



Identification of novel TMPRSS2 inhibitors against SARS-CoV-2 infection: a structure-based virtual screening and molecular dynamics study

Suman Manandhar¹ · K. Sreedhara Ranganath Pai¹ · Praveen T. Krishnamurthy² · Kiran Ammu V. V. V. Ravi² · Garikapati Kusuma Kumari²

Received: 5 January 2022 / Accepted: 13 March 2022 / Published online: 23 March 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

The scientific insights gained from the severe acute respiratory syndrome (SARS) and the middle east respiratory syndrome (MERS) outbreaks are helping scientists to fast-track the antiviral drug discovery process against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Coronaviruses, as well as influenza viruses, depend on host type 2 transmembrane serine protease, TMPRSS2, for entry and propagation in the human cell. Recent studies show that SARS-CoV-2 also uses TMPRSS2 for its cell entry. In the present study, a structure-based virtual screening of 52,337, protease ligands downloaded from the Zinc database was carried out against the homology model of TMPRSS2 protein followed by the molecular dynamics-based simulation to identify potential TMPRSS2 hits. The virtual screening has identified 13 hits with a docking score range of -10.447 to -9.863 and glide energy range of -60.737 to -40.479 kcal/mol. The binding mode analysis shows that the hit molecules form H-bond (Asp180, Gly184 & Gly209), Pi-Pi stacking (His41), and salt bridge (Asp180) type of contacts with the active site residues of TMPRSS2. In the MD simulation of ZINC000013444414, ZINC000137976768, and ZINC000143375720 hits show that these molecules form a stable complex with TMPRSS2. The complex equilibrates well with a minimal RMSD and RMSF fluctuation. All three structures, as predicted in Glide XP docking, show a prominent interaction with the Asp180, Gly184, Gly209, and His41. Further, MD simulation also identifies a notable H-bond interaction with Ser181 for all three hits. Among these hits, ZINC000143375720 shows the most stable binding interaction with TMPRSS2. The present study is successful in identifying TMPRSS2 ligands from zinc data base for a possible application in the treatment of COVID-19.

Keywords TMPRSS2 · SARS-CoV-2 · COVID-19 · Virtual screening · Molecular docking · Molecular dynamics

Introduction

The coronaviruses (CoV) are genotypically and serologically divided into 4-subfamilies: α , β , γ , and δ -CoVs. Among them, α and β -CoV affect mammals, while γ and

δ -CoVs affect birds. The SARS-CoV-2, a β -coronavirus, is an enveloped non-segmented positive-sense RNA virus (Fig. 1). The SARS-CoV-2 is round or oval (diameter of 60~100 nm) shaped with a 29.9 kb genome size (one of the giant RNA virus genomes). The viral RNA, along with nucleoprotein (N), forms a helical nucleocapsid. A lipid-bilayer membrane encapsulates the nucleocapsid and is associated with transmembrane protein (M), spike glycoprotein (S), and a non-glycosylated envelope protein (E) (Fig. 1). S-glycoprotein facilitates cellular entry by binding to angiotensin-converting enzyme 2 (ACE2) host cell receptors. The transmembrane M protein facilitates the transport of nutrients, bud release, and formation of the envelope [1, 2]. Recent WHO situation report published on January 2022 stated that globally ~ 330 million people were infected with more than 5 million deaths due to coronavirus disease 2019

Suman Manandhar and K. Sreedhara Ranganath Pai contributed equally to this work

✉ Praveen T. Krishnamurthy
praveentk7812@gmail.com; praveentk@jssuni.edu.in

¹ Department of Pharmacology, Manipal College of Pharmaceutical Sciences, Manipal Academy of Higher Education, Manipal 576 104, India

² Department of Pharmacology, JSS College of Pharmacy (JSS Academy of Higher Education & Research), Ooty, The Nilgiris 643 001, Tamil Nadu, India

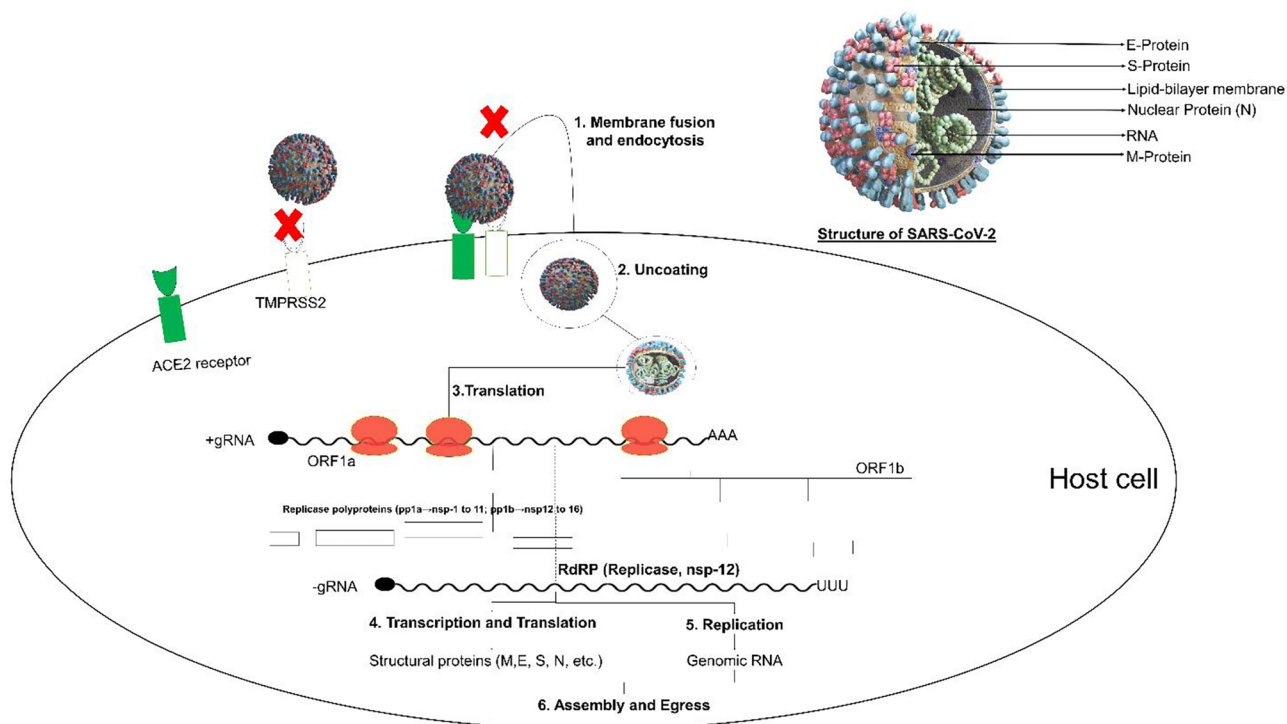


Fig. 1 Structure and replication mechanism of SARS-CoV-2 (SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; RdRp, RNA-dependent RNA polymerase; ORF, open reading frame intermediate compartment; pp1a, polyprotein 1a; pp1ab, polyprotein 1ab;

nsp, non-structural proteins; TMPRSS2, type 2 transmembrane serine protease; ACE2, angiotensin-converting enzyme 2; N, nucleoprotein; M, transmembrane protein; S, spike glycoprotein; E, non-glycosylated envelope protein)

(COVID-19) [3]. The scientific insights acquired from SARS and MERS outbreaks is fast-tracking the antiviral drug discovery process against SARS-CoV-2. Druggable targets for antiviral drugs include non-structural viral proteins (3-chymotrypsin-like protease, papain-like protease, RNA-dependent RNA polymerase, and its helicase), viral structural proteins (S-glycoprotein) and host protein, TMPRSS2 (Fig. 1). Several preclinical and clinical trials were underway across the globe to test the potential treatments against these protein targets [4–8].

Human coronaviruses could enter cell via two pathways, the cathepsins mediated endosomal pathway and TMPRSS2 mediated the cell-surface or the early endosome pathway [9]. SARS-CoV-2 is reported to utilize the later path, where its spike glycoprotein (S) binds with host ACE2 and TMPRSS2 receptors to enable cellular entry [10–13]. Uncoating facilitates the use of genomic RNA as mRNA to translate the replicase polyproteins. The translation of the replicase gene produces the polyprotein 1a (pp1a) and polyprotein 1ab (pp1ab). Autoproteolytic cleavage of pp1a and pp1ab produces 11 (nsp1–nsp11) and 15 (nsp1–nsp10 and nsp12–nsp16) non-structural proteins, respectively (Fig. 1). The nsp12 is an RNA-dependent RNA polymerase (replicase, RdRp). The replicase uses genomic RNA as a template to produce negative-sense genomic RNAs (gRNAs), which is used for preparing

progeny positive-sense RNA genomes (Fig. 1) [12]. A nested set of sub-genomic RNAs (sgRNAs) are synthesized by replicase through discontinuous transcription of the genome. The sgRNAs are later translated into structural and accessory proteins (Fig. 1). The structural proteins, S, M, and E produced in the endoplasmic reticulum (ER) are transported to the ER–Golgi intermediate compartment (ERGIC) for virion assembly [2]. The N proteins associate with progeny genomic RNA to form the nucleocapsids. The smooth-wall vesicles transport the assembled virions in the ERGIC to the cell membranes for the release of the mature virus particles [14].

Computational-based studies including molecular docking and network pharmacology play a major role in the identification as well as shortlisting of the hit molecules and the pathway involved in their mechanism of action. Several in silico-based virtual screening using molecular docking studies have been conducted for the identification of the potential hit molecules from the natural molecules, food components, marine source molecules, repurposing of the existing drugs, and traditional medicine [15–22]. The are several approaches targeting proteins responsible for the replication of the virus like main protease (Mpro), spike protein, and RNA-dependent RNA polymerase (RdRp) and protein targeting the entry of virus cellular angiotensin converting enzyme 2 (ACE2) receptor [15, 19, 23–25]. In the

current approach, we have chosen *in silico* docking as well as molecular dynamics approach to identify the potential hit molecule from the library of the molecules with proven protease inhibitory activity against TMPRSS2 protein.

TMPRSS2 belongs to S1A class of serine proteases family like factor Xa and trypsin. The S-protein is processed by TMPRSS2 into two functional subunits, N-terminal receptor-binding domain (S1) and a C-terminal membrane fusion domain (S2) at the S1/S2 cleavage site. The S1 domain facilitates ACE2 recognition and initiates a conformational change in the S2 subunit, assisting membrane fusion via insertion of fusion peptide into the host cell membrane and delivering the viral nucleocapsid into the host cytoplasm [9, 13, 26–36]. The S2 domain contains a fusion peptide (FP), a second proteolytic site (S2), an internal fusion peptide (IFP), and two heptad-repeat domains (HR1 & HR2) before the transmembrane domain (TM) (Fig. 2). It is hypothesized that both FP and IFP are involved in the facilitation of the viral entry process, and, therefore, S-protein must be cleaved at both S1/S2 and S2' cleavage sites to facilitate viral entry [34–36].

TMPRSS2 was first identified in prostate cancer, where it was found to be overexpressed and controlled by androgen receptor signalling and it has approximately 41% similarity with TMPRSS15. It promotes prostate cancer metastasis by activating the hepatocyte growth factor (HGF) and was inhibited by use of TMPRSS2 inhibitors [33, 37–39]. In humans, TMPRSS2 is expressed in the prostate, stomach, colon, salivary gland, and epithelia of the gastrointestinal, urogenital, and respiratory tracts [40]. Previous studies confirmed the role TMPRSS2 in activating proteases for respiratory influenza viruses [29, 41]. The purpose of host TMPRSS2 in spike protein activation was well demonstrated in animal models of SARS-CoV and MERS-CoV infection. The results demonstrated that lack of TMPRSS2 (in TMPRSS2-knockout mice) significantly reduced the airway infection and spread [28]. Further, when *Tmprss2*(-/-)-deficient mice were infected with a re-assorted influenza A virus (IAV) H10 subtype hemagglutinin (HA), they show no abnormal clinical signs, lung lesions, viral antigen, and body weight loss, when compared to wild-type mice [42]. Another

study suggests that TMPRSS2 is a significant HA-activating protease of IAV and IBV (influenza B viruses) in primary human type II pneumocytes and human bronchial cells [27]. TMPRSS2-positive VeroE6 cells are highly susceptible to SARS-CoV-2 infection, suggesting the critical role played by TMPRSS2 in viral entry into the host cell [31]. SARS-CoV-2 receptors, ACE2, and TMPRSS2, are reported to be primarily expressed in bronchial transient secretory cells [32]. A recent study confirms that SARS-CoV-2 takes advantage of the host ACE2 for entry and the serine protease, TMPRSS2, for S-protein priming [30]. Further, TMPRSS2 inhibitor, Camostat, is reported to block the host cell entry of SARS-CoV-2 [30]. The above results strongly suggest that TMPRSS2 is a vital protein required for host cell entry of SARS-CoV-2 and, therefore, constitutes a treatment option. The 3D crystal structure of TMPRSS2 is not available, and hence, we have used the previously reported homology model of TMPRSS2, generated by using TMPRSS15 (PDB ID. 4DGJ) [34].

Methodology

Structure-based virtual screening

Virtual screening workflow of the Schrödinger software suite 2018–3 version (Maestro 11.7, Schrödinger, LLC, New York, NY, 2020) was used to carry out the screening of the Zinc database [43] against the active site of TMPRSS2.

Database and ligand preparation

A total of 52,337 molecules belonging to proteases category have been downloaded in 2D SDF format. The preparation of these 2D structures was carried out using the LigPrep module of the software. Briefly, the molecules were desalted and converted from 2D structure to corresponding low energy 3D structures, tautomeric, ionization (between pH 6.8 and 7.2 by choosing Epik module), and all possible stereoisomeric states were generated. The generated structures were energy minimized using OPLS 2005 force field.

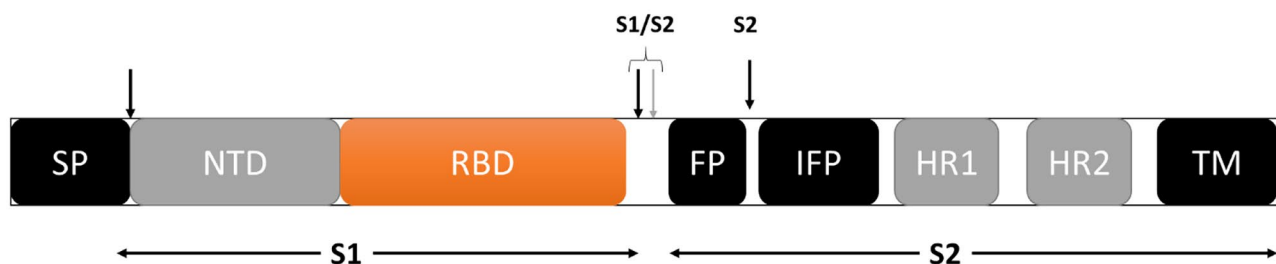


Fig. 2 Structure of TMPRSS2. S1, N-terminal receptor-binding domain; S2, C-terminal membrane fusion domain; SP, signal peptide; NTD, N-terminal domain (NTD), RBD, receptor-binding domain; FP, fusion

peptide; IFP, internal fusion peptide; HR1, heptad repeat 1; HR2, heptad repeat 2; TM, transmembrane domain. The SP, S1/S2, and S2' cleavage sites are indicated by arrows

Homology model of TMPRSS2

The homology model of TMPRSS2 protein was obtained from the TMPRSS15 crystallographic structure (PDB ID. 4DGJ) (with 41% of similarity on their peptide sequence) [20]. The obtained template is an aligned sequences of all the available S1A proteases followed by the identification of TMPRSS15 as the most suitable for the current study. PDB id 4DGJ was selected to build the homology model using Prime module of Schrodinger. The generated model was further validated by performing the MD simulation of 100 ns.

Protein preparation and receptor grid generation

The homology protein of TMPRSS2, generated by using TMPRSS15 (PDB ID. 4DGJ), was prepared by using the Protein Preparation Wizard. The protein was processed for bond orders, missing atoms, tautomer/ionization states, water orientations, and hydrogen bond networking. Constrained energy minimization was then performed using OPLS 3e 2005 force field. The receptor grid was generated using the prepared protein from the previous step. The size and position of the receptor grid box were defined using the centroid of workspace ligand (Benzamide; A: BEN 245) with van der Waals scaling factor of 1.0 and partial charge cutoff at 0.25.

Virtual screening

The virtual screening was carried out using the Glide Virtual Screening Workflow. The previously prepared ligands were used as input structures. Before subjecting to Glide Docking, the ADME properties of these ligands were predicted and prefiltered by Lipinski's rule and reactive functional group criteria. All ligands which pass through these prefilters were docked into the previously prepared receptor grid structure in three stages: in the first stage, the molecules were flexibly docked in Glide HTVS (high-throughput virtual screening) mode using the default settings. In the second stage, 10% of the lower-scoring hits from the previous step were flexibly docked in Glide SP (standard precision) mode. In the final stage, 10% of the lower-scoring hits of the last step were flexibly docked in Glide XP (extra precision) mode.

Binding energy calculation

The ligand binding energy of the selected hit molecules was calculated using prime module MMGBSA (molecular mechanics-generalized Born surface area).

Prime MM-GBSA uses the VSGB solvation model [38] which is dependent on the variable-dielectric generalized Born model and solvent as water under the OPLS3e force field.

The binding free energy of all the selected proteins–ligand complex is calculated using the formula:

$$\Delta G_{bind} = G_{complex} - (G_{protein} + G_{ligand})$$

where G = MME (molecular mechanics energy) + GSGB (SGB solvation model for polar solvation) + GNP (nonpolar solvation).

Molecular dynamics (MD) simulation

MD simulation for 100 ns was performed using Desmond software (Schrodinger software suite 2018–3 version) for the three hits having the highest docking score (ZINC000013444414), glide energy (ZINC000137976768), and MMGBSA dG bind free energy (ZINC000143375720). The system comprising of SPC solvent model, ligand, and protein complex was built in orthorhombic boundary with a dimension of $10 \text{ \AA} \times 10 \text{ \AA} \times 10 \text{ \AA}$, and the negative charge was neutralized with counter sodium ions. The system was then minimized into local energy minimum using hybrid method of steepest descent with minimum ten steps of steepest descent and the limited memory Broyden-Fletcher-Goldfarb-Shanno (LBFSGS) algorithms considering 3 LBFSGS vectors until gradient threshold of 25 kcal/mol/\AA was reached. Default options of minimization setup of Desmond were used that utilizes minimization with maximum of 2000 iterations and convergence threshold of 1 kcal/mol/\AA . The simulation was then carried out using an NPT ensemble system with 300.0 K temperature and 1.0 bar pressure for 100 ns duration. The trajectory was recorded for every 100.0 ps, and 1000 frames were captured, which is used for the calculation of root-mean-square deviation (RMSD) and root-mean-square fluctuation (RMSF).

Results and discussion

The structure-based virtual screening

The ligand preparation was performed using 52,337 protease molecules from zinc database which yielded 77,956 structures. The HTVS docking of these structures against the active site of the homology model of TMPRSS2 resulted in 4,567 hits. The further docking of these hits in the SP mode yielded in 615 hits. The final docking of these hits in the XP mode produced 13 hits with a docking score range of -10.447 to -9.863 and glide energy range of -60.737 to -40.479 kcal/mol (Table 1 and Fig. 3).

Table 1 Results of Structure-Based Virtual Screening of ZINC database against TMPRSS2 active site

Zinc ID	Docking score	Glide energy	MMGBSA dG bind	Glide evdw	Glide ecoul	Glide emodel	Residue interactions		
							H-bond	Pi-Pi stacking	Salt bridge
Nafamostat	-6.075	-25.636	-32.661	-16.027	-9.609	-32.630	Asp180, Gly209	—	Asp 180
Camostat	-4.732	-43.818	-56.08	-37.847	-5.971	-56.678	Asp180, Gly209	—	Asp 180
ZINC000013444414	-10.447	-47.167	-59.21	-30.569	-16.598	-64.055	Asp180, Gly184, Gly209	His41	Asp180
ZINC000013444423	-10.38	-48.337	-61.21	-31.781	-16.555	-67.916	Asp180, Gly184, Gly209	His41	Asp180
ZINC000003591050	-10.3	-48.82	-63.19	-28.034	-20.786	-64.646	Asp180, Gly184, Gly209	His41	Asp180
ZINC000013444417	-10.196	-47.467	-63.12	-31.115	-16.352	-62.99	Asp180, Gly184, Gly209	His41	Asp180
ZINC000137976768	-10.101	-57.952	-76.33	-39.566	-18.386	-84.707	Asp180, Gly209, His41	—	Asp180
ZINC000013644383	-10.059	-46.243	-58.64	-28.823	-17.42	-55.185	Asp180, Gly209, Lys135, Asp162	—	Asp180
ZINC000013444408	-10.046	-48.339	-67.21	-29.337	-19.002	-63.549	Asp180, Gly184, Gly209	His41	Asp180
ZINC000013444358	-10.039	-41.761	-50.86	-25.801	-15.96	-61.018	Asp180, Gly184, Gly209	—	Asp180
ZINC000013444405	-9.995	-48.679	-65.89	-30.246	-18.433	-62.942	Asp180, Gly184, Gly209	His41	Asp180
ZINC000013444396	-9.99	-48.3	-64.75	-28.939	-19.361	-63.132	Asp180, Gly184, Gly209	His41	Asp180
ZINC000013444353	-9.989	-40.479	-62.92	-25.828	-14.651	-58.481	Asp180, Gly184, Gly209	—	Asp180
ZINC000143375720	-9.967	-60.737	-74.39	-41.044	-19.692	-87.584	Asp180, Gly209, His41, Lys87	—	—
ZINC000143375943	-9.863	-57.137	-69.63	-39.532	-17.605	-82.619	Asp180, Gly209	His41	Asp180

Otamixaban was docked and reported to have interactions with TMPRSS2 homology model i.e., pi stacking interaction with His 41, salt bridges with Glu 44, Lys 45, and H-bond with between benzamidine and plasma kallikrein in their template structure 2ANY [44, 45]. Similarly, in our analysis

of the docking complexes, we revealed that the ligands form H-bond, Pi-Pi stacking, and salt bridge-type interactions with the active site residues of TMPRSS2 (Table 1 and Fig. 4). The amine groups of benzamidine moiety present in these molecular structures formed H-bond interactions

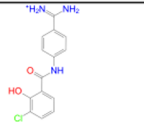
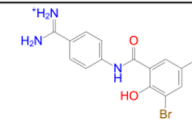
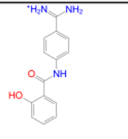
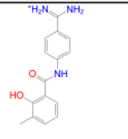
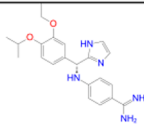
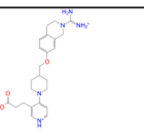
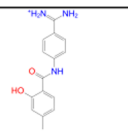
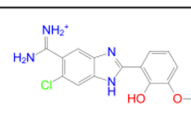
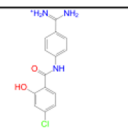
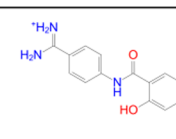
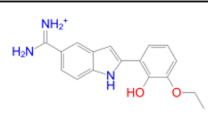
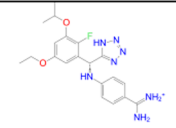
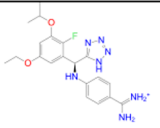
 <p>zinc_id: ZINC000013444414 docking score: -10.447 glide energy: -47.167</p>	 <p>zinc_id: ZINC000013444423 docking score: -10.38 glide energy: -48.337</p>	 <p>zinc_id: ZINC000003591050 docking score: -10.3 glide energy: -48.82</p>	 <p>zinc_id: ZINC000013444417 docking score: -10.196 glide energy: -47.467</p>	 <p>zinc_id: ZINC000137976768 docking score: -10.101 glide energy: -57.952</p>
 <p>zinc_id: ZINC000013644383 docking score: -10.059 glide energy: -46.243</p>	 <p>zinc_id: ZINC000013444408 docking score: -10.046 glide energy: -48.339</p>	 <p>zinc_id: ZINC000013444358 docking score: -10.039 glide energy: -41.761</p>	 <p>zinc_id: ZINC000013444405 docking score: -9.995 glide energy: -48.679</p>	 <p>zinc_id: ZINC000013444396 docking score: -9.99 glide energy: -48.3</p>
 <p>zinc_id: ZINC000013444353 docking score: -9.989 glide energy: -40.479</p>	 <p>zinc_id: ZINC000143375720 docking score: -9.967 glide energy: -60.737</p>	 <p>zinc_id: ZINC000143375943 docking score: -9.863 glide energy: -57.137</p>		

Fig. 3 Structure of TMPRSS2 hits with their zinc IDs, docking scores, and glide energies

with active site residues Asp180 and Gly209. Further, the aromatic hydroxyl groups formed H-bond interactions with Gly184 residue. Most of the hits formed H-bond interactions with the above active site residues. Besides, three-hit molecules show additional H-bond interactions with His41, Lys135, Asp162, and Lys87 residues (Table 1 and Fig. 4). The Pi-Pi stacking interaction between the benzene groups and the His41 residues was observed in all the hits, except five. The salt bridge interaction was observed between the amine groups of benzamidine moiety, and Asp180 residue was found in all the hits excepting ZINC000143375720 (Table 1 and Fig. 4). The above analysis highlights that the benzamidine moiety present in these molecules plays a critical role in establishing H-bond and salt bridge interactions with the active site residues of TMPRSS2.

ADME properties

The in silico ADME analysis results of TMPRSS2 hits are given in Table 2. All the molecules show properties within the permitted limits of Lipinski rules of 5 and Jorgensen's rules of 3. The results, therefore, suggest that all the hit molecules have acceptable ADME properties.

MD simulation

MD simulation of the selected hits was carried out to assess the physical movements of atoms and molecules of the ligand-receptor complex under the physiological conditions to gain insights into the protein–ligand interactions. The MD simulation analysis of ligand-receptor complex of ZINC000013444414 and TMPRSS2 shows a ligand RMSD of 1.5 Å for early the first 20 ns, which later changes to 3.3 Å and becomes stable for the remaining 80 ns without much variation. In contrast, the protein RMSD was steady at around 1.68 Å for the entire period of simulation (Fig. 5). The protein RMSF graph shows that TMPRSS2 residues remain stable during the period of simulation, except for residues in the range of 160–165 (Fig. 5). Protein–ligand contact analysis shows that molecule, ZINC000013444414, interacts with more than 12 active site residues of TMPRSS2. The interaction types observed include H-bonds, hydrophobic, ionic, and water bridges. Among these, a prominent H-bond interaction lasting throughout the simulation period was observed with Asp180, Asp181, and Gly209 residues. Other residues involved in the ligand interaction which lasted less

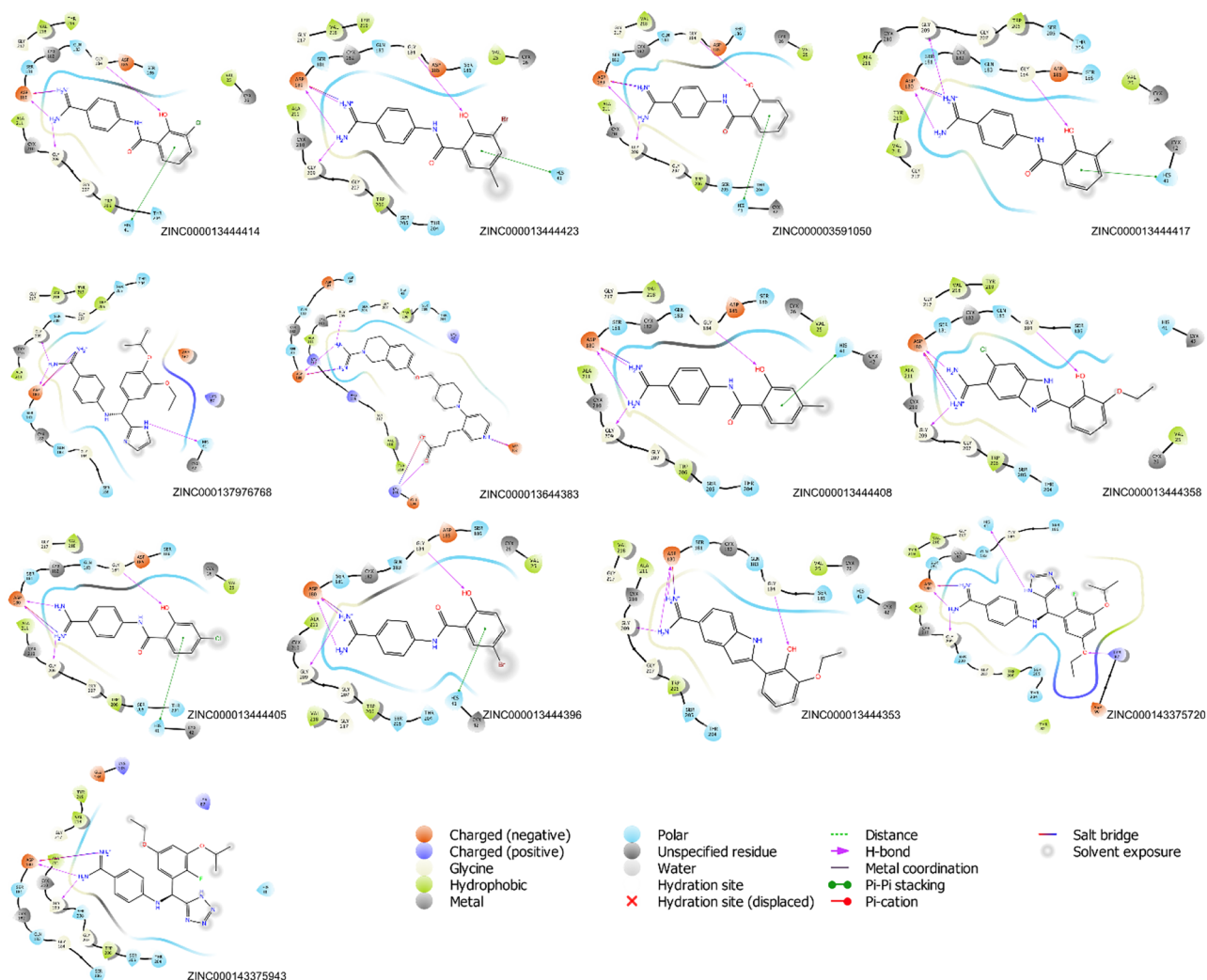


Fig. 4 2D interaction diagrams of TMPRSS2 hits

than 30% of the simulation time include His41, Lys87, Gl134, Cys182, Gln183, Gly184, Ser186, Thr204, Ser20, and Gly207 (Fig. 5).

The MD simulation result analysis of ZINC000137976768 and TMPRSS2 shows a relatively fluctuating ligand RMSD which seems to be stable at around 1.5 Å for the first 20 ns, which later changes to 2.25 Å at 30 ns and somewhat stabilizes between the 2.25 and 3.5 Å during the remaining period of the simulation. In contrast, the protein RMSD seems to be relatively stable at around 2.1 Å after 30 ns (Fig. 6). The protein RMSF graph shows that TMPRSS2 residues remain stable during the period of simulation (Fig. 6). Protein–ligand contact analysis shows a similar pattern of interaction with the active site residues of TMPRSS2. Here also a prominent H-bond interaction lasting throughout the simulation period

was observed with Asp180, Asp181, and Gly209 residues (Fig. 6). Among the three hits, ZINC000143375720 shows a most stable interaction with the TMPRSS2. The RMSD of protein and ligand remains steady at around 1.75 and 3.6 Å, respectively, throughout the simulation (Fig. 7). Further, as observed with the previous two ligands, a similar pattern of protein contact was seen with ZINC000143375720.

Overall, the comparison of MD simulation results of all the hits with that of Glide XP docking suggests a significant correction for ligand interaction with active site residues. The prominent residues commonly observed include Asp180, Gly184, Gly209, and His41. Further, MD simulation also identifies a notable H-bond interaction with Ser181 for all the three hits, which was not recognized during Glide XP docking.

Table 2 Results of ADME property analysis of TMPRSS2 hits

ZINC ID	Stars	Mol. Wt	DHB	AHB	logP	logS	SASA	Volume	PCaco	logBB	Met	HOA	%HOA	Rule of 5	Rule of 3
ZINC000013444414	0	289.721	4	3.75	1.93	-3.405	528.983	880.819	165.623	-1.403	1	3	77.937	0	0
ZINC000013444423	0	348.198	4	3.75	2.3	-4.041	566.184	949.898	169.527	-1.443	2	3	80.304	0	0
ZINC000003591050	0	255.276	4	3.75	1.42	-2.762	507.293	839.163	143.26	-1.596	1	3	73.873	0	0
ZINC000013444417	0	269.302	4	3.75	1.78	-3.231	536.28	895.943	181.044	-1.538	2	3	77.79	0	0
ZINC000137976768	0	393.488	5	5.5	3.83	-5.546	759.253	1322.55	443.928	-1.709	5	3	96.766	0	0
ZINC000013644383	0	437.541	4	7.25	3.72	-6.063	764.798	1395.81	18.931	-2.251	6	2	71.587	0	2
ZINC000013444408	0	269.302	4	3.75	1.72	-3.29	539.891	899.83	144.827	-1.649	2	3	75.685	0	0
ZINC000013444358	0	330.773	5	4.5	2.21	-4.022	590.762	995.549	280.548	-1.292	2	3	83.722	0	0
ZINC000013444405	0	289.721	4	3.75	1.9	-3.463	531.345	883.273	143.326	-1.465	1	3	76.662	0	0
ZINC000013444396	0	334.172	4	3.75	1.97	-3.569	536.328	892.127	143.317	-1.461	1	3	77.088	0	0
ZINC000013444353	0	295.34	5	3	2.42	-3.703	573.77	966.208	341.481	-1.313	2	3	86.447	0	0
ZINC000143375720	1	413.454	5	7	2.57	-5.252	756.123	1312.33	64.705	-2.656	6	2	74.425	0	0
ZINC000143375943	1	413.454	5	7	2.62	-5.289	755.728	1315.3	63.783	-2.629	6	2	74.593	0	0

Optimum values: Stars (number of property or descriptor values that fall outside the 95% range of similar values for known drugs): 0–5; Mol Wt. (molecular weight): 130–725; DHB (hydrogen bond donors): 0–6; AHB (hydrogen bond acceptors): 2–20; logP (predicted octanol/water partition coefficient): 2.0–6.5; SASA (solvent accessible surface area): 300–1000 Å²; volume: 500–2000 Å³; PCaco (predicted apparent Caco-2 cell permeability in nm/s): <25 poor, >500 great; logBB (predicted brain/blood barrier partition coefficient): -3.0 to 1.2; Met (number of likely metabolic reactions): 1–8; HOA (human oral absorption): 1 low, 2 medium, 3 high; % HOA (% human oral absorption): >80% high, <25% low; number of violations of Lipinski's rule of 5 (the rules are: Mol. Wt. <500, logPo/w <5, DHB ≤5, AHB ≤10. Compounds that satisfy these rules are considered drug-like): maximum 4; number of violations of Jorgensen's rule of 3 (the three rules are: QPlogS > -5.7, QPPCaco > 22 nm/s, # Primary Metabolites <7. Compounds with fewer and preferably no violations of these rules are more likely to be orally available: maximum 3. Ref from [46]

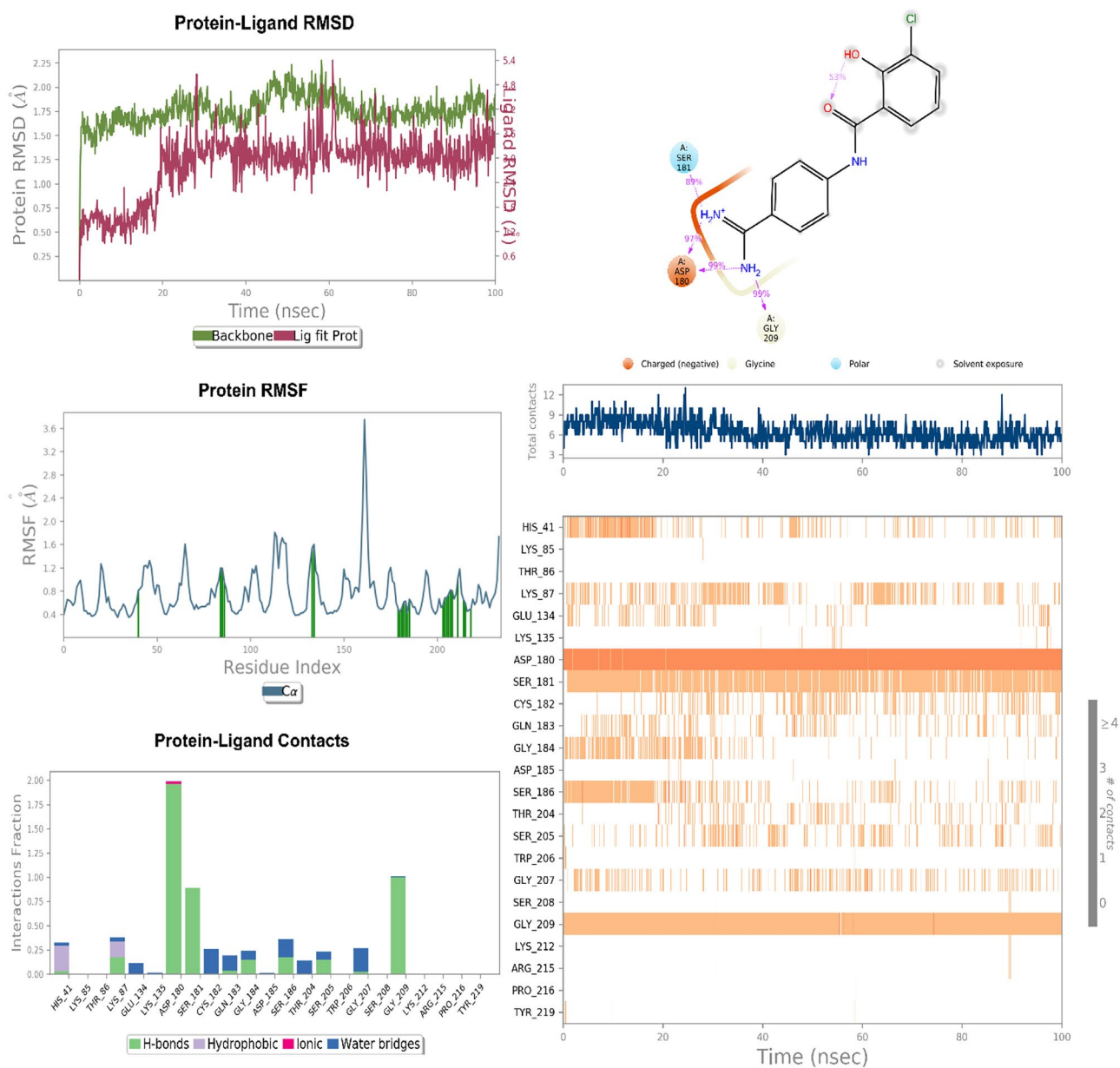


Fig. 5 RMSD, RMSF, and protein–ligand contact diagram of ZINC000013444414 with TMRSS2

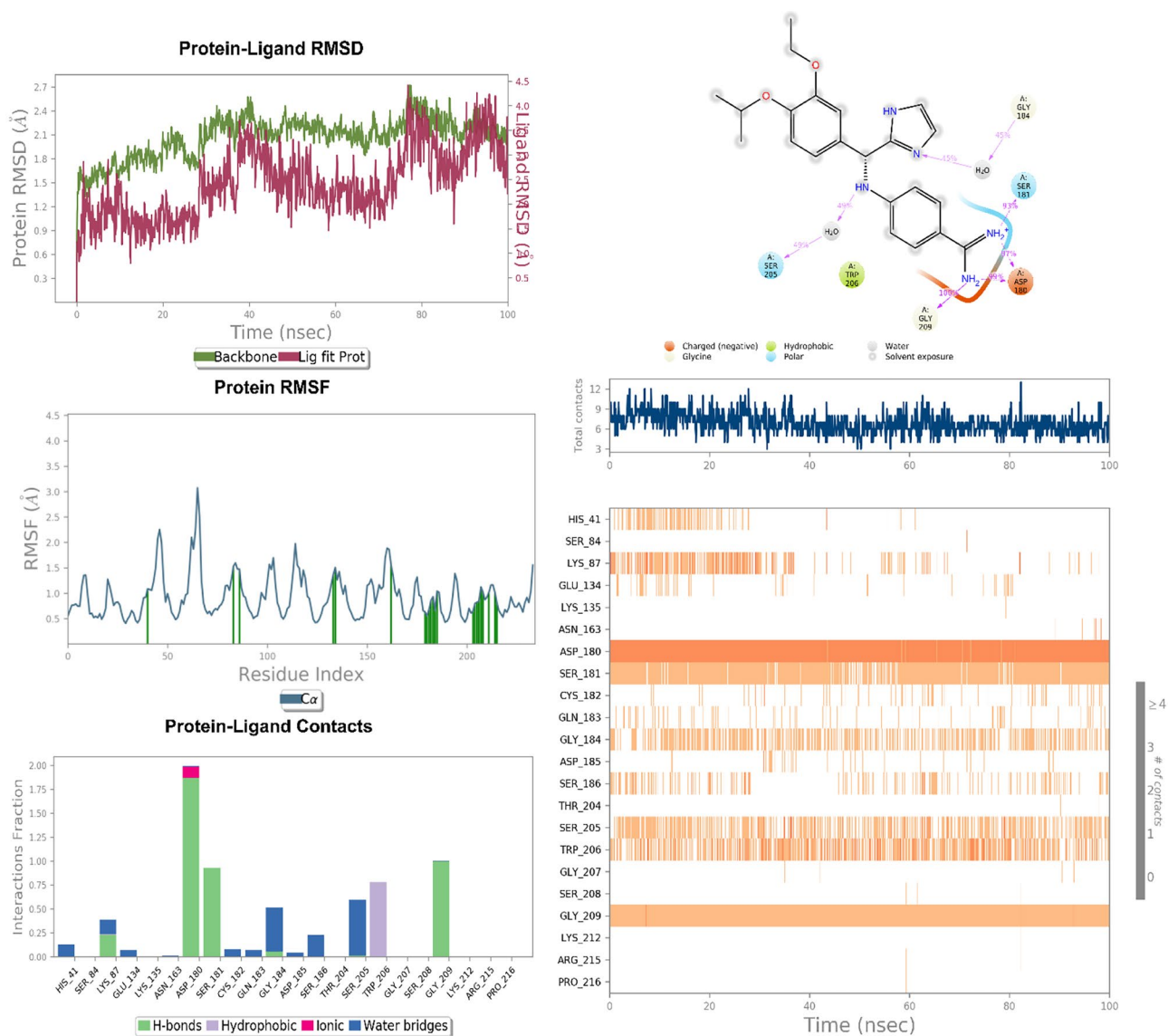


Fig. 6 RMSD, RMSF, and protein–ligand contact diagram of ZINC000137976768 with TMPRSS2

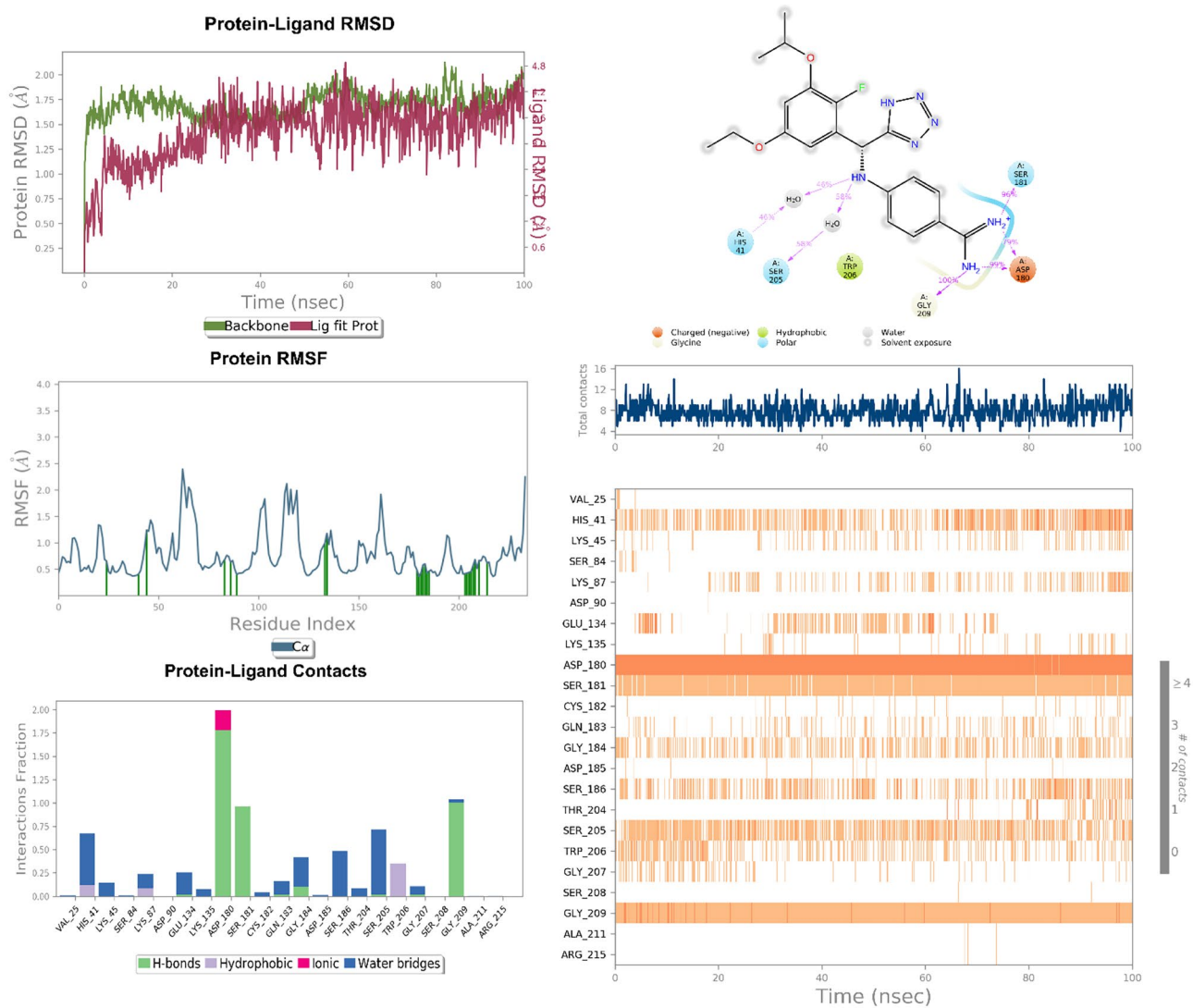


Fig. 7 RMSD, RMSF, and protein–ligand contact diagram of ZINC000143375720 with TMPRSS2

Conclusion

Developing antiviral drugs against SARS-CoV-2 is challenging. Researchers across the globe are working towards developing lead molecules against various target proteins of this virus. TMPRSS2 is one of the vital host target proteins which is well recognized as an essential antiviral drug target against SARS-CoV-2 infection. To date, there are no solved crystal structures of TMPRSS2 available in the protein databank; however, homology structure derived from TMPRSS15 has been useful in the discovery and development of lead molecules against this target. In the current study, structural-based virtual screening and molecular dynamics-based computational research has identified the potential hit molecules from the Zinc database. Further, in vitro and in vivo, studies with these molecules may shed more light on their possible benefits in the treatment of COVID-19.

Acknowledgements The authors acknowledge the infrastructure support provided DST-FIST, Government of India, New Delhi (Grant No. SR/FST/LSI-574/2013) to the Dept. of Pharmacology, JSS College of Pharmacy, Ooty. The authors also thank Manipal – Schrödinger Centre for Molecular Simulations and Manipal College of Pharmaceutical Sciences, Manipal Academy of Higher Education, Manipal, India, for providing facility to carry out this research work.

Author contribution Conceptualization, K Sreedhara Ranganath Pai, Praveen T Krishnamurthy. Investigation, Suman Manandhar, Ravi Kiran Ammu V V V, Kusuma Kumari Garikapati. Supervision, K Sreedhara Ranganath Pai, Praveen T Krishnamurthy. Writing — original draft — Suman Manandhar, Ravi Kiran Ammu V V V, Kusuma Kumari Garikapati. Writing — review and editing — K Sreedhara Ranganath Pai, Praveen T Krishnamurthy.

Availability of data and material The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare no competing interests.

References

- Guo YR, Cao QD, Hong ZS, Tan YY, Chen SD, Jin HJ et al (2020) The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak - an update on the status. *Mil Med Res* 7(1):11
- Li F (2016) Structure, function, and evolution of coronavirus spike proteins. *Annu Rev Virol* 3(1):237–261
- Coronavirus Disease (2022) (COVID-19) Situation reports. Available from: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>
- CoVID-19 Clinical trials (2020) 20 Apr 2020. Available from: <https://clinicaltrials.gov/ct2/results?cond=COVID-19>
- Wu C-Y, Lin Y-S, Yang Y-H, Shu L-H, Cheng Y-C, Liu HT (2020) GB-2 inhibits ACE2 and TMPRSS2 expression: in vivo and in vitro studies. *Biomed Pharmacother* 132:110816
- Li K, Meyerholz David K, Bartlett Jennifer A, McCray Paul B, Frieman Matthew B, Griffin Diane E (2021) The TMPRSS2 inhibitor nafamostat reduces SARS-CoV-2 pulmonary infection in mouse models of COVID-19. *mBio* 12(4):e00970–21
- Sun YJ, Velez G, Parsons DE, Li K, Ortiz ME, Sharma S et al (2021) Structure-based phylogeny identifies avoralstat as a TMPRSS2 inhibitor that prevents SARS-CoV-2 infection in mice. *J Clin Investig* 131(10)
- Ianevski A, Yao R, Lysvand H, Grødeland G, Legrand N, Oksenyh V et al (2021) Nafamostat–interferon- α combination suppresses SARS-CoV-2 infection in vitro and in vivo by cooperatively targeting host TMPRSS2. *Viruses* 13(9)
- Shirato K, Kawase M, Matsuyama S (2018) Wild-type human coronaviruses prefer cell-surface TMPRSS2 to endosomal cathepsins for cell entry. *Virology* 517:9–15
- Fehr AR, Perlman S (2015) Coronaviruses: an overview of their replication and pathogenesis. *Methods Mol Biol* 1282:1–23
- Chu H, Chan JF, Wang Y, Yuen TT, Chai Y, Hou Y et al (2020) Comparative replication and immune activation profiles of SARS-CoV-2 and SARS-CoV in human lungs: an ex vivo study with implications for the pathogenesis of COVID-19. *Clin Infect Dis*
- Fung TS, Liu DX (2014) Coronavirus infection, ER stress, apoptosis and innate immunity. *Front Microbiol* 5:296
- Sanders JM, Monogue ML, Jodlowski TZ, Cutrell JB (2020) Pharmacologic treatments for coronavirus disease 2019 (COVID-19): a review. *JAMA*
- Shang J, Ye G, Shi K, Wan Y, Luo C, Aihara H et al (2020) Structural basis of receptor recognition by SARS-CoV-2. *Nature*
- Teli DM, Shah MB, Chhabria MT (2021) In silico screening of natural compounds as potential inhibitors of SARS-CoV-2 main protease and spike RBD: Targets for COVID-19. *Front Mol Biosci* 7:429
- Santibáñez-Morán MG, López-López E, Prieto-Martínez FD, Sánchez-Cruz N, Medina-Franco JL (2020) Consensus virtual screening of dark chemical matter and food chemicals uncover potential inhibitors of SARS-CoV-2 main protease. *RSC Adv* 10(42):25089–25099
- Bharti R, Shukla SK (2021) Molecules against covid-19: an in silico approach for drug development. *J Electron Sci Technol* 19(1):100095
- Gupta RK, Nwachuku EL, Zusman BE, Jha RM, Puccio AM (2021) Drug repurposing for COVID-19 based on an integrative meta-analysis of SARS-CoV-2 induced gene signature in human airway epithelium. *PLoS One* 16(9):e0257784
- Cavasotto CN, Di Filippo JJ (2021) In silico Drug repurposing for COVID-19: targeting SARS-CoV-2 proteins through docking and consensus ranking. *Mol Inf* 40(1):2000115
- Kumar BH, Manandhar S, Mehta CH, Nayak UY, Pai KSR (2022) Structure-based docking, pharmacokinetic evaluation, and molecular dynamics-guided evaluation of traditional formulation against SARS-CoV-2 spike protein receptor bind domain and ACE2 receptor complex. *Chem Pap* 76(2):1063–83
- Zhang D-h, Wu K-l, Zhang X, Deng S-q, Peng B (2020) In silico screening of Chinese herbal medicines with the potential to directly inhibit 2019 novel coronavirus. *J Integr Med* 18(2):152–8
- Elmezayen AD, Al-Obaidi A, Şahin AT, Yeleği K (2021) Drug repurposing for coronavirus (COVID-19): in silico screening of known drugs against coronavirus 3CL hydrolase and protease enzymes. *J Biomol Struct Dyn* 39(8):2980–2992
- Hosseini M, Chen W, Xiao D, Wang C (2021) Computational molecular docking and virtual screening revealed promising SARS-CoV-2 drugs. *Precis Clin Med* 4(1):1–16
- Tejera E, Munteanu CR, López-Cortés A, Cabrera-Andrade A, Pérez-Castillo Y (2020) Drugs repurposing using QSAR, docking and molecular dynamics for possible inhibitors of the SARS-CoV-2 Mpro protease. *Molecules* 25(21)

25. Choudhary S, Malik YS, Tomar S (2020) Identification of SARS-CoV-2 cell entry inhibitors by drug repurposing using in silico structure-based virtual screening approach. *Front Immunol* 11
26. Sakai K, Ami Y, Tahara M, Kubota T, Anraku M, Abe M et al (2014) The host protease TMPRSS2 plays a major role in in vivo replication of emerging H7N9 and seasonal influenza viruses. *J Virol* 88(10):5608–5616
27. Limburg H, Harbig A, Bestle D, Stein DA, Moulton HM, Jaeger J et al (2019) TMPRSS2 is the major activating protease of influenza A virus in primary human airway cells and influenza B virus in human type II pneumocytes. *J Virol* 93(21)
28. Iwata-Yoshikawa N, Okamura T, Shimizu Y, Hasegawa H, Takeda M, Nagata N (2019) TMPRSS2 contributes to virus spread and immunopathology in the airways of murine models after coronavirus infection. *J Virol* 93(6)
29. Abe M, Tahara M, Sakai K, Yamaguchi H, Kanou K, Shirato K et al (2013) TMPRSS2 is an activating protease for respiratory parainfluenza viruses. *J Virol* 87(21):11930–11935
30. Hoffmann M, Kleine-Weber H, Schroeder S, Kruger N, Herrler T, Erichsen S et al (2020) SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* 181(2):271–80 e8
31. Matsuyama S, Nao N, Shirato K, Kawase M, Saito S, Takayama I et al (2020) Enhanced isolation of SARS-CoV-2 by TMPRSS2-expressing cells. *Proc Natl Acad Sci U S A* 117(13):7001–7003
32. Lukassen S, Chua RL, Trefzer T, Kahn NC, Schneider MA, Muley T et al (2020) SARS-CoV-2 receptor ACE2 and TMPRSS2 are primarily expressed in bronchial transient secretory cells. *EMBO J* e105114
33. Stopsack KH, Mucci LA, Antonarakis ES, Nelson PS, Kantoff PW (2020) TMPRSS2 and COVID-19: serendipity or opportunity for intervention?. *Cancer Discov*
34. Huggins DJ (2020) Structural analysis of experimental drugs binding to the SARS-CoV-2 target TMPRSS2. *J Mol Graph Model* 100:107710
35. Coutard B, Valle C, de Lamballerie X, Canard B, Seidah NG, Decroly E (2020) The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade. *Antiviral Res* 176:104742
36. Jaimes JA, Andre NM, Chappie JS, Millet JK, Whittaker GR (2020) Phylogenetic analysis and structural modeling of SARS-CoV-2 spike protein reveals an evolutionary distinct and proteolytically sensitive activation Loop. *J Mol Biol* 432(10):3309–3325
37. Phillips R (2014) Prostate cancer: TMPRSS2 promotes metastasis through proteolysis. *Nat Rev Urol* 11(10):546
38. Clyne M (2013) Prostate cancer: TMPRSS2:ERG—the root of the problem?. *Nat Rev Urol* 10(5):248
39. Wang Z, Wang Y, Zhang J, Hu Q, Zhi F, Zhang S et al (2017) Significance of the TMPRSS2:ERG gene fusion in prostate cancer. *Mol Med Rep* 16(4):5450–5458
40. Vaarala MH, Porvari KS, Kellokumpu S, Kyllonen AP, Vihko PT (2001) Expression of transmembrane serine protease TMPRSS2 in mouse and human tissues. *J Pathol* 193(1):134–140
41. Hatesuer B, Bertram S, Mehnert N, Bahgat MM, Nelson PS, Pohlmann S et al (2013) Tmprss2 is essential for influenza H1N1 virus pathogenesis in mice. *PLoS Pathog* 9(12):e1003774
42. Lambertz RLO, Gerhauser I, Nehlmeier I, Leist SR, Kollmus H, Pohlmann S et al (2019) Tmprss2 knock-out mice are resistant to H10 influenza A virus pathogenesis. *J Gen Virol* 100(7):1073–1078
43. Sterling T, Irwin JJ (2015) Zinc 15–ligand discovery for everyone. *J Chem Inf Model* 55(11):2324–2337
44. Rensi S, Altman RB, Liu T, Lo Y-C, McInnes G, Derry A et al (2020) Homology modeling of TMPRSS2 yields candidate drugs that may inhibit entry of SARS-CoV-2 into human cells. *ChemRxiv*. <https://doi.org/10.26434/chemrxiv.12009582.v1>
45. Hu X, Shrimp JH, Guo H, Xu M, Chen CZ, Zhu W et al (2021) Discovery of TMPRSS2 inhibitors from virtual screening as a potential treatment of COVID-19. *ACS Pharmacol Trans Sci* 4(3):1124–1135
46. Schrödinger Release (2021) QikProp, Schrödinger, LLC, New York, NY

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.