org/10.1016/j.jiph.2020.12.020>.
- Glowska I, Bertram S, Muller MA, Allen P, Soilleux E, Pfefferle S, Steffen I, Tsegaye TS, He Y, Gnirss K, Niemeier D, Schneider H, Drosten C, Pöhlmann S. Evidence that TMPRSS2 activates the severe acute respiratory syndrome coronavirus spike protein for membrane fusion and reduces viral control by the humoral immune response. *J Virol* 2011;85:4122–34. <https://doi.org/10.1128/jvi.02232-10>.
- Zhou Y, Vedantham P, Lu K, Agudelo J, Carrion R, Nunneley JW, Barnard D, Pöhlmann S, McKerron JH, Renslo AR, Simmons G. Protease inhibitors targeting coronavirus and filovirus entry. *Antivir Res* 2015;116:76–84. <https://doi.org/10.1016/j.antiviral.2015.01.011>.
- Gierer S, Bertram S, Kaup F, Wrensch F, Heurich A, Kramer-Kuhl A, Welsch K, Winkler M, Meyer B, Drosten C, Dittmer U, von Hahn T, Simmons G, Hofmann H, Pöhlmann S. The spike protein of the emerging betacoronavirus EMC uses a novel coronavirus receptor for entry, can be activated by TMPRSS2, and is targeted by neutralizing antibodies. *J Virol* 2013;87:5502–11. <https://doi.org/10.1128/jvi.00128-13>.
- Ou Xiuyuan, Zheng Wangliang, Shan Yiwei, Mu Zhixia, Dominguez Samuel R, Holmes Kathryn V, Qian Zhaohui. Identification of the fusion peptide-containing region in Betacoronavirus spike glycoproteins. *J Virol* 2016 May 27;90(12): 5586–600. <https://doi.org/10.1128/JVI.00015-16>.
- Hussain M, Jabeen N, Amanullah A, Baig AA, Aziz B, Shabbir S, Raza F, Uddin N. Molecular docking between human tmprss2 and sars-cov-2 spike protein: conformation and intermolecular interactions. *AIMS Microbiol* 2020;6:350–60. <https://doi.org/10.3934/microbiol.2020021>.
- Tang T, Bidon M, Jaimes JA, Whittaker GR, Daniel S. Coronavirus membrane fusion mechanism offers a potential target for antiviral development. *Antivir Res* 2020;178:104792. <https://doi.org/10.1016/j.antiviral.2020.104792>.
- Shulla A, Heald-sargent T, Subramanya G, Zhao J, Perlman S, Gallagher T, Shulla A, Heald-sargent T, Subramanya G, Zhao J, Perlman S, Gallagher T. A transmembrane serine protease is linked to the severe acute respiratory syndrome coronavirus receptor and activates virus entry A transmembrane serine protease is linked to the severe acute respiratory syndrome coronavirus receptor and activates virus E. <https://doi.org/10.1128/JVI.02062-10>; 2011.
- Matsuyama S, Nagata N, Shirato K, Takeda M, Taguchi F. Efficient activation of the severe acute respiratory syndrome coronavirus spike protein by the transmembrane protease efficient activation of the severe acute respiratory syndrome coronavirus spike protein by the transmembrane protease TMPRSS2. <https://doi.org/10.1128/JVI.01542-10>; 2010.
- Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, Schiergens TS, Herrler G, Wu NH, Nitsche A, Müller MA, Drosten C, Pöhlmann S. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* 2020;181:271–80. <https://doi.org/10.1016/j.cell.2020.02.052>. e8.
- Heurich A, Hofmann-winkler H, Gierer S, Liepold T, Jahn O. Proteolysis by TMPRSS2 augments entry driven by the severe. *Acute* 2014;88:1293–307. <https://doi.org/10.1128/JVI.02202-13>.
- Kuhn JH, Li W, Choe H, Farzan M. Cellular and molecular life sciences angiotensin-converting enzyme 2 : a functional receptor for SARS coronavirus, 61; 2004. p. 2738–43. <https://doi.org/10.1007/s00018-004-4242-5>.
- Simmons G, Zmora P, Gierer S, Heurich A, Pöhlmann S. Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information; 2020.
- Antalis TM, Thomas H. Membrane-anchored serine proteases in health and disease I. Introduction. *first ed. Elsevier Inc.*; 2011. <https://doi.org/10.1016/B978-0-12-385504-6.00001-4>.
- Singh H, Choudhari R, Nema V, Khan AA. Division of molecular biology Jo ur I P re. *Microb Pathog* 2021;104621. <https://doi.org/10.1016/j.micpath.2020.104621>.
- Stopsack KH, Mucci LA, Mph S, Antonarakis ES, Nelson PS, Kantoff PW. TMPRSS2 and COVID-19 : serendipity or opportunity for intervention ? <https://doi.org/10.1158/2159-8290.CD-20-0451>; 2020.
- Iwata-yoshikawa N, Okamura T, Shimizu Y, Hasegawa H, Takeda M, Nagata N, Health G. TMPRSS2 contributes to virus spread and immunopathology in the airways of murine models after coronavirus infection. <https://doi.org/10.1128/JVI.01815-18>; 2019.
- Zhu H, Du W, Song M, Liu Q, Herrmann A, Huang Q. Spontaneous binding of potential COVID-19 drugs (Camostat and Nafamostat) to human serine protease TMPRSS2. *Comput Struct Biotechnol J* 2021;19:467–76. <https://doi.org/10.1016/j.csbj.2020.12.035>.
- Ragia G, Manolopoulos VG. Inhibition of SARS-CoV-2 entry through the ACE2/TMPRSS2 pathway: a promising approach for uncovering early COVID-19 drug therapies. *Eur J Clin Pharmacol* 2020;76:1623–30. <https://doi.org/10.1007/s00228-020-02963-4>.
- Idris MO, Yekeen AA, Alakanse OS, Durojaye OA. Computer-aided screening for potential TMPRSS2 inhibitors: a combination of pharmacophore modeling, molecular docking and molecular dynamics simulation approaches. *J Biomol Struct Dyn* 2020;5638–56. <https://doi.org/10.1080/07391102.2020.1792346>. 1–19.
- Vyas VK, Ukawala RD, Ghate M, Chintha C. Homology modeling a fast tool for drug discovery: current perspectives. *Indian J Pharmaceut Sci* 2012;74:1–17. <https://doi.org/10.4103/0250-474X.102537>.
- Singh K, Sharma MC, Sharma S, V Jain S, Avchar MH. An approach to design potent anti-Alzheimer's agents by 3D-QSAR studies on fused 5-, 6-bicyclic heterocycles as c-secretase modulators using kNN – MFA methodology. *Arab. J. Chem.* 2013;924–35. <https://doi.org/10.1016/j.arabj.2013.02.002>.
- Amnerkar ND, Bhusari KP. Synthesis, anticonvulsant activity and 3D-QSAR study of some prop-2-enamido and 1-acetyl-pyrazolin derivatives of aminobenzothiazole. *Eur J Med Chem* 2010;45:149–59. <https://doi.org/10.1016/j.ejmech.2009.09.037>.
- Bendix F, Wolber G, Seidel T. 3D pharmacophore elucidation and virtual screening strategies for 3D pharmacophore- based virtual screening. <https://doi.org/10.1016/j.ddtec.2010.11.004>; 2010.

- [38] Singh K, Kumawat NK, V Bhavthankar S, Avchar MH, Dhumal DM, Patil SD, V Jain S. Exploring 2D and 3D QSARs of benzimidazole derivatives as transient receptor potential melastatin 8 (TRPM8) antagonists using MLR and kNN-MFA methodology. *J. SAUDI Chem. Soc.* 2012;8. <https://doi.org/10.1016/j.jscs.2012.11.001>.
- [39] Tmprss2 - transmembrane protease serine 2 precursor - Homo sapiens (Human) - Tmprss2 gene & protein, (n.d.). <https://www.uniprot.org/uniprot/O15393>. [Accessed 21 September 2021]. accessed.
- [40] Protein BLAST: search protein databases using a protein query, (n.d.). <https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>. [Accessed 21 September 2021]. accessed.
- [41] Li Z, Partridge J, Silva-Garcia A, Rademacher P, Betz A, Xu Q, Sham H, Hu Y, Shan Y, Liu B, Zhang Y, Shi H, Xu Q, Ma X, Zhang L. Structure-guided design of novel, potent, and selective macrocyclic plasma kallikrein inhibitors. *ACS Med Chem Lett* 2017;8:185–90. <https://doi.org/10.1021/acsmchemlett.6b00384>.
- [42] Lyne PD, Lamb ML, Saeh JC. Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring. *J Med Chem* 2006;49:4805–8. <https://doi.org/10.1021/jm060522a>.
- [43] Sastry GM, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des* 2013;27:221–34. <https://doi.org/10.1007/s10822-013-9644-8>.
- [44] Elbadwi FA, Khairy EA, Alsamani FO, Mahadi MA, Abdalrahman SE, Alsharf Z, Ahmed M, Elsayed I, Ibraheem W, Alzain AA. Informatics in Medicine Unlocked Identification of novel transmembrane Protease Serine Type 2 drug candidates for COVID-19 using computational studies. *Informatics Med. Unlocked.* 2021;26:100725. <https://doi.org/10.1016/j.imu.2021.100725>.
- [45] Databases - enamine, (n.d.). <https://enamine.net/11-databases>. [Accessed 10 September 2021]. accessed.
- [46] Salam NK, Nuti R, Sherman W. Novel method for generating structure-based pharmacophores using energetic analysis. *J Chem Inf Model* 2009;49:2356–68. <https://doi.org/10.1021/ci900212v>.
- [47] Suryanarayanan V, Singh SK. Assessment of dual inhibition property of newly discovered inhibitors against PCAF and GCN5 through in silico screening, molecular dynamics simulation and DFT approach. *J Recept Signal Transduct Res* 2015;35:370–80. <https://doi.org/10.3109/10799893.2014.956756>.
- [48] Wilson S, Greer B, Hooper J, Zijlstra A, Walker B, Quigley J, Hawthorne S. The membrane-anchored serine protease. Tmprss2 , activates PAR-2 in prostate cancer cells 2005;972:967–72.