JOURNAL OF BIOMOLECULAR STRUCTURE AND DYNAMICS
https://doi.org/10.1080/07391102.2020.1868335

Taylor & Francis
Taylor & Francis Group

Fisetin 8-C-glucoside as entry inhibitor in SARS CoV-2 infection: molecular modelling study

Abha Mishra[a], Upinder Kaur[b] and Amit Singh[c]

[a]School of Biochemical Engineering, Indian Institute of Technology, Banaras Hindu University, Varanasi, India; [b]Department of Pharmacology, All India Institute of Medical Sciences, Gorakhpur, India; [c]Department of Pharmacology, Institute of Medical Sciences, Banaras Hindu University, Varanasi, India

Communicated by Ramaswamy H. Sarma

ABSTRACT

Coronaviruses are RNA viruses that infect varied species including humans. TMPRSS2 is gateway for SARS CoV-2 entry into the host cell. It causes proteolytic activation of spike protein and discharge of the peptide into host cell. The TMPRSS2 inhibition could be one of the approaches to stop the viral entry, therefore, interaction pattern and binding energies for Fisetin and TMPRSS2 have been explored in the present study. TMPRSS2 peptide was used for homology modelling and then for further study. Molecular docking score and MMGBSA Binding energy of Fisetin was better than Nafamostat, a known inhibitor of TMPRSS2. Post docking MM-GBSA free energy for Fisetin and Nafamostat was −42.78 and −21.11 kcal/mol, respectively. Fisetin forms H bond with Val 25, His 41, Lys 42, Lys 45, Glu 44, Ser186. Nafamostat formed H bonds with Lys 85, Asp 90, Asp 203. RMSD plots of TMPRSS2, TMPRSS2-Fisetin and TMPRSS2-Nafamostat complex showed stable profile with very small fluctuation during entire simulation of 150 ns. Significant decrease in TMPRSS2-Fisetin and TMPRSS2-Nafamostat complex fluctuation occurred around His 41, Glu 44, Gly 136, Ser 186 in RMSF study. During simulation Fisetin interaction was observed with residues Val 25, His 41, Glu 44, Lys 45, Lys 87, Gly 136, Gln 183, Ser 186 likewise interaction of Nafamostat with Lys 85, Asp 90, Asn 163, Asp 203 and Ser 205. Post simulation MM-GBSA free energy was found to be −51.87 ± 4.3 and −48.23 ± 4.39 kcal/mol for TMPRSS2 with Fisetin and Nafamostat, respectively.

ARTICLE HISTORY
Received 3 June 2020
Accepted 18 December 2020

KEYWORDS
TMPRSS2; homology modelling; Fisetin; MD simulation; MMGBSA

1. Introduction

Recent outbreak of Corona virus SARS CoV-2 infection has created a serious health problem to the whole world. There is urgent need to understand the exact mechanism of action of virus on human system and to develop a specific drug as no specific line of treatment is currently available. SARS-CoV-2 is a single-stranded RNA-enveloped virus, its spike (S) protein binds to the host angiotensin-converting enzyme 2 (ACE2) and primed by trans membrane serine protease, TMPRSS2 (Transmembrane serine protease 2), that leads to entry of virus into the cell. TMPRSS2 plays a role in the proteolytic activation of viral S protein and helps in respiratory disease caused by influenza (Böttcher et al., 2006), SARS-CoV and MERS (Iwata-Yoshikawa et al., 2019) viruses. TMPRSS2 may also regulate SARS-CoV-2 assembly and release from the cell membrane (Alireza et al., 2020; Shen et al., 2017). After entry into the cells there is synthesis of polyproteins and RNA with the help of replicase-transcriptase and RNA-dependent RNA polymerase. These proteins like host TMPRSS2, ACE2 and viral proteins like M^{pro} Protease and RNA-dependent RNA polymerase can be a potential drug targets for SARS-CoV-2 (Glowacka et al., 2011; Hoffmann et al., 2020; Shen et al., 2017; Shulla et al., 2011; Singh & Mishra, 2020; Sinha et al., 2020).

TMPRSS2 consists of 492 amino acids which is attached to the cell membrane. The mature proteases after autocatalytic breakdown between Arg 255 and Ile 256 are generally membrane-bound. TMPRSS2 catalytic domain contains a catalytic triad consisting of the amino acid residues His 296, Asp 345 and Ser 441 (Lin et al., 1999; Park, 2010). The exact physiological function of TMPRSS2 in *vivo* remains unknown. TMPRSS2 protein is highly expressed in bronchial epithelial cells than in type II pneumocytes (Bertram et al., 2012).

There are mainly three classes of antiviral drugs under clinical trial for the effectiveness and safety against COVID-19 infection. First is RNA dependent RNA Polymerase Inhibitors (Remdesivir, Favipiravir, Ribavirin) and second is Protease Inhibitors (Lopinavir, Ritonavir) and various miscellaneous drugs like Interferons, Anti-cytokines, hydroxychloroquine comes in third group. Bromhexine is found to be an inhibitor of the TMPRSS2 with an IC50 of 0.75 μM (Lucas et al., 2014). The Camostat, Nafamostat, Aprotinine and bromhexine although not developed specifically for targeting TMPRSS2 were reported to inhibit TMPRSS2 protease activity (Lucas et al., 2014). Secondary metabolites, phytochemicals have therapeutic potential which has been explored in the present study against TMPRSS2 [Transmembrane serine protease 2] to prevent SARS CoV-2 infection. Ligand we selected for

CONTACT Amit Singh amisingh@bhu.ac.in Department of Pharmacology, Institute of Medical Sciences, Banaras Hindu University, Varanasi, India.

present study is a flavanol Fisetin 8-C-glucoside (Fisetin), generally present in fruits and vegetables such as apple, onion, strawberry, grape, cucumber etc. and are known to possess anticancer, antioxidant activities (Khan et al., 2013). Present study was initiated to understand how viral entry into target cells can be blocked by use of phytochemicals. Nafamostat, protease inhibitor, was taken for comparative study.

2. Method

Homology modelling method used for building a three-dimensional model of the TMPRSS2 by comparing its amino acid sequence with an experimental model of related homologous template protein. The primary sequence of the target human TMPRSS2 was obtained from UniProtKB database with a sequence id: O15393, entry name TMPS2_HUMAN sequence length 492aa (UniProt). The amino acid sequence of TMPRSS2 subjected to BLASTp program with default parameters in Uniprot to find out template for sequence alignment (Altschul et al., 1990; Jakubík et al., 2013). Peptide chain which consists of catalytic domain, from 256 to 492, was used for modelling. So, in our model Amino acid 1 corresponds to amino acid 256. The model was generated using Prime in Schrödinger Suite and loop refinement was done. To compare the model iTASSER model prediction tool was also used for model preparation. The validation of Target model obtained from Prime was performed by analysing the psi (ψ) and phi (φ) torsion angles using Ramachandran plot. The 3D structure of target model obtained from homology modelling was optimized by hydrogen bond assignments (Ramachandran et al., 2016). PROCHECK was used to assess the quality of the model structure (Laskowski et al., 1993). The modelled structure was further minimized using 10 ns MD simulation using OPLS3 force field as per default protocol.

2.1. Active site prediction

Schrödinger suite was used to predict receptor sites of TMPRSS2 structure by using Sitemap module. It had generated potential site in the receptors with quantitative site-score values for each site.

2.2. Molecular docking

Protein preparation wizard of Schrödinger suite was used to optimize TMPRSS2. Protonation of the TMPRSS2 at the physiological pH performed by Epik module of Schrödinger suite. Water molecules greater than 3 Å away from the Fisetin (Fisetin 8-C-glucoside:Pub chem i.d. 5281615) were deleted before subjecting to minimization and hydrogen bond optimization. Impact refinement module utilizing the OPLS3e force field for restrained molecular minimization to make relaxed structure after deleting all water molecules, heteroatom and adding hydrogen atoms. OPLS3e force field used in minimization studies to relieve steric clashes. LigPrep module was used for ligands preparation. Energy minimisation (in gas phase) was done using Macromodel with the OPLS3e force field. Glide software was used to carry out molecular docking and grid box was generated around catalytic triad His 41, Asp 90 and Serine 186 residues as centroid (Halgren, 2007). The docking was carried out using extra precision (XP) mode with default parameters to analyze the binding modes of compounds under study (Glide, 2015). The molecular docking generated H-bond between ligand and target, and the consequent glide energy. We have used 250 phytochemicals compiled from published literature with antimicrobial properties. Compounds were ranked based on GlideScore and the best pose of the ligand and chosen for further study (Friesner et al., 2004, 2006; Halgren et al., 2004; Jacobson et al., 2004; Kumar et al., 2016; Sastry et al., 2013).

2.3. Molecular dynamics simulation

Desmond module of Schrodinger Suite used for Molecular dynamics (MD) of ligand protein complex. The Fisetin-TMPRSS2 complex was solvated using SPC aqueous solvation of orthorhombic solvent box. Solvent buffer extended 10 Å beyond the TMPRSS2 in all directions and OPLS3 force field was applied. System was neutralized by adding the required number of counter ions. Initial two stages with steepest descent minimization comprised of 2000 steps with and without a restraint of 50 kcal/mol/Å^2 on the solute atoms carried out for relaxation protocol before production runs (Harder et al., 2016). The MD simulations done for short four times with 12 ps in NVT ensemble at 10 K with solute heavy atoms restrained with a force constant of 50 kcal/mol/Å^2. NPT ensemble at10 K with the same restraint used for 12 ps simulation further NPT ensemble at 300 K with the same restraint for12 ps and NPT ensemble at 300 K without any restraints used finally for 24 ps simulation. Simulation with NPT ensemble at 300 K temperature and 1.01325 bar pressure run under unrestrained condition (Nosé-Hoover chain thermostat relaxation time of 1 ps and the isotropic Martyna-Tobias-Klein barostat with a relaxation time of 2 ps). Short range interactions studied with 9 Å cut off and long-range coulombic interactions analysed with smooth particle mesh Ewald method (PME). r-RESPA integrator used for non-bonded forces evaluation. After every three steps long range forces and at every step short-range force were recorded. MD simulations run for 150 ns. At every 150 ps trajectories were recorded and energy recording interval was maintained at 1.2 ps. OPLS3 (Optimized Kanhesia for Liquid Simulations) force fields used for analysing RMSD and RMSF of the complex structure. At every 150.0 ps intervals trajectories were recorded for evaluation. Molecular dynamic simulation study was done with Nafamostat and TMPRSS2 complex by the same protocol (MacroModel, 2015; Singh & Mishra, 2020).

Binding free energy was studied with Prime-MM/GBSA (Molecular Mechanics/Generalized Born Surface Area) (Mobley & Dill, 2009; Prime, 2015). Surface GB model with Gaussian surface in place of van der Waals surface used for better depiction of the solvent accessible surface area (Jakubík et al., 2013). Equations used for binding energy (ΔG_{bind}) calculations were as follows:

$$\Delta G_{bind} = \Delta E + \Delta G_{solv} + \Delta G_{SA},$$

$$\Delta E = E_{complex} - E_{protein} - E_{ligand},$$

where $E_{complex}$, $E_{protein}$ and E_{ligand} representing minimized energies of the complex, TMPRSS2 and Fisetin, respectively (Das et al. 2009).

$$\Delta G_{solv} = G_{solv-complex} - G_{solv-protein} - G_{solv-ligand}$$

where solvation free energies were $G_{solv\text{-}complex}$, $G_{solv\text{-}protein}$ and $G_{solv\text{-}ligand}$ for complex, TMPRSS2, and Fisetin, respectively.

The surface area energies were calculated for the complex, TMPRSS2, and Fisetin, by $G_{SA\text{-}complex}$, $G_{SA\text{-}protein}$ and $G_{SA\text{-}ligand}$, respectively;

$$\Delta G_{SA} = G_{SA-complex} - G_{SA-protein} - G_{SA-ligand}$$

The scoring and interaction parameters of XP docking was used as selection parameter for selection of compound for MD simulation study (Kevin et al., 2006; Shivakumar et al., 2010; Wang et al., 2015). Prime program of Schrödinger software suite employed for all docking poses which further rescored with the MM/GBSA approach to allow more flexibility to protein and assess ligand affinity thereby providing more reliable results. The variable dielectric solvent model VSGB 2.0 was used for the study which comprises empirical corrections for modelling directionality of hydrogen-bond and π-stacking interactions. This method gives significant binding free energies for a wide range of protein–ligand complexes (Li et al., 2011; Mulakala & Viswanadhan, 2013). MM minimisation of the complex was carried out by allowing residues within 5.0 Å of the ligand to relax and keeping the rest of the structure fixed (Srivastava & Sastry, 2012).

3. Results and discussion

SARS CoV-2 entry into host cells involves two important steps like binding to host cell receptor and entry of viral genetic material into the host cell. Viral entry requires a main protease (TMPRSS2) which causes proteolytic processing of viral Spike (S) protein. The TMPRSS2 could be an important target to prevent viral entry into host cells. Hoffman et al. (2020) in an *in vitro* study showed that Camostat mesylate, an inhibitor of TMPRSS2, partially inhibit the entry of SARS-CoV-2 into primary pulmonary cells and cell lines. In another study it was reported that Camostat blocks the entry of SARS-CoV in human Calu-3 airway epithelial cells (Kawase et al., 2012). Serine protease inhibitor such as Nafamostat, inhibits TMPRSS2 and resists S-protein facilitated membrane fusion of MERS-CoV (Yamamoto et al., 2016). Bromhexine also reported to be an inhibitor of the TMPRSS2 with an IC-50 of 0.75 μM (Lucas et al., 2014). Inhibition of TMPRSS2 by Nafamostat was reported at 10-fold lower concentration than Camostat (Sielaff et al., 2011; Hoffman et al., 2020). There are reports on several other small molecule which inhibits TMPRSS2 with low nanomolar affinity (Yang & Honig, 2000). We tried to explore interaction pattern of some phytochemicals against TMPRSS2 and results were compared with Nafamostat.

3.1. Homology modelling (HM)

The three-dimensional structure of TMPRSS2 was not available in protein data bank (PDB) so we tried HM for Model preparation. With a similarity of 36.9% the sequence of *Oryzias latipes enteropeptidase light chain* (PDB ID: 3w94) has shown as best homology in the BLAST search for Human TMPRSS2 amino acid sequence. If the sequence similarity between two proteins is higher than 25% then their structures are considered to be similar (Rost, 1999; Yang & Honig, 2000). Therefore, the 3D structure of Human TMPRSS 2 built through homology modelling using the 3D structure of *Oryzias latipes* enteropeptidase light chain as the template for modelling TMPRSS2 protein. Model accuracy is related to the degree of sequence identity and similarity between template and TMPRSS2 (target). The crystal structure of *Oryzias latipes* enteropeptidase light chain was retrieved from PDB (PDB ID: 3w94) (Xu et al., 2014). PROCHECK analysis in Figure 1(a), showed that 81.6% of the residues are in the most favoured regions, 16.3% in additional allowed regions, 2% residues in generously allowed regions and 0.0% of the residue were obtained in disallowed regions (Laskowski et al., 1993). PROCHECK analysis of the model prepared by i-Tasser showed that 76.4% of the residues are in the most favoured regions, 16.6% in additional allowed regions, 3.5% residues in generously allowed regions and 3.5% of the residue were obtained in disallowed regions (Figure 1b).

3.2. Active site prediction

Active site residues of TMPRSS2 were predicted by Site Map. It identified the potential sites with residues with a site score of 1.06 considering parameters which include volume, size, contact, amino acid exposure, hydrophilicity, hydrophobicity and donor/acceptor ratio. The sites with site score of 1 and above are consider as suitable site for the ligand binding (Xie .& Huang 2012). The site included His 41 and Asp 90 also. The protease catalytic domain of TMPRSS2 contains a catalytic triad consisting of the amino acid residues His 296, Asp 345 and Ser 441 (Lin et al., 1999; Park, 2010). Peptide chain which consists of catalytic domain, from 256 to 492, was used for modelling. So, in our model Amino acid 1 correspond to amino acid 256 [His 41 correspond to His 296, Asp 90 correspond to Asp 345, Ser 186 correspond to Ser 441] (Kumar et al., 2020).

3.3. Molecular docking

An attempt was made in the present study to find inhibitor of TMPRSS2 through molecular docking and simulations. The receptor grid was formed around His 41, Asp 90 and Serine 186 in homology model of TMPRSS2 to conduct molecular docking. 250 phytochemicals with antimicrobial properties were chosen for docking against TMPRSS2. The phytochemicals were analysed by glide score, interaction with particular residue, His 41, Asp 90 and Ser 186 which forms catalytic triad of TMPRSS2 and Post docking MM-GBSA binding energies. We used a known inhibitor Nafamostat for comparative

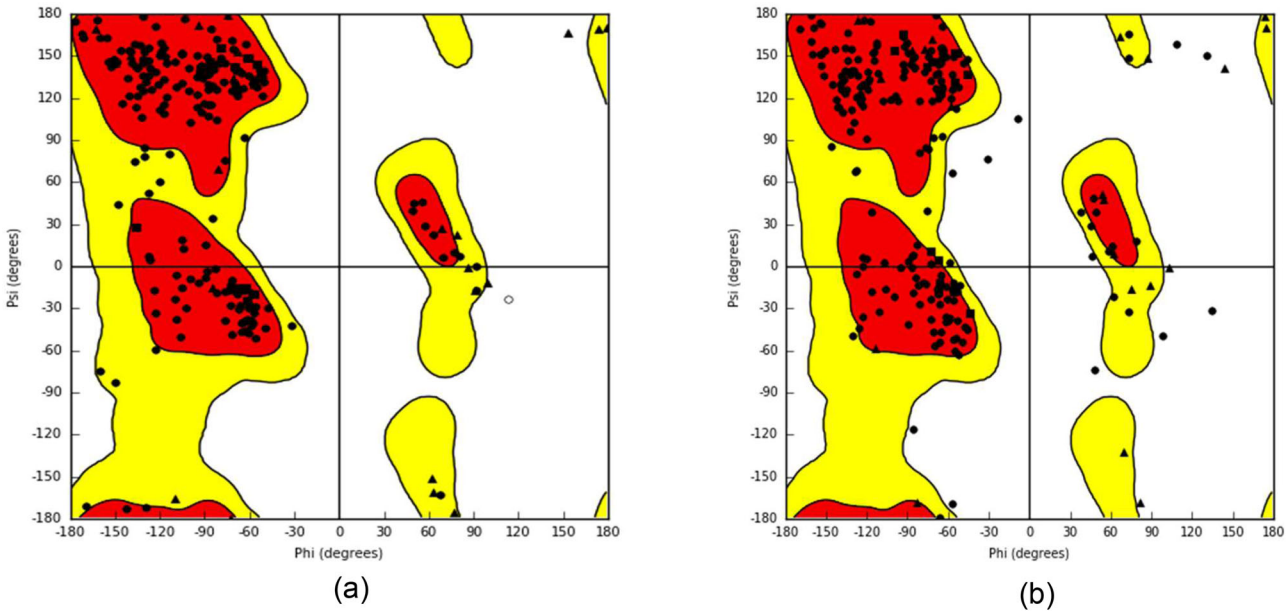

Figure 1. (a) Ramchandran plot of modelled TMPRSS2 protein (target protein) ′. (b) Ramachandran plot of modelled TMPRSS2 protein (target protein) [i-tasser].

Table 1. Molecular docking studies of TMPRSS2 with different phytochemicals.

S.No.	Compounds	XP (kcal/mol)	MMGBSA (kcal/mol)
1	Nafamostat	−6.99	−21.11
2	Fisetin 8-C-glucoside	−8.073	−42.78
3	Isorhamnetin	−7.618	−40.31
4	Glucoiberverin	−6.862	−42.67
5	Castanospermine	−6.044	−12.8
6	Allixin	−5.71	−35.14
7	Astragalin	−5.239	−20.87
8	Trillin	−5.101	−34.32
9	Maclurin	−5.094	−26.38
10	Swainsonine	−5.04	−23.2

study. Out of 250 compound we identified 10 best phytochemicals from Extra precision (XP) docking and MMGBSA study (Table 1). Fisetin showed a good interaction at catalytic domain by forming hydrogen bonds and other non-bonded interactions with TMPRSS2 and lowest MM-GBSA binding energy. The Hydrogen bond interaction occurred at His 41, Ser 186 and with several other amino acids. His 41, Ser 186 are part of catalytic triad of TMPRSS2. Hydrophobic interactions were also seen with Cys 42, Tyr 82, Trp 206 (Figure 2a). The mode of binding and interaction of residues were similar to Nafamostat (Figure 2b). The Fisetin had produced much better MM-GBSA binding energies than Nafamostat. The post docking MM-GBSA binding free energy with TMPRSS2 was found to be for Fisetin and Nafamostat as −42.78 and −21.11 (kcal/mol), respectively (Table 1). Considering all factors, Fisetin was our choice phytochemical for molecular dynamic simulation study.

3.4. Molecular dynamics simulation

MD simulations of the TMPRSS2 complexed with Fisetin were performed using Desmond. It evaluates the dynamic interactions and analyse structural stability and binding site adaptations to the docked Fisetin during 150 ns simulation. MD simulations were run for the TMPRSS2 and Fisetin for 150 ns. The stability and fluctuations of the TMPRSS2 and Fisetin complex was investigated using carbon alpha (Cα) atoms and ligand during 150 ns simulation. RMSD is a measure of protein conformational stability. It can be observed that how fast and how far the protein deviates from the starting structure for a more quantitative analysis of the protein fluctuations. All protein frames are first aligned on the backbone of the reference frame, and then the RMSD was calculated on the basis of the selection of the atom. RMSD plot of the complex showed some fluctuations in the beginning then attained equilibrium at 25 ns and remained stable for next 125 ns, i.e. fluctuation remains within 1 Å. Changes of the order of 1–3 Å are considered satisfactory for small, globular proteins (Figure 3).

Structural stability of protein gives an indication of the shape of the molecule at each time during entire simulation in radius of gyration plot (Rg). It was calculated to analyse structural changes of TMPRSS2, when the Fisetin was bound. The plot of radius of gyration in simulation time in different condition of the TMPRSS2 has given in Figure 4. TMPRSS2-Fisetin complex during 150 ns molecular dynamics simulation did not bring significant conformational changes in its structure.

Figure 5(a) and 5(b) showed H-bond formation between TMPRSS2-Fisetin and TMPRSS2-Nafamostat, respectively, during 150 ns MD simulation study. The average H-bond between TMPRSS2 and Fisetin was 3.04 while 2.67 found with TMPRSS2 and Nafamostat. Upon general hydrogen bond analysis throughout 150 ns simulation, Fisetin was found to form more than 76% hydrogen bonds with Glu44, Lys45 and Lys87. The average distance was also less than 3 Å signifies moderate type of binding (Table 2).

The residues involved in the interactions with a Fisetin during simulation were depicted in Root mean square fluctuations (RMSF) (Figure 6). Root mean square fluctuations

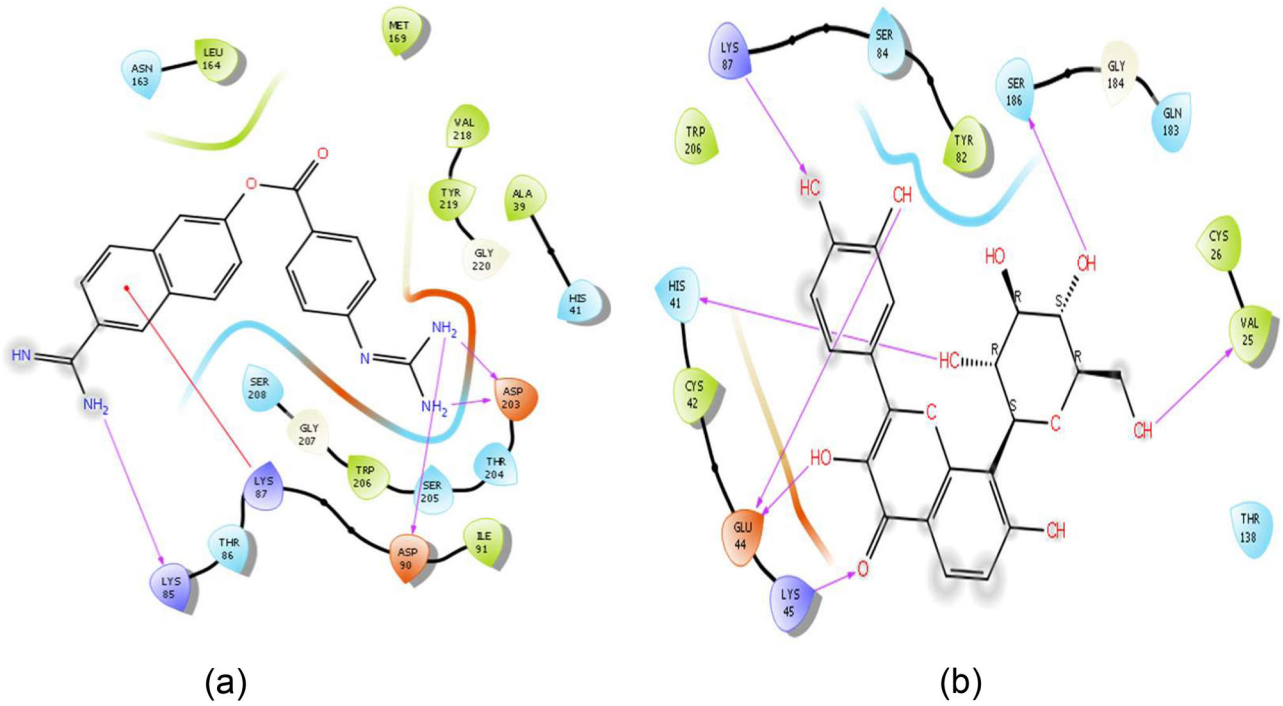

Figure 2. (a) TMPRSS2 with Nafamostat (amino acid 1 corresponds to amino acid 256). (b) TMPRSS2 with Fisetin 8-C-glucoside (amino acid 1 correspond to amino acid 256).

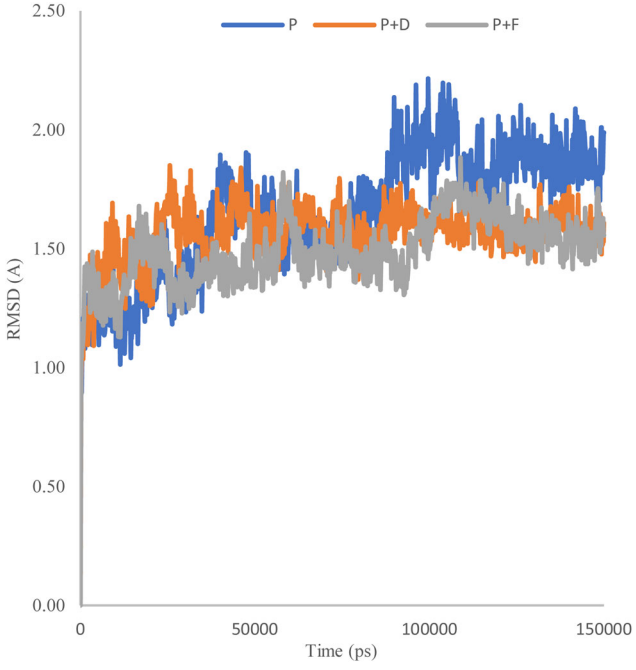

Figure 3. RMSD profile of TMPRSS2 and during 150 ns MD simulation (blue—TMPRSS2; orange—TMPRSS2-Nafamostat; grey—TMPRSS2-Fisetin).

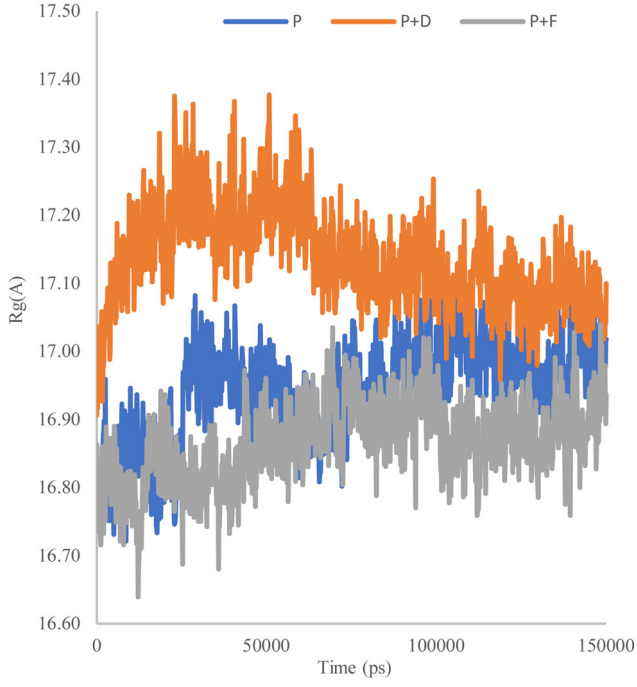

Figure 4. Compactness of TMPRSS2 in terms of radius of gyration (Rg) during 150 ns MD simulation (blue—TMPRSS2; orange—TMPRSS2-Nafamostat; grey—TMPRSS2-Fisetin).

(RMSF) was calculated to identify the residues involved in the key interactions with a ligand. Significant movement was observed around Amino acid His 41, Glu44, Lys45, Pro46, Leu47, Lys87, Gly 136 and Ser 186 which are critical residues in bonds formation. These amino acids might be playing significant role in complex formation (Figure 6). The secondary structures like alpha helices and beta strands displayed less fluctuations as compared to loop regions. The essential hydrophobic residues in the protein binding site for the ligand showed rigid behaviour with very less amounts of fluctuations, which reflects the potential of ligand for forming stable interactions with protein during 150 ns simulation (Figure 6). Figure 7 shows H bond, hydrophobic, ionic and water bridges formation between TMPRSS2 and Fisetin during entire period of simulation. Top panel indicates interaction that TMPRSS2 makes with the Fisetin in entirety. Interaction between specific residue and ligand in each

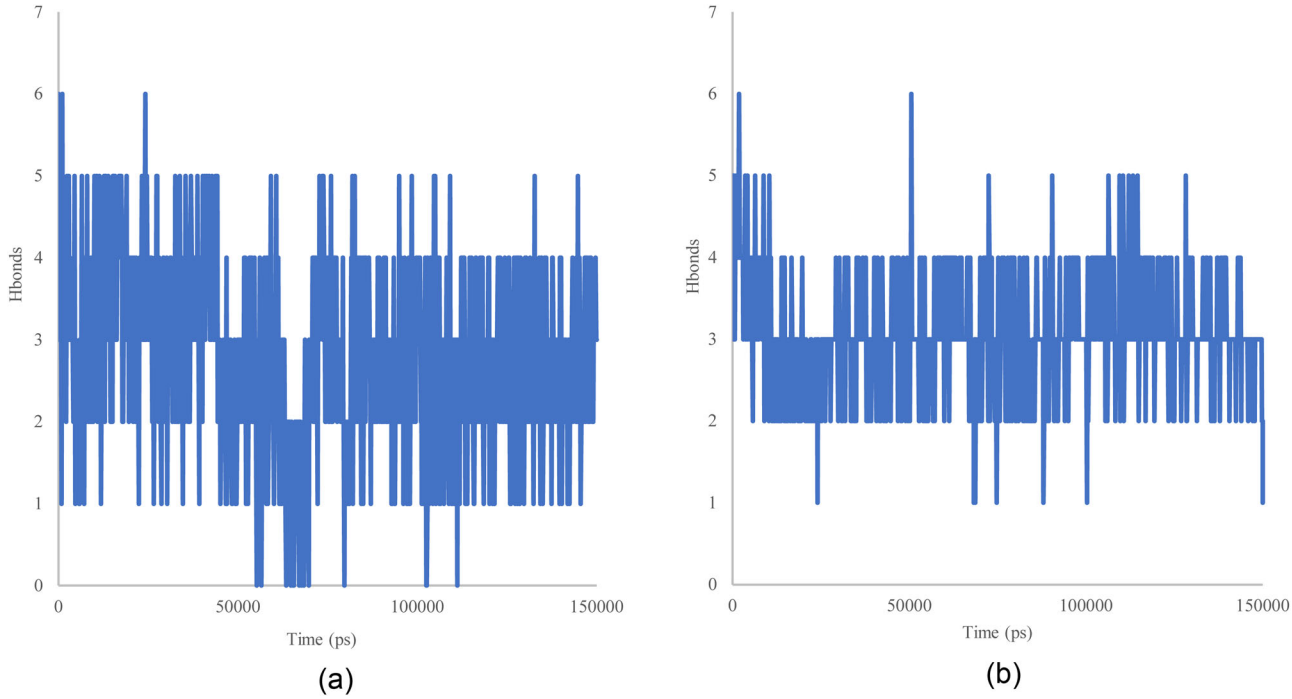

Figure 5. (a) H-bond formation between TMPRSS2 and Nafamostat during 150 ns MD simulation study. (b) H-bond formation between TMPRSS2 and Fisetin during 150 ns MD simulation study.

Table 2. Hydrogen bonds analysis with MD simulation for the TMPRSS2 inhibitor Fisetin within the active site.

	% Occupancy	Average distance (Å)
Glu44	99	2.29 ± 0.65
Lys45	76	2.45 ± 0.85
Lys87	93	2.82 ± 0.14

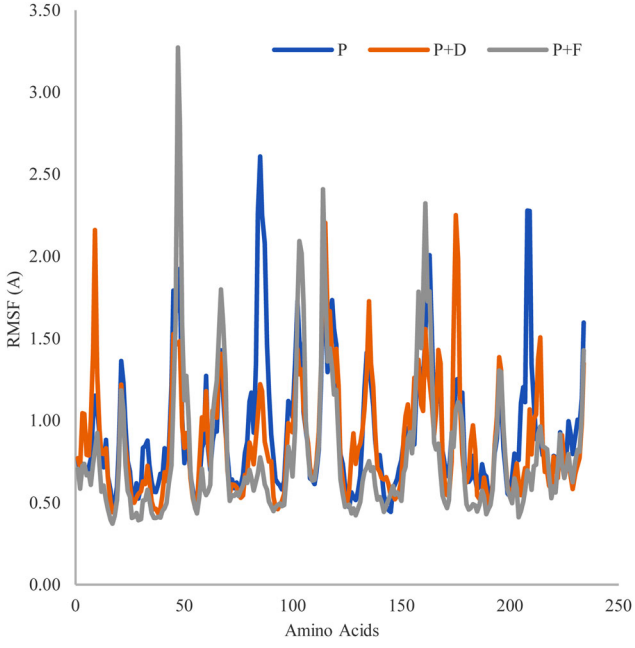

Figure 6. RMSF plot to represents local changes around TMPRSS2 chain (blue—TMPRSS2; orange—TMP-Nafamostat; grey—TMP-Fisetin).

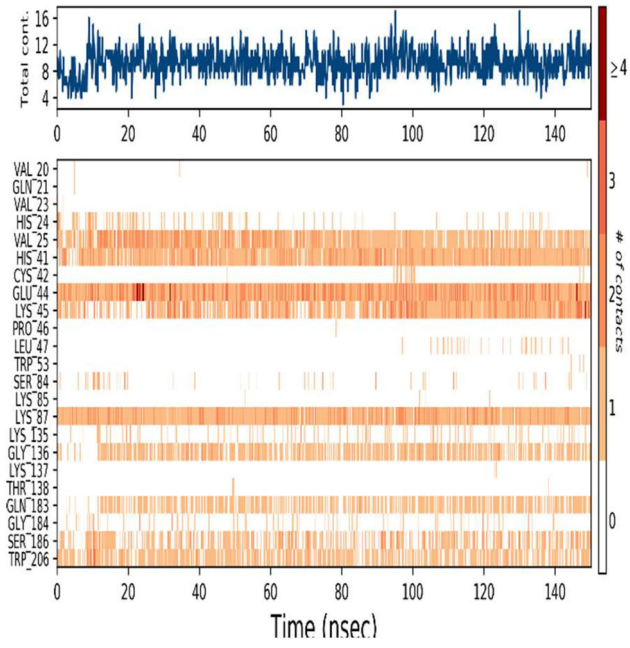

Figure 7. Top panel show specific contact and interaction study of Fisetin and TMPRSS2 during 150 ns simulation.

trajectory frame was given in bottom panel. Dark orange shade indicates formation of more than one specific contacts (Figure 7). Figures 8a, 8b, 9a, and 9b showed the interaction fraction analysis of Fisetin and Nafamostat with TMPRSS2 on the basis of hydrogen, water bridges, hydrophobic and π–π interactions throughout simulation period. Fisetin showed hydrogen bonding with His 41, Glu 44, Lys 45, Lys 85 and Ser 186. Hydrophobic occupancy showed during simulation with Val 25 and Trp 206 residues. Very long contacts of ligand with water molecules were seen which probably facilitating bridged interactions with the residues Val 25, His 41, Glu 44, Lys 135 and Gly 136. These interactions suggested that TMPRSS2 had strong binding affinity for Fisetin (Figure 8a and 8b). It was also evident that Nafamostat was making

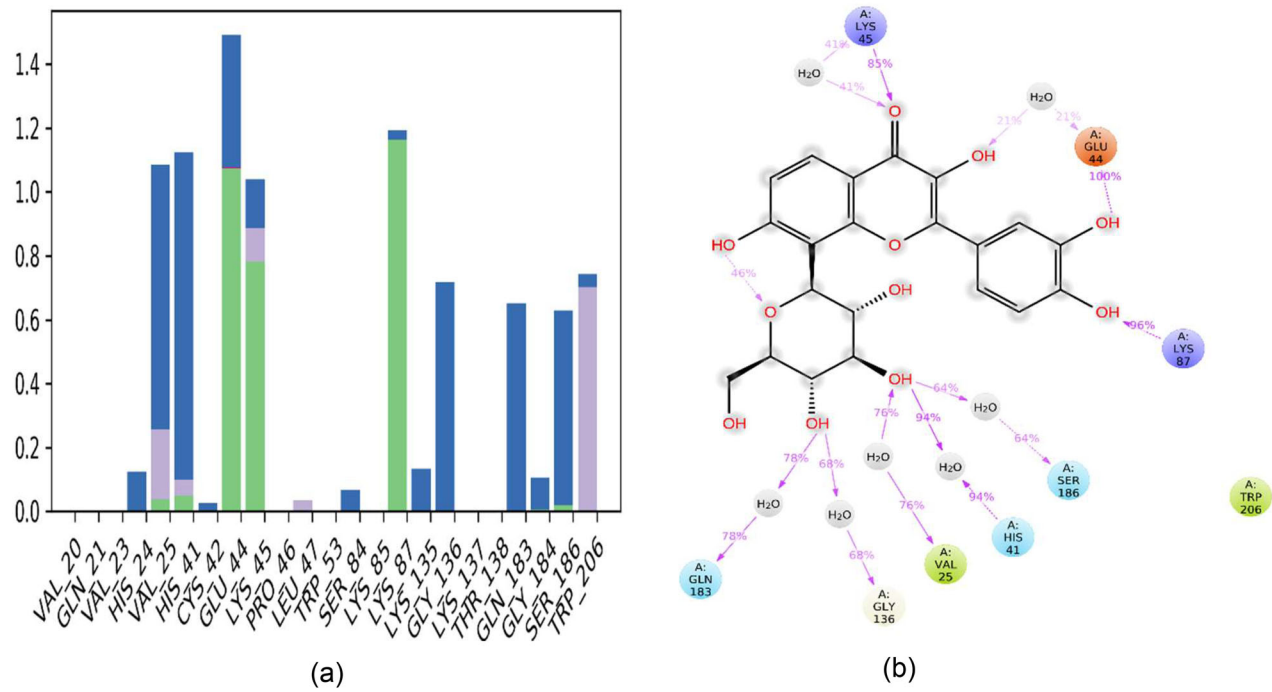

Figure 8. (a) Fisetin and TMPRSS2 interaction study during 150 ns simulation (water bridges—blue color; H-bond—green color; hydrophobic contacts—violet). (b) Schematic representation of Fisetin and TMPRSS2 interaction (green—hydrophobic, light blue—polar, red charged—salt bridge; amino acid 1 will correspond to amino acid 256).

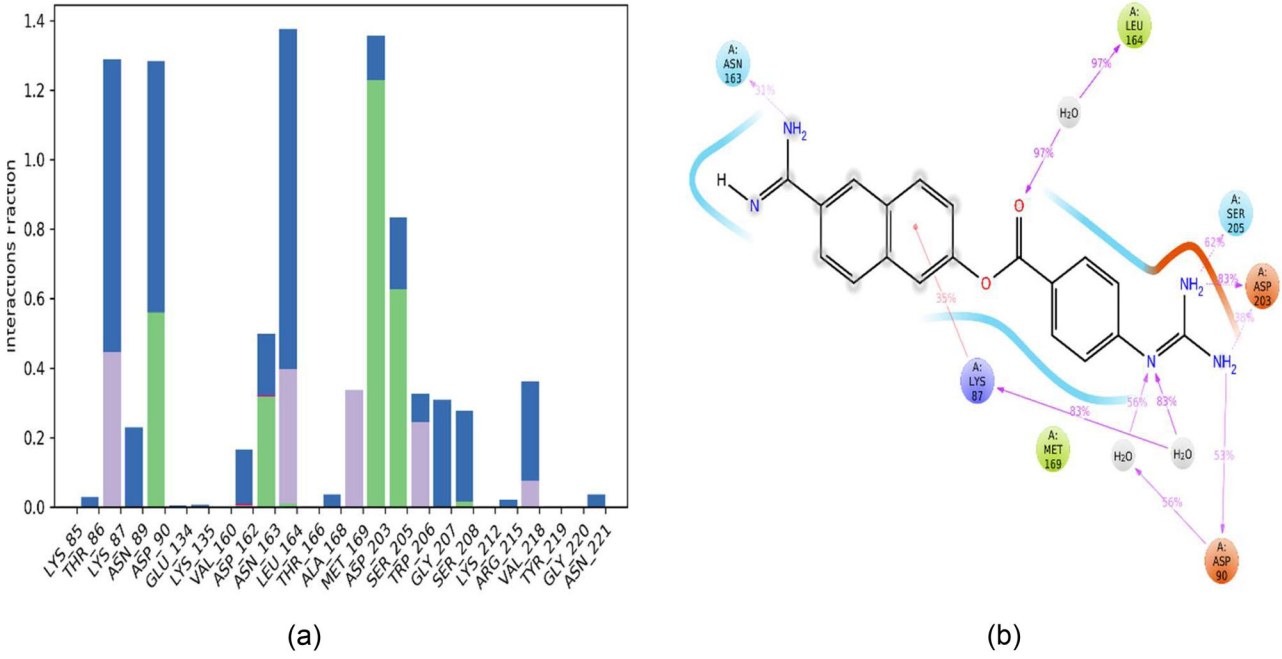

Figure 9. (a) TMPRSS2-Nafamostat interaction during 150 ns simulation (water bridges—blue color; H-bond—green color; hydrophobic contacts—violet). (b) Schematic representation of TMPRSS2 and Fisetin interaction during 150 ns simulation (green—hydrophobic, blue—polar, red charged—salt bridge; amino acid 1 will correspond to amino acid 256).

similar pattern of hydrogen bonds, hydrophobic interactions and water bridges with TMPRSS2 (Figure 9a and 9b).

Free energy determines most of the molecular processes, like molecular interaction, chemical reaction, protein folding and others. The enthalpy component influences strength of the interactions in terms of free energy. These interactions include hydrogen, van der Waals bonding, ionic, electrostatic and polarisation of the interacting groups. The entropy term of the binding free energy greatly influenced by solvation effects (due to release of bound water molecules on Fisetin binding to TMPRSS2). Molecular dynamics simulation (MD) helps to calculate free energies of molecular interactions by using theoretical calculations (Christ et al. 2009). In binding free energy estimation, docking and scoring are generally efficient but not very accurate. These methods can differentiate between binders and non-binders to certain extent

Table 3. Post MD simulation binding energies of Fisetin and Nafamostat interaction with TMPRSS2.

	Fisetin (kcal/mol)	Nafamostat (kcal/mol)
Electrostatic	−26.43	−19.50
H-bond	−3.31	−2.47
Vander Waals energy	−37.89	−42.17
Lipophilic energy	−12.84	−2.80
Pi-pi packing correction	−1.49	−10.17
Solv GB	30.09	28.88
ΔG_{bind}	−51.87	−48.23

(Gohlke & Klebe 2002). Alchemical perturbation (AP) method is another alternative which is accurate but requires large data sets and computationally intensive. Therefore, not often used in drug design (Homeyer et al. 2014; Shirts et al. 2007). There are third group of methods with intermediate performance which is again statistical method but only of the end states. These end point methods are MM/PBSA (molecular mechanics [MM] with Poisson–Boltzmann [PB] and surface area solvation) and GBSA, which is inexpensive than AP, more precise than the scoring functions (Kollman et al., 2000). MM/PB(GB)SA have been extensively studied for free energy calculations (Wang et al., 2001, 2006, Gohlke et al. 2003, Hou et al., 2008, 2009).

$$\Delta G_{bind} = \Delta H - T\Delta S = \Delta E_{MM} + \Delta G_{sol} - T\Delta S \quad (1)$$

$$\Delta E_{MM} = \Delta E_{internal} + \Delta E_{electrostatic} + \Delta E_{vdw} \quad (2)$$

$$\Delta G_{sol} = \Delta G_{PB/GB} + \Delta G_{SA} \quad (3)$$

$$G_{non-polar} = \gamma SASA + b$$

where ΔG is binding free energy, ΔH enthalpy, ΔS entropy, and T represent temperature. Surface tension of the solvent related with γ a coefficient and b is fitting parameter. ΔG_{sol}, ΔE_{MM} and $-T\Delta S$, are the variations of the solvation free energy, gas phase energy and the conformational entropy, respectively. ΔE_{MM} calculate ΔE_{vdw} (van der Waals) energies and $\Delta E_{internal}$ (bond, angle, and dihedral energies) and $\Delta E_{electrostatic}$ (electrostatic). ΔG_{solv} is the sum of the nonelectrostatic solvation component (nonpolar contribution), ΔG_{SA} electrostatic solvation energy (polar contribution). $\Delta G_{PB/GB}$, GB or PB model were used for polar contributions. Solvent accessible surface area (SASA) is used for nonpolar energy contribution. Conformational snapshots of MD simulations were used for calculation of $-T\Delta S$. Post simulation MM-GBSA binding energy estimation for 1000 snapshots was performed for the Complex. The binding energy of Fisetin to TMPRSS2 was found to be -51.87 ± 4.3 kcal/mol and range as −67.22 to −34.25 kcal/mol. The binding energy of Nafamostat to TMPRSS2 was observed as -48.23 ± 4.39 kcal/mol and range as −61.46 to −30.69 kcal/mol (Table 3). In an another study on TMPRSS2 with Camostat mesylate interaction the average MMGBSA free binding energy was reported as -65.20 ± 1.64 kcal/mol in 100 snapshots (Kumar et al., 2020).

4. Conclusion

SARS CoV-2 entry into host cells involves two important steps like binding to host cell receptor and release of virus genome into the host cell's cytoplasm. Viral entry requires a main protease (TMPRSS2) which causes proteolytic processing of viral Spike (S) protein. In the present study we demonstrated that Fisetin 8-C-glucoside stably interact with TMPRESS2. MMGBSA binding energy after docking for Fisetin 8-C-glucoside and Nafamostat was −42.78 and −21.11 kcal/mol, respectively. Post simulation MMGBSA for Fisetin and Nafamostat for 1000 snapshots was -51.87 ± 4.3 and -48.23 ± 4.39 kcal/mol. Molecular docking and simulation studies revealed that Fisetin 8-C-glucoside TMPRSS2 protein complex showed better hydrogen bonds, hydrophobic interactions than known inhibitor Nafamostat. The findings suggest a probable role of Fisetin as potential lead compound in drug development against host protease TMPRSS2. Thus, the Fisetin 8-C-glucoside can be helpful in the management of COVID-19.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Banaras Hindu University.

Reference

Alireza, P., Hosseini, M. M., & Akhavan-Niaki, H. (2020). First comprehensive computational analysis of functional consequences of TMPRSS2 SNPs in susceptibility to SARS-CoV-2 among different populations. *Journal of Biomolecular Structure and Dynamics*, https://doi.org/10.1080/07391102.2020.1767690

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215* (3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Bertram, S., Heurich, A., Lavender, H., Gierer, S., Danisch, S., Perin, P., Lucas, J. M., Nelson, P. S., Pöhlmann, S., & Soilleux, E. J. (2012). Influenza and SARS-coronavirus activating proteases TMPRSS2 and HAT are expressed at multiple sites in human respiratory and gastrointestinal tracts. *PLoS One*, *7*(4), e35876. https://doi.org/10.1371/journal.pone.0035876

Böttcher, E., Matrosovich, T., Beyerle, M., Klenk, H. D., Garten, W., & Matrosovich, M. (2006). Proteolytic activation of influenza viruses by serine proteases TMPRSS2 and HAT from human airway epithelium. *Journal of Virology*, *80*(19), 9896–9898. https://doi.org/10.1128/JVI.01118-06

Christ, C. D., Mark, A. E., & van Gunsteren, W. F. (2009). Basic ingredients of free energy calculations: A review. *Journal of Computational Chemistry*, *31*, 1569–1582.

Das, D., Koh, Y., Tojo, Y., Ghosh, A. K., & Mitsuya, H. (2009). Prediction of potency of protease inhibitors using free energy simulations with polarizable quantum mechanics-based ligand charges and a hybrid water model. *Journal of Chemical Information and Modeling*, *49*(12), 2851–2862. https://doi.org/10.1021/ci900320p

Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., Repasky, M. P., Knoll, E. H., Shaw, D. E., Shelley, M., Perry, J. K., Francis, P., & Shenkin, P. S. (2004). Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of Medicinal Chemistry*, *47*(7), 1739–1749.

Friesner, R. A., Murphy, R. B., Repasky, M. P., Frye, L. L., Greenwood, J. R., Halgren, T. A., Sanschagrin, P. C., & Mainz, D. T. (2006). Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *Journal of Medicinal Chemistry*, *49*(21), 6177–6196. https://doi.org/10.1021/jm051256o

Glide. (2015). Version 6.7. Schrödinger, LLC, New York, NY.

Glowacka, I., Bertram, S., Müller, M. A., Allen, P., Soilleux, E., Pfefferle, S., Steffen, I., Tsegaye, T. S., He, Y., Gnirss, K., Niemeyer, D., Schneider, H., Drosten, C., & Pöhlmann, S. (2011). Evidence that TMPRSS2 activates the severe acute respiratory syndrome coronavirus spike protein for membrane fusion and reduces viral control by the humoral immune response. *Journal of Virology*, *85*(9), 4122–4134. https://doi.org/10.1128/JVI.02232-10

Gohlke, H., & Klebe, G. (2002). Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angewandte Chemie International Edition*, *41*(15), 2644–2676. https://doi.org/10.1002/1521-3773(20020802)41:15<2644::AID-ANIE2644>3.0.CO;2-O

Gohlke, H., Kiel, C., & Case, D. A. (2003). Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-Ral GDS complexes. *Journal of Molecular Biology*, *330* (4), 891–913. https://doi.org/10.1016/S0022-2836(03)00610-7

Halgren, T. (2007). New method for fast and accurate binding-site identification and analysis. *Chemical Biology & Drug Design*, *69*(2), 146–148. https://doi.org/10.1111/j.1747-0285.2007.00483.x

Halgren, T. A., Murphy, R. B., Friesner, R. A., Beard, H. S., Frye, L. L., Pollard, W. T., & Banks, J. L. (2004). Glide: A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *Journal of Medicinal Chemistry*, *47*(7), 1750–1759. https://doi.org/10.1021/jm030644s

Harder, E., Damm, W., Maple, J., Wu, C., Reboul, M., Xiang, J. Y., Wang, L., Lupyan, D., Dahlgren, M. K., Knight, J. L., Kaus, J. W., Cerutti, D. S., Krilov, G., Jorgensen, W. L., Abel, R., & Friesner, R. A. (2016). OPLS3: A force field providing broad coverage of drug-like small molecules and proteins. *Journal of Chemical Theory and Computation*, *12*(1), 281–296. https://doi.org/10.1021/acs.jctc.5b00864

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T. S., Herrler, G., Wu, N.-H., Nitsche, A., Müller, M. A., Drosten, C., & Pöhlmann, S. (2020). SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell*, *181*(2), 271–280. https://doi.org/10.1016/j.cell.2020.02.052

Homeyer, N., Stoll, F., Hillisch, A., & Gohlke, H. (2014). Binding free energy calculations for lead optimization: Assessment of their accuracy in an industrial drug design context. *Journal of Chemical Theory and Computation*, *10*(8), 3331–3344. https://doi.org/10.1021/ct5000296

Hou, T. J., Xu, Z., Zhang, W., McLaughlin, W. A., Case, D. A., Xu, Y., & Wang, W. (2009). Characterization of domain-peptide interaction interface: A generic structure-based model to decipher the binding specificity of SH3 domains. *Molecular & Cellular Proteomics*, *8* (4), 639–649. https://doi.org/10.1074/mcp.M800450-MCP200

Hou, T. J., Zhang, W., Case, D. A., & Wang, W. (2008). Characterization of domain-peptide interaction interface: A case study on the amphiphysin-1 SH3 domain. *Journal of Molecular Biology*, *376* (4), 1201–1214. http://www.uniprot.org/ https://doi.org/10.1016/j.jmb.2007.12.054

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. *Journal of Virology*, *93*(6), 1815–1818. https://doi.org/10.1128/JVI.01815-18

Jacobson, M. P., Pincus, D. L., Rapp, C. S., Day, T. J. F., Honig, B., Shaw, D. E., & Friesner, R. A. (2004). A hierarchical approach to all-atom protein loop prediction. *Proteins*, *55*(2), 351–367.

Jakubík, J., Randáková, A., & Doležal, V. (2013). On homology modeling of the M2 muscarinic acetylcholine receptor subtype. *Journal of Computer-Aided Molecular Design*, *27* (6), 525–538. https://doi.org/10.1007/s10822-013-9660-8

Kawase, M., Shirato, K., van der Hoek, L., Taguchi, F., & Matsuyama, S. (2012). Simultaneous treatment of human bronchial epithelial cells with serine and cysteine protease inhibitors prevents severe acute respiratory syndrome coronavirus entry. *Journal of Virology*, *86*(12), 6537–6545. https://doi.org/10.1128/JVI.00094-12

Kevin, J. B., Chow, E., Xu, H., Dror, R. O., Eastwood, M. P., Gregersen, B. A., Klepeis, J. L., Kolossvary, I., Moraes, M. A., Sacerdoti, F. D., Salmon, J. K., Shan, Y., Shaw, D. E. (2006). *Scalable algorithms for molecular dynamics simulations on commodity clusters*. Proceedings of the ACM/IEEE Conference on Supercomputing (SC06), Tampa, 11–17.

Khan, N., Syed, D. N., Ahmad, N., & Mukhtar, H. (2013). Fisetin: A dietary antioxidant for health promotion. *Antioxidants & Redox Signaling*, *19*(2), 151–162. https://doi.org/10.1089/ars.2012.4901

Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S. H., Chong, L., Lee, M., Lee, T., Duan, Y., Wang, W., Donini, O., Cieplak, P., Srinivasan, J., Case, D. A., & Cheatham, T. E. (2000). Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Accounts of Chemical Research*, *33* (12), 889–897. https://doi.org/10.1021/ar000033j

Kumar, V., Kaur, J., Priyanshu, D., Kaul, B. A., Wang, J., Zhang, H., Kaul, S. C., Wadhwa, R., & Sundar, D. (2020). Withanone and withaferin-A are predicted to interact with transmembrane protease serine 2 (TMPRSS2) and block entry of SARS-CoV-2 into cells. *Journal of Biomolecular Structure and Dynamics*. https://doi.org/10.1080/07391102.2020.

Kumar, H., Raj, U., Srivastava, S., Gupta, S., & Varadwaj, P. K. (2016). Identification of dual natural inhibitors for chronic myeloid leukemia by virtual screening, molecular dynamics simulation and ADMET analysis. *Interdisciplinary Sciences, Computational Life Sciences*, *8*(3), 241–252. https://doi.org/10.1007/s12539-015-0118-7

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: A program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, *26*(2), 283–291. https://doi.org/10.1107/S0021889892009944

Li, J., Abel, R., Zhu, K., Cao, Y., Zhao, S., & Friesner, R. A. (2011). The VSGB 2.0 model: A next generation energy model for high resolution protein structure modeling. *Proteins*, *79*(10), 2794–2812. https://doi.org/10.1002/prot.23106

Lin, B., Ferguson, C., White, J. T., Wang, S., Vessella, R., True, L. D., Hood, L., & Nelson, P. S. (1999). Prostate-localized and androgen-regulated expression of the membrane-bound serine protease TMPRSS2. *Cancer Research*, *59*(17), 4180–4184.

Lucas, J. M., Heinlein, C., Kim, T., Hernandez, S. A., Malik, M. S., True, L. D., Morrissey, C., Corey, E., Montgomery, B., Mostaghel, E., Clegg, N., Coleman, I., Brown, C. M., Schneider, E. L., Craik, C., Simon, J. A., Bedalov, A., & Nelson, P. S. (2014). The androgen-regulated protease TMPRSS2 activates a proteolytic cascade involving components of the tumor microenvironment and promotes prostate cancer metastasis. *Cancer Discovery*, *4*(11), 1310–1325. https://doi.org/10.1158/2159-8290.CD-13-1010

MacroModel. (2015). Version 10.8. Schrödinger, LLC.

Mobley, D. L., & Dill, K. A. (2009). Binding of small-molecule ligands to proteins: "What you see" is not always "what you get". *Structure (London, England: 1993)*, *17*(4), 489–498. https://doi.org/10.1016/j.str.2009.02.010

Mulakala, C., Viswanadhan, V. N. (2013). Could MM-GBSA be accurate enough for calculation of absolute protein/ligand binding free energies? *J Mol Graph Model, 46*, 41–51. https://doi.org/10.1016/j.jmgm.2013.09.005

Park, Y. (2010). TMPRSS2 (transmembrane protease, serine 2) *Atlas. Genetics and Cytogenetics in Oncology and Haematology*, *14*, 1163–1165.

Prime. (2015). Version 4.0. Schrödinger, LLC.

Ramachandran, B., Kesavan, S., & Rajkumar, T. (2016). Molecular modeling and docking of small molecule inhibitors against NEK2. *Bioinformation*, *12*(2), 62–68. https://doi.org/10.6026/97320630012062

Rost, B. (1999). Twilight zone of protein sequence alignments. *Protein Engineering*, *12* (2), 85–94. https://doi.org/10.1093/protein/12.2.85

Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R., & Sherman, W. (2013). Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *Journal of Computer-Aided Molecular Design*, *27*(3), 221–234. https://doi.org/10.1007/s10822-013-9644-8

Shen, L. W., Mao, H. J., Wu, Y. L., Tanaka, Y., & Zhang, W. (2017). TMPRSS2: A potential target for treatment of influenza virus and coronavirus infections. *Biochimie*, *142*, 1–10. https://doi.org/10.1016/j.biochi.2017.07.016

Shirts, M. R., Mobley, D. L., & Chodera, J. D. (2007). Alchemical free energy calculations: Ready for prime time. *Annual Reports in Computational Chemistry*, *3*, 41–59.

Shivakumar, D., Williams, J., Wu, Y., Damm, W., Shelley, J., & Sherman, W. (2010). Prediction of Absolute solvation free energies using molecular dynamics free energy perturbation and the OPLS force field. *Journal of Chemical Theory and Computation*, *6*(5), 1509–1519. https://doi.org/10.1021/ct900587b

Shulla, A., Heald-Sargent, T., Subramanya, G., Zhao, J., Perlman, S., & Gallagher, T. (2011). A transmembrane serine protease is linked to the severe acute respiratory syndrome coronavirus receptor and activates virus entry. *Journal of Virology*, *85*(2), 873–882. https://doi.org/10.1128/JVI.02062-10

Sielaff, F., Böttcher-Friebertshäuser, E., Meyer, D., Saupe, S. M., Volk, I. M., Garten, W., & Steinmetzer, T. (2011). Development of substrate analogue inhibitors for the human airway trypsin-like protease HAT. *Bioorganic & Medicinal Chemistry Letters*, *21*(16), 4860–4864. https://doi.org/10.1016/j.bmcl.2011.06.033

Singh, A., & Mishra, A. (2020). Leucoefdin a potential inhibitor against SARS CoV-2 M^{pro}. *Journal of Biomolecular Structure and Dynamics*. https://doi.org/10.1080/07391102.2020.1777903.

Sinha, S. K., Shakya, A., Prasad, S. K., Singh, S., Gurav, N. K., Prasad, R. S., & Gurav, S. S. (2020). An in-silico evaluation of different Saikosaponins for their potency against SARS-CoV-2 using NSP15 and fusion spike glycoprotein as targets. *Journal of Biomolecular Structure and Dynamics*. https://doi.org/10.1080/07391102.2020.1762741

Srivastava, H. K., & Sastry, G. N. (2012). Molecular dynamics investigation on a series of HIV protease inhibitors: Assessing the performance of MM-PBSA and MM-GBSA approaches. *Journal of Chemical Information and Modeling*, *52*(11), 3088–3098. https://doi.org/10.1021/ci300385h

Wang, W., Donini, O., Reyes, C. M., & Kollman, P. A. (2001). Biomolecular simulations: Recent developments in force fields, simulations of enzyme catalysis, protein-ligand, protein-protein, and protein-nucleic acid noncovalent interactions. *Annual Review of Biophysics and Biomolecular Structure*, *30*, 211–243. https://doi.org/10.1146/annurev.biophys.30.1.211

Wang, J. M., Hou, T. J., & Xu, X. J. (2006). Recent advances in free energy calculations with a combination of molecular mechanics and continuum models. *Current Computer Aided-Drug Design*, *2* (3), 287–306. https://doi.org/10.2174/157340906778226454

Wang, L., Wu, Y., Deng, Y., Kim, B., Pierce, L., Krilov, G., Lupyan, D., Robinson, S., Dahlgren, M. K., Greenwood, J., Romero, D. L., Masse, C., Knight, J. L., Steinbrecher, T., Beuming, T., Damm, W., Harder, E., Sherman, W., Brewer, M., … Abel, R. (2015). Accurate and reliable prediction of relative ligand binding potency in prospective drug discovery by way of a modern free-energy calculation protocol and force field. *Journal of the American Chemical Society*, *137*(7), 2695–2703. https://doi.org/10.1021/ja512751q

Xie, Z., & Hwang, M. J. (2012). Ligand-binding site prediction using ligand-interacting and binding site-enriched protein triangles. *Bioinformatics (Oxford, England)*, *28*(12), 1579–1585. https://doi.org/10.1093/bioinformatics/bts182

Xu, J., Hu, S., Wang, X., Zhao, Z., Zhang, X., Wang, H., Zhang, D., & Guo, Y. (2014). Structure basis for the unique specificity of medaka enteropeptidase light chain. *Protein & Cell*, *5*(3), 178–181. https://doi.org/10.1007/s13238-013-0008-x

Yamamoto, M., Matsuyama, S., Li, X., Takeda, M., Kawaguchi, Y., Inoue, J.-I., & Matsuda, Z. (2016). Identification of nafamostat as a potent inhibitor of Middle East respiratory syndrome coronavirus S protein-mediated membrane fusion using the split-protein-based cell-cell fusion assay. *Antimicrobial Agents and Chemotherapy*, *60*(11), 6532–6539. https://doi.org/10.1128/AAC.01043-16

Yang, A. S., & Honig, B. (2000). An integrated approach to the analysis and modeling of protein sequences and structures. III. A comparative study of sequence conservation in protein structural families using multiple structural alignments. *Journal of Molecular Biology*, *301*(3), 691–711. https://doi.org/10.1006/jmbi.2000.3975