Identification of Some Dietary Flavonoids as Potential Inhibitors of TMPRSS2 Through Protein-ligand Interaction Studies and Binding Free Energy Calculations

Jibin K Varughese
Sacred Heart College Department of Chemistry
Kavitha J
Sacred Heart College Department of Chemistry
Sindhu K S
Morning Star College
Dhiya Francis
Sacred Heart College Department of Chemistry
Joseph Libin K L
Sacred Heart College Department of Chemistry
ABI T G (abitg@shcollege.ac.in)
Sacred Heart College https://orcid.org/0000-0002-5915-8609

Research Article

Keywords: TMPRSS2, COVID-19, Flavonoids, Molecular Docking, Molecular Dynamics

Posted Date: January 12th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1209434/v1

Abstract

The alarming increase in COVID-19 cases and deaths calls for an urgent cost-effective pharmacological approach. Here, we examine the inhibitory activity of a group of dietary bioactive flavonoids against the human protease TMPRSS2, which plays a major role in SARS CoV-2 viral entry. After the molecular docking studies of a large number of flavonoids, four compounds with high binding scores were selected and studied in detail. The binding affinities of these four ligands, Amentoflavone, Narirutin, Eriocitrin, and Naringin, at the active site of TMPRSS2 target were investigated using MD simulations followed by MM-PBSA binding energy calculations. From the studies, a number of significant hydrophobic and hydrogen bonding interactions between the ligands and binding site amino residues of TMPRSS2 are identified which showcase their excellent inhibitory activity against TMPRSS2. Among these ligands, Amentoflavone and Narirutin showed MM-PBSA binding energy values of -155.48 and -138.13 kJ/mol respectively. Our previous studies of the inhibitory activity of these compounds against main protease of SARS-COV2 and the present study on TMPRSS2 strongly highlighted that Amentoflavone and Naringin can exhibit promising multi-target activity against SARS-CoV-2. Moreover, due to their wide availability, no side effects and low cost, these compounds could be recommended as dietary supplements for COVID patients or for the development of SARS-CoV-2 treatments.

1. Introduction

The coronavirus disease 2019 (COVID-19) caused by the novel coronavirus severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has emerged as a global pandemic and severe threat to human beings.[1] SARS-CoV-2 is a beta coronavirus similar to SARS-CoV and MERS viruses. Compared to the other two pathogenic human coronaviruses, SARS-CoV-2 has extraordinary transmissibility which underscore the urgent need for pharmacological approaches to combat the virus.[2] Cost-effective treatments should be developed quickly to support the available vaccines for avoiding the chaos all over the world.

The SARS-CoV-2 virus is enveloped and it contains four major structural proteins - the spike protein (S), nucleocapsid protein (N), membrane protein (M), and the enveloped protein (E). Once the virus enters the host cell, its RNA is released to the cell cytoplasm and there it undergoes replication and gets converted into an effector protein by viral proteases. Our previous studies were focused on the main protease (MPro) of SARS -CoV-2 which plays a key role in viral replication and transcription. We have used computational tools to study the flavonoids like Amentoflavone, Naringin, Eriocitrin, and Narirutin and they were found to possess strong inhibitory activity against the MPro receptor.[3] After this, we have been thoroughly investigating whether these flavonoids can be used to inhibit those protein targets which assist the viral cell entry. Here we present a selection of relevant compounds that inhibit them as it would prevent the viral load[4].

Recently, it is reported that the cell entry of SARS-CoV-2 occurs as a result of binding of its spike glycoprotein S to human host cell ACE2 receptor and it is a human protease TMPRSS2 which primes and

activates S protein for its binding and fusion.[2] The spike protein has 2 subunits - S1 and S2. The S1 subunit with a receptor binding domain binds to ACE2 receptor whereas S2, the membrane anchored subunit helps in the fusion of virus and host membrane. From the studies, it is clear that TMPRSS2, the endothelial cell surface protein, helps in SARS -COV-2 viral entry and infection. The S1 subunit binds to the host cell receptor ACE2, which is followed by the priming of spike protein by the host TMPRSS2. TMPRSS2 breaks the viral S-protein at the right upstream of fusion peptide and results in membrane fusion. This points out the importance of inhibiting TMPRSS2 to avoid the host cell entry of SARS-CoV-2 virus. [1] Another reason which makes TMPRSS2 a promising target is the fact that it is a human protease, thereby this eliminates the problem of causing drug resistance like viral protein targets. [2]

Many attempts were made to develop a potential drug that can target TMPRSS2. The repurposing efficacy of some drugs were studied and from the studies, Camostat, a drug for treating chronic pancreatitis was verified as a potential inhibitor of TMPRSS2. Later, after the screening of several FDA approved drugs, Nafamostat, the drug for pancreatitis and disseminated intravascular coagulation was found to block the SARS-CoV-2 fusion at a concentration less than that required for Camostat. [2] Apart from these proven antiviral drugs, a natural compound Lutonarin [5], the polyphenols like Mangiferin, Glucogallin and Phlorizin[6] were also identified as safe TMPRSS2 inhibitors.

In this study, we focused on investigating the efficacy of bioactive compounds like flavonoids in inhibiting the serine protease ‘TMPRSS2’. The flavonoids are powerful antioxidants and they are found to be effective against diseases such as cancer, obesity, hypertension and other diseases [7]. The literature studies also underline the effectiveness of flavonoids as potential antiviral compounds. The virtual screening of a library of bioactive flavonoids was conducted based on molecular docking towards the binding site of TMPRSS2. Even though the crystal structure was available (PDB ID 7MEQ), it has some missing residues. So, its homology model was generated using 7MEQ template. Among the screened flavonoids, Amentoflavone, Narirutin, Eriocitrin and Naringin showed strong inhibitory activity against the receptor.

The bioflavonoid Amentoflavone can be isolated from Ginkgo biloba plant whereas the flavonoids Naringin is mostly found in citrus fruits. These two dietary flavonoids, which possess strong antiviral activity are available as nutritional supplements from different commercial sources [3]. Eriocitrin is also a citrus flavonoid found in lemon peel which has antioxidant and enzyme inhibitory activity.[8] Narirutin is a flavanone originally isolated from C. paradisi that has anti-inflammatory and antidepressant-like activities [9]. These natural flavonoids have several benefits over other therapeutic agents a) available as dietary supplements b) rarely have side effects c) cost-effective d) can be easily absorbed in the intestine and e) can be used effectively for home care patients with mild symptoms. These four selected compounds were then subjected to molecular dynamics simulations followed by MM-PBSA calculations. Their nonbonding interactions with the protein residues were also investigated. The MM-PBSA binding free energy calculations and nonbonding interaction studies along with docking results confirmed Amentoflavone and Narirutin has maximum antiviral activity against TMPRSS2 protein.

2. Materials And Methods

2.1 Preparation of target protein using Homology modeling

The target TMPRSS2 structure was prepared using homology modeling method in which the sequence of TMPRSS2 was obtained from Uniprot database (Id O15393). The amino acid sequence which belongs to catalytic region of TMPRSS2 [10] (256 to 491) was selected as our target sequence. The crystal structure of TMPRSS2 (PDB Id 7MEQ) with a few missing residues (Figure 1) and 98.5% sequence identity is used as the standard template. Using the homology modeling tools of Swiss model webserver [11],[12] and Modeller 9.2 [13] interface in UCSF Chimera, we generated the 3D structure of TMPRSS2 target. We checked the reliability and consistency of the modeled structure obtained from Swiss model webserver using the validation parameters such as QMEAN [14] (normal score= 0.89 and z score= -0.55) and GMQE value = 0.93. The DOPE [15] score of modeled structure from Modeller was found to be -1.12. The RMSD of the two aligned model structures is only 0.5 Å. We further validated the stereo chemical quality of both modeled structures from their Ramachandran Plots (RC-plot) using the Procheck web server [16]. The percentage of residues belong to the most favoured regions of Ramachandran plot were found to be 86.9% and 86.8% respectively for modeled structures of SWISS Model and Modeller 9.2. The modeled structure obtained from SWISS Model was utilized for further studies and their structural and energetic validations were performed using the molecular dynamic simulation of 50 ns duration.

2.2 Preparation of Ligands and Molecular Docking calculations

The structure of TMPRSS2 target obtained from homology modeling is used in molecular docking studies. After the screening of a group of flavonoids using docking studies, Amentoflavone, Narirutin, Eriocitrin and Naringin with high binding free energy (docking score) with the target were chosen for extensive study and analysis. The representative 2D structures of these flavonoids are illustrated in Fig. 2. The structure data files (SDF) of these flavonoids,, were downloaded from the Pubchem database and a DFT based structure optimization was carried out using B3LYP/SVP[17] method and Turbomole package[18]. The molecular docking calculations were performed using Autodock Vina Software [19]. The Cartesian coordinates of the center of ligand binding site in TMPRSS2 structure were fixed as x =21.65, y = -5.157, z =17.58. The location of this site lies in the neighborhood of amino acid residues of substrate binding site as per the structural information [10] of TMPRSS2 protein. The grid box centered on the binding site coordinates was used for the docking calculations. There were 6400 grid points per map where 40 grid points each were fixed in x, y, and z dimensions respectively. We fixed the grid point spacing as 1 Å and exhaustiveness as 8 in our current calculation. The best binding poses with higher docking scores were identified and then subjected to further MD studies.

2.3 Molecular dynamics simulations

The molecular dynamics simulations were executed using the conformations of the protein-ligand complex generated from molecular docking calculations. GROMACS 2018.1[20] molecular dynamics package was employed to conduct the simulations. We have used OPLS All Atom Force Field parameters[21] in the simulation procedure and the OPLS topology parameters for the ligands were created using LigPargen Server[22]. The neutralization of the protein-ligand complex structure was done by adding sodium ions to it followed by its solvation using SPC/E waters[23] in a cubic box with periodic boundary conditions. To optimize the solvated system, it was subjected to an energy minimization process using the steepest-descent minimization algorithm. [24] A 1 ns simulation of NVT and NPT equilibration was performed on the protein-ligand complex. The equilibrated system was then subjected to a 100 ns final production run with 2 fs timestep. The temperature and pressure of the system were fixed as 300K and 1 bar respectively. They were maintained throughout using Berendsen thermostat and Berendsen barostat[25]. The short-range forces of the simulating system were computed using the Verlet neighbor list and long-range electrostatic forces using the particle-mesh Ewald summation method[26].

After every 10 ps, the coordinates from the trajectory were written so that a sufficient number of frames were obtained for analysis. The stabilization of protein-ligand structure inside was monitored through various trajectory analysis tools such as RMSD, RMSF and radius of gyration. The hydrogen bonding and nonbonding analysis information throughout the MD trajectory is extracted using the GROMACS tools and Pycontact[27] tool.

2.4 MM-PBSA binding free energy calculations

MM-PBSA method, developed by Kollman and coworkers, can be used to evaluate the binding affinity of a ligand in terms of binding free energy [28]. It is proved to be useful in drug designing, studying the stabilities of macromolecular conformations, and identifying the hotspots [29]. It is also helpful in investigating the dominant binding interactions between ligand and protein [30]. This method proves to be effective in the drug discovery process, particularly for finding inhibitors against target proteins [31]. We used MM-PBSA implemented g_mmpbsa [32] code to perform binding energy calculations. The energy values corresponding to protein(P), ligand(L), and protein-ligand complex(C) conformations were obtained from the MD trajectory using a single trajectory approach. The binding free energy of the protein-ligand complex in the solution phase was calculated using Equation (1).

$$\Delta G_{\mathrm{bind}}=\mathrm{G}_{C}-\left(G_{P}+\mathrm{G}_{L}\right)$$

1

The free energy of any individual component is given by

$$\mathrm{G}=\mathrm{E}_{\mathrm{MM}}+\mathrm{G}_{\mathrm{solv}}-\mathrm{TS}$$

$$E_{\mathrm{MM}}=\mathrm{E}_{\mathrm{bonded}}+\mathrm{E}_{\mathrm{elec}}+\mathrm{E}_{\mathrm{vdW}}$$

2

The (E_{MM}) term in Equation (2) represents the energy of molecular mechanics (MM) potential with both bonding and non-bonding (van der Waals + electrostatic) terms. The value of E_{bonded} energy can be taken as zero under the assumption that the bound and unbound forms of protein and ligand conformations in the single trajectory method are similar. The (TS) term is the conformational entropy term associated with complex and isolated protein is calculated in the vacuum environment. Instead of considering absolute binding free energy, we focused on the contribution of individual residues of protein and ligands to the individual components of E_{MM} and Gsolv terms given in Equations (2, 3, and 4). As the change in entropy term does not affect the relative binding energy of ligands, it was neglected [33].

$$G_{\mathrm{solv}}=\mathrm{G}_{\mathrm{PB}}+\mathrm{G}_{\mathrm{SA}}$$

3

$$G_{\mathrm{SA}}=\gamma_{\mathrm{SASA}}+\mathrm{b}$$

4

The free energy of solvation (Gsolv) in Equation (3) is evaluated as the sum of the non-linearized version of the Poisson-Boltzmann Equation (GPB) that gives the energy of polar interactions. The nonpolar energy (GSA) term is calculated using the solvent-accessible surface area (SASA). The γ term in Equation (4) is a coefficient related to the surface tension of the solvent whereas b stands for the fitting parameter. The dielectric constants used for solvent, solute, and vacuum in the g_mmpbsa calculations have the values 80, 2, and 1, respectively, and the solvent probe radius is 1.4 Å. We conducted the MMPBSA calculations using the frames obtained from 80-100ns interval of equilibrated MD trajectory. Further we conducted 5 MD replicate simulations of 20 ns duration using the protein-ligand conformations obtained from the 100 ns MD trajectory and validated the MMPBSA results. The studies related to the energy decomposition per residues that contributed to the estimated MM-PBSA binding energy of ligand in the protein-ligand complex were also conducted.

3. Results And Discussion

3.1 Molecular docking studies

The 3D representation of the best docking poses of 4 selected flavonoids in the active site of TMPRSS2 obtained from Autodock Vina calculations is depicted in Fig. 3. The protein residues which show hydrogen bond /van der Waals interactions with the ligands in these docking poses were extracted using the freeware Maestro 12.5 (Schrödinger Release 2019-3: Maestro, Schrödinger, LLC, New York, NY). The higher negative docking scores of Amentoflavone (-9.2 kcal/mol), Narirutin (-9.1 kcal/mol), Eriocitrin (-9 kcal/mol) and Naringin (-8.1 kcal/mol) are indications of their potential binding towards TMPRSS2 target. The catalytic binding site of TMPRSS2[10] comprises of amino acid residues from residue number 256-492 and the residues SER441, HIS296, and ASP345 form the catalytic triad. The ligand Amentoflavone with the highest binding affinity showed prominent van der Waals interactions with the

catalytic residues HIS296 and SER441. Furthermore, these residues exhibit either van der Waal interactions or hydrogen bonding interactions with all other flavonoids discussed in the current work. However, the catalytic residue ASP345 shows interactions only with Narirutin and Eriocitrin. The other amino acid residues MET424, LEU419, GLN438, TRP461, CYS465, TYR474, VAL473 and CYS437 also show van der Waals interactions with Amentoflavone whereas the residue SER436 form hydrogen bond. The residues responsible for van der Waal interactions with Narirutin are HIS296, ASP345, CYS437, GLN438, TRP461 and CYS465 and that form hydrogen bonding interactions are SER436 and SER441. Eriocitrin was found to form hydrogen bonds with the residues GLY462, SER436 and hydrophobic contacts with the residues like ASP345, GLU299, GLU389, TRP461, HIS296, GLN438, SER441, and GLY464. Naringin was the ligand with least docking score with van der Waals interactions owned by the residues CYS297, HIS296, SER463, GLN438, SER441, GLY464 and hydrogen bond interactions with the residues GLU299, GLY462, TRP461 and SER436. More number of hydrogen bonding interactions were observed in the docking pose of Naringin compared to Eriocitrin, Narirutin, and Amentoflavone.

For a benchmarking, we also carried out the docking calculations of known drugs Nafamostat and Camostat [2] which are active against TMPRSS2. The docking scores of Nafamostat and Camostat were found to be -8.4 kcal/mol and -6.6 kcal/mol respectively and were closer to the reported values[34]. We also analyzed the amino acid residues involved in the binding interactions between the TMPRSS2 target and ligands reported in the previous studies [34],[35]. We also found significant nonbonding interactions between the reported residues HIS296 and SER441 and the selected flavonoids in our study. Other than these, GLN438, ASP435, GLU299, and TRP461 residues of TMPRSS2 were reported to have significant interactions with the ligands Gabexate, Camostat and Nafamostat respectively [34],[35]. The residues GLN438 and TRP461 also showed Van der Waals interactions with the selected flavonoids and GLU299 showed van der interaction with Eriocitrin and hydrogen bond interactions with Naringin. The analysis shows that our results are in good agreement with the reported literature. To check the validity of docking results, we again performed the docking of flavonoids against the modeled structure of TMPRSS2 using modeler 9.2. The docking scores and binding interactions of all flavonoids are also reproduced with negligible variations in the latter calculations (supporting data). Identification of the crucial residues belonging to the catalytic site of TMPRSS2 gave insights for further analysis.

3.2 Stability profile analysis of MD trajectory

After the docking studies, the conformations with the best binding poses of ligands were obtained. It was then minimized and subjected to MD simulation studies. The steady energies of the simulating protein ligand system with respect to time were observed and proper equilibration was ensured. Further detailed RMSD calculations of simulating complex system and protein backbone with respect to equilibrated initial structure were conducted and plotted in Fig. 4 (a) and (b) respectively.

We reported here the RMSD values of 50-100 ns interval at the end of equilibrated trajectory. The root mean square deviations of Amentoflavone, Narirutin, and Eriocitrin complex lie within the range of 0.35 – 0.45, which indicates the higher stability of these complexes in a solvated system. However, slight

deviations (0.4nm-0.5nm) were shown by Naringin and Eriocitrin complex systems. When analyzing the RMSD's of the protein backbone (Fig. 4b), similar fluctuations were observed for Narirutin (0.25nm-0.35nm), Amentoflavone (0.25nm-0.4nm), Eriocitrin (0.4nm-0.55nm), and Naringin (0.4nm-0.5nm) systems. Furthermore, we calculated the RMSD of the complex w.r.t the average structure from the MD trajectory and were plotted in Fig. 5 (a). Narirutin complex showed negligible deviation from its average structure with an RMSD of 0.7 nm. The RMSD's of Amentoflavone (1.4 nm), Eriocitrin (1.4 nm) and Naringin (1.5 nm) systems showed minimum deviations from their average structure. Thus RMSD results of all protein-ligand systems confirmed their stable structure in the 100 ns MD trajectory. Next we analyzed the contribution of individual protein residues to the structural fluctuations and its intensity in terms of RMSF (Fig. 5b). The protein residues in all complexes showed minimum fluctuations within the range 0.6 nm. An average of 4-5 residues (ARG255, ARG252, SER251 and SER254) showed comparatively higher fluctuations in the range 0.4-0.6 nm throughout the simulation. The lesser number of fluctuating residues is again an indication of stable protein ligand systems in MD simulation.

3.3 Contact map analysis

The hydrophobic and hydrogen bonding interactions between ligands and TMPRSS2 target in the MD trajectory were analyzed using GROMACS and Pycontact tools [27]. The bond distance range for all hydrogen bond interactions was fixed as 1.5 to 3.5 Å and the minimum D-H-A bond angle was fixed as 120^{o}. The maximum cutoff bond distance for all hydrophobic interactions was fixed as 5.0 Å. The classes of interactions we considered in hydrophobic interactions include pi-stacking, pi-cation, pi-alkyl, pi-amide, and other vdW interactions. The occupancy percentage of all interactions in the entire trajectory at 100 ps intervals is depicted in Fig. 6 and 7 respectively. The hydrophobic-VdW interactions play a vital in effective protein-ligand binding. TRP461 found as a crucial residue with a 100% occupancy, which showed a significant hydrophobic interaction with Amentoflavone through amide-pi stacking interaction. The other residues that form hydrophobic interactions with Amentoflavone ligand are in the order of TRP461>SER463>LYS342>GLY472>LEU419. These residues interact with Amentoflavone through vdW contacts. The residues GLN438 and THR459 show high hydrogen bond occupancy of 91.7% and 83.3% respectively with Amentoflavone. The occupancy of other amino acid residues follow the order GLY464 >GLY462 >SER460 >CYS465 >CYS437 >GLU389 >TYR474. The occupancy percentage of hydrogen-bonded residues of Narirutin complex is in the order ASP440> ASN398= ASP435> GLU395> GLY259> GLY391>LYS392 > ASN433 whereas it's hydrophobic residues follow the order ALA400=ILE381=SER436> CYS437> VAL434> ALA399> SER382. Among these residues, ALA400 and ILE381 interact with Narirutin via alkyl interactions and all other residues interact through vdW contacts. When it comes to Eriocitrin, strong hydrogen bonding interactions are shown by the residues GLY462, TRP461, and SER463 with occupancy percentages of 100%, 69.6%, and 47.8% respectively whereas the order of hydrophobic interaction is CYS465> TYR416> LEU419. The residue CYS465 interacts with Eriocitrin via alkyl interactions whereas TYR416 and LEU419 interact through other vdW contacts. In the Naringin complex, the residues HIS296 and GLN438 show high hydrogen bond occupancy of 73.91% which is followed by SER436. The residues GLU389, GLU299, and CYS297 are found to have an occupancy percentage of 39.1%. Based on the profile, the significant residues that contribute to hydrophobic interactions with the

Naringin ligand are in the order SER441> SER463=CYS465> GLY462> VAL280> LYS390> ASP345. The type of hydrophobic interactions in these residues are found to be pi-alkyl(LYS 390), pi-sulfur(CYS465), and other vdW(SER441, SER463, GLY462, VAL280, ASP345) respectively.

The catalytic triad SER441, ASP345 and HIS296 were found to be involved in significant non- bonding interactions with all the selected ligands.

3.4 MM-PBSA Binding Energies

We reported here the results of MM-PBSA calculations using 1000 snapshots obtained from the MD-trajectory between 81-100 ns at 20 ps intervals. The MM-PBSA binding energies obtained from the MD replicates (provided in the supporting information) are found to be closer with the results reported in Table 1. It is the van der Waal energy (vdW-E) that contributes mostly towards the total negative binding energy value of all the four ligands. Along with this, the electrostatic energy (ESE) and the solvent accessible surface area energy (SASA-E) also contribute to the MM-PBSA binding energy with negative values. However, the increase in the Polar Solvation Energy (PSE) with positive energy values reduces the total negative value of the binding energy. The increase in PSE is a consequence of unfavorable interaction of residues with solvent molecules. The summation of the energies contributed by both protein and ligand residues to the components such as vdW-E, ESE, PSE and SASA-E respectively gave the net MM-PBSA binding energy. The energy contribution of individual protein-ligand residues to the total binding energy was calculated using the code “energy2bfac” implemented in g_mmpbsa.

Table 1
Binding energy values and individual component energy calculated with MM-PBSA method for ligands. All reported energy values have a standard deviation within the range of 10 kJ/mol.

Sl No	Ligand	vdW-E	ESE	PSE	SASA-E	BE	SD
1	Amentoflavone	-215.400	-43.411	123.089	-19.758	-155.48	7.574
2	Narirutin	-182.604	-92.529	158.412	-21.413	-138.133	9.594
3	Eriocitrin	-208.316	-83.681	185.788	-23.592	-129.802	9.185
4	Naringin	-165.278	-88.363	159.666	-21.147	-115.122	10.888
5	Camostat	-86.431	-21.941	62.013	-11.823	-58.182	7.964
6	Nafamostat	-64.647	-60.267	90.462	-10.467	-44.920	10.495

The amino acid residues which contribute to the MM-PBSA binding energy with negative energy values are termed hotspot residues and those which contribute positive energy values are bad contact residues. From the MM-PBSA protein residue-free energy decomposition analysis, we identified the hot spots and bad contact residues of all four complexes. Among the four ligands, binding energy of Amentoflavone towards the active site of TMPRSS2 is found to possess the highest negative binding energy value of

-155.48 kJ/mol. The energy contribution of Amentoflavone residue to this total binding energy is calculated as -84.99kJ/mol which is higher compared to other ligand residues.

In the case of Amentoflavone, the hot spot residues (Figure 8a and 10a) with negative binding energy are TRP461(-16.49), GLN438(-8.54), CYS465(-5.93), SER463(-5.82), TYR474(-3.96), LEU419(-3.48), CYS437(-3.11), and GLU389(-3.22) respectively. However, ASP435, a bad contact residue which create steric clashes inside the complex, contributes a positive binding energy of 10.44 kJ/mol. The contact map analysis gave the evidence that the protein residues GLN438, CYS465, TYR474, CYS437, and GLU389 interact with Amentoflavone using hydrogen bond interactions whereas TRP461, SER463, and LEU419 through hydrophobic interactions. The residues that form hydrogen bonds with Amentoflavone contributed to the electrostatic energy part of MM-PBSA energy. TRP461 exhibits amide-pi stacking interaction and the other two residues have vdW contacts. These three hydrophobic contacts gave its contribution to the vdW component of total BE. A higher energy contribution of TRP461 residue (-16.49 KJ/mol) is pointing out its crucial involvement in the catalytic activity of TMPRSS2. Potential binding of Amentoflavone to TMPRSS is again confirmed from these results.

The next potential inhibitor of TMPRSS2 is found to be Narirutin with a protein ligand binding energy of -138.13 kJ/mol. The ligand residue contribution, in this case, is -74.92 kJ/mol. The hot spot residues (Figure 9a and 10b) with negative binding energy contribution are identified as VAL434(-7.06), ASN398(-5.71), ILE381(-5.57), GLU395(-5.39), CYS437(-5.23), SER436(-4.98), ASN433(-4.52), GLY385(-4.03) and ALA400(-3.93). As per the nonbonding interaction profile, the higher occupancy of hydrophobic contacts of the residues ALA400, ILE381, SER436, CYS437, and VAL 434 with Narirutin, contribute significantly to the vdW component of MM-PBSA binding energy. Meanwhile ASN398 and GLU395 residues in Narirutin complex contribute through hydrogen bonding interactions. The ARG316 residue with unfavourable interactions resulted in a contribution of positive binding energy value of 2.96 kJ/mol to MM-PBSA energy.

Coming to the Eriocitrin complex, the total binding energy and the ligand residue contribution are -129.80 kJ/mol and -78.63 kJ/mol respectively. The hotspot residues (Figure 9b and 10c) and their binding energy contribution, in this case, include TRP461(-9.79), SER463(-9.59), TYR416(-8.61), GLN438(-7.27) GLY462(-4.99) ,ASP417(-4.75), LEU419(-4.26), GLU389(-3.57) and GLY464(-3.22) respectively. Two residues (GLU388 and ARG470) create steric clashes and hinder the effective protein-ligand binding. From the contact map analysis, it is found that TRP461, GLY462 and SER463 interact with Eriocitrin through hydrogen bonding interactions and the residues TYR416 and LEU419 through hydrophobic-vdW interactions.

When TMPRSS2-Naringin complex was analyzed, it is found that the residues (Figure 8b and 10d) HIS296(-10.19), GLU299(-7.25), GLN438(-5.87), SER463(-4.11), CYS297(-3.02), GLU389(-2.94), GLY462(-2.51), and LYS342(-1.24) contribute to the negative binding energy whereas the residues ASP345 and LYS300 show unfavorable interactions. The contact map analysis results showed that the interaction of the protein residues GLN438, HIS296, GLU389, GLU299, and CYS297 with Naringin are

through hydrogen bonds. The residues SER463 and GLY462 show hydrophobic-vdW contact with the receptor. This complex has the least negative value of -115.12 kJ/mol. The rank the ligands in the order of its inhibitory activity was found to be as Amentoflavone>Narirutin>Eriocitrin> Naringin. For a better comparison, we calculated MM-PBSA binding free energy of Camostat and Nafomastat using the same MM-PBSA method. The binding energies of Camostat and Nafamostat were found to be only -58.182 kJ/mol and -44.92 kJ/mol respectively. The plots of MM-PBSA energy decomposition analysis of these compounds are given in the supporting data. We also analyzed the hot spot residues of TMPRSS2 obtained from the MM-PBSA calculations of other protein-ligand systems reported earlier.[34–35] The MM-PBSA energy decomposition studies of TMPRSS2 residues in Camostat / Gabexate complexes identified GLN438 as the key hot spot residue. Similarly TRP461 was identified as a hot spot residue when UKI-1complexed with TMPRSS2 [34] and GLU389 was found as a hot spot residue when Bromhexine hydrochloride complexed with TMPRSS2 [35].

These hot spot residues reported earlier found to be significant in our calculations also and they contributed with negative MM-PBSA binding energies in our calculations. The common hot spot residues of TMPRSS2 that interact with all the selected ligands are found to be GLN438, SER463, GLU389, GLY462, TRP461 and LEU419. The major roles of these residues in effective protein ligand binding were previously observed in our nonbonding and molecular docking calculations. Concluding this session, we observed that all the four flavonoids are better inhibitors than Nafamostat and Camostat in terms of MM-PBSA binding energies.

4. Conclusion

The use of natural compounds as pharmacological agents will be highly beneficial in preventing the widespread of COVD-19. All the flavonoids we have reported, except the biflavonoid Amentoflavone, are dietary compounds and therefore they can be used by the patients without the fear of side effects. From our studies, we have identified Amentoflavone and Narirutin as flavonoids with very high inhibitory activity against TMPRSS2 protein. By inhibiting TMPRSS2, they help in preventing viral entry and thus reducing the viral load inside the cell. These flavonoids bind well to the active site of the protein via hydrophobic and hydrogen bond interactions. Molecular docking, MD simulation, contact map analysis, and the MM-PBSA calculations have established Amentoflavone as the flavonoid with maximum potency for TMPRSS2 inhibition. Since all these flavonoids are widely available at a low cost, they can be used for further studies to assess their pharmacological relevance.

Declarations

Associated Content:

We provide the following data and figures in the supporting file a) PDB file of modeled TMPRSS2 structure b) Cartesian Coordinates of Ligands used in docking c) Vina Results of docking between modeled TMPRSS2 structure-2 (from Modeler) and selected ligands c) Docking poses of Camostat and

Nafamostat d) MM-PBSA Results of MD replicates e) MM-PBSA residue decomposition of Camostat and Nafamostat f) Ramachandran Plots of Modeled TMPRSS2 structures

Acknowledgements

We acknowledge DST-FIST Delhi, India, for providing the financial assistance of setting computational facilities in Heartian Center for Molecular Modeling (HCMM) of Sacred Heart College Thevara. We are also thankful to Dr. Girinath Pillai (Scientist, Nyro Research India) for giving the technical assistance.

Funding

Not applicable

Conflict of Interest Statement

The author's declared that there is no conflict of interest.

Availability of Data and Material

The preparation of target protein used in the article is performed with the help of the data obtained from RCSB Protein databank and Uniprot database. The ligand structures are obtained from Pubchem database.

Code Availability

We used licensed (academic) version of Turbomole in the DFT optimization calculation of ligands. Except Turbomole all other molecular modeling tools used to execute computational and theoretical calculations are free-wares.

Author Contribution Statement

Dr. Abi, Dr Sindhu and Jibin K Varughese designed the research project and performed the whole theoretical and computational calculations. Kavitha J, Dhiya Francis and Joseph Libin K L compiled the manuscript. Dr. Abi supervised the whole preparation of the manuscript.

References

1. Barge S, Jade D, Gosavi G, Talukdar NC, Borah J (2021) Eur J Pharm Sci 162:105820. doi:10.1016/j.ejps.2021.105820
2. Zhu H, Du W, Song M, Liu Q, Herrmann A, Huang Q (2021) Computational and Structural Biotechnology. Journal 19:467. doi:https://doi.org/10.1016/j.csbj.2020.12.035
3. Varughese JK, Joseph Libin KL, Sindhu KS (2021) 1. doi:10.1080/07391102.2021.1891139
4. Baby K, Maity S, Mehta CH, Suresh A, Nayak UY, Nayak Y (2021) Eur J Pharmacol 896:173922. doi:10.1016/j.ejphar.2021.173922

5. Roomi M, Khan Y (2020) Potential Compounds for the Inhibition of TMPRSS2
6. Singh R, Gautam A (2020) 25. doi:10.3390/molecules25204604
7. Middleton E Jr, Kandaswami C, Theoharides TC (2000) Pharmacol Rev 52:673
8. Horowitz RM, Gentili B (1960) J Am Chem Soc 82:2803. doi:10.1021/ja01496a033
9. Mizelle JW, Dunlap WJ, Hagen RE, Wender SH, Lime BJ, Albach RF, Griffiths FP (1965) Anal Biochem 12:316. doi:10.1016/0003-2697(65)90097-7
10. Shulla A, Heald-Sargent T, Subramanya G, Zhao J, Perlman S, Gallagher T (2011) J Virol 85:873. doi:10.1128/jvi.02062-10
11. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R, Schwede T (2018) Nucleic Acids Res 46:W296. doi:10.1093/nar/gky427
12. Bienert S, Waterhouse A, de Beer TA, Tauriello G, Studer G, Bordoli L, Schwede T (2017) 45: D313. doi:10.1093/nar/gkw1132
13. Webb B, Sali A (2016) Curr Protoc Bioinformatics 54: 5.6.1. doi:10.1002/cpbi.3
14. Benkert P, Biasini M, Schwede T (2010) Bioinformatics 27:343. doi:10.1093/bioinformatics/btq662
15. Shen MY, Sali A (2006) Protein Sci 15:2507. doi:10.1110/ps.062416606
16. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) J Appl Crystallogr 26:283. doi:https://doi.org/10.1107/S0021889892009944
17. Becke AD (1988) Phys Rev A 38:3098. doi:10.1103/PhysRevA.38.3098
18. Bauernschmitt R, Häser M, Treutler O, Ahlrichs R (1997) Chem Phys Lett 264:573. doi:https://doi.org/10.1016/S0009-2614(96)01343-7
19. Trott O, Olson AJ (2010) J Comput Chem 31:455. doi:https://doi.org/10.1002/jcc.21334
20. Berendsen HJC, van der Spoel D, van Drunen R (1995) Comput Phys Commun 91:43. doi:https://doi.org/10.1016/0010-4655(95)00042-E
21. Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) J Am Chem Soc 118:11225. doi:10.1021/ja9621760
22. Dodda LS, Cabeza de Vaca I, Tirado-Rives J, Jorgensen WL (2017) Nucleic Acids Res 45:W331. doi:10.1093/nar/gkx312
23. Toukan K, Rahman A (1985) Phys Rev B 31:2643. doi:10.1103/PhysRevB.31.2643
24. Chen C, Huang Y, Ji X, Xiao Y (2013) J Chem Phys 138:164122. doi:10.1063/1.4799236
25. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) J Chem Phys 81:3684. doi:10.1063/1.448118
26. Di Pierro M, Elber R, Leimkuhler B (2015) J Chem Theory Comput 11:5624. doi:10.1021/acs.jctc.5b00648
27. Scheurer M, Rodenkirch P, Siggel M, Bernardi RC, Schulten K, Tajkhorshid E, Rudack T (2018) Biophys J 114:577. doi:10.1016/j.bpj.2017.12.003

28. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA, Cheatham TE (2000) Acc Chem Res 33:889. doi:10.1021/ar000033j
29. Zoete V, Irving MB, Michielin O (2010) J Mol Recognit 23:142. doi:https://doi.org/10.1002/jmr.1005
30. Li S, Xu W, Chu S, Ma N, Liu S, Li X, Wang T, Jiang X, Li F, Li Y, Zhang D, Luo Q, Liu J (2019) Chemistry – A European Journal 25:10350. doi:https://doi.org/10.1002/chem.201901480
31. Genheden S, Ryde U (2015) Expert Opin Drug Discov 10:449. doi:10.1517/17460441.2015.1032936
32. Kumari R, Kumar R, Lynn A (2014) Journal of Chemical Information and Modeling 54: 1951. doi:10.1021/ci500020m
33. Hou Q, Zhang L (2020) Langmuir 36:1813. doi:10.1021/acs.langmuir.9b03832
34. Hu X, Shrimp JH, Guo H, Xu M, Chen CZ, Zhu W, Zakharov AV, Jain S, Shinn P, Simeonov A, Hall MD, Shen M (2021) ACS Pharmacology Translational Science 4:1124. doi:10.1021/acsptsci.0c00221
35. Sonawane KD, Barale SS, Dhanavade MJ, Waghmare SR, Nadaf NH, Kamble SA, Mohammed AA, Makandar AM, Fandilolu PM, Dound AS, Naik NM, More VB (2021) Informatics in medicine unlocked 24: 100597. doi:10.1016/j.imu.2021.100597

Figures

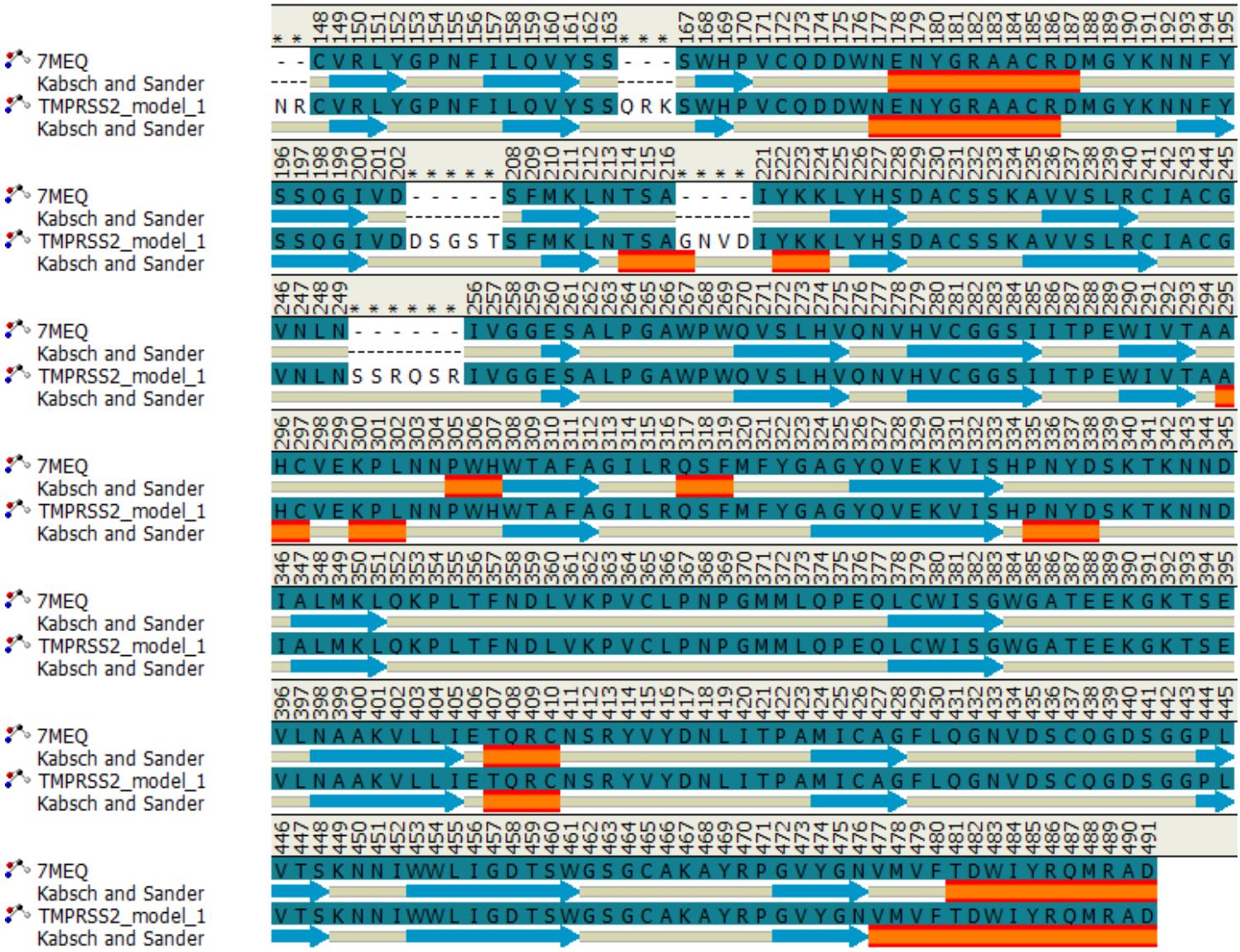

Figure 1

Amino acid sequence and Kabsch and Sander secondary structure representation of 7MEQ template and modelled structure of TMPRSS2

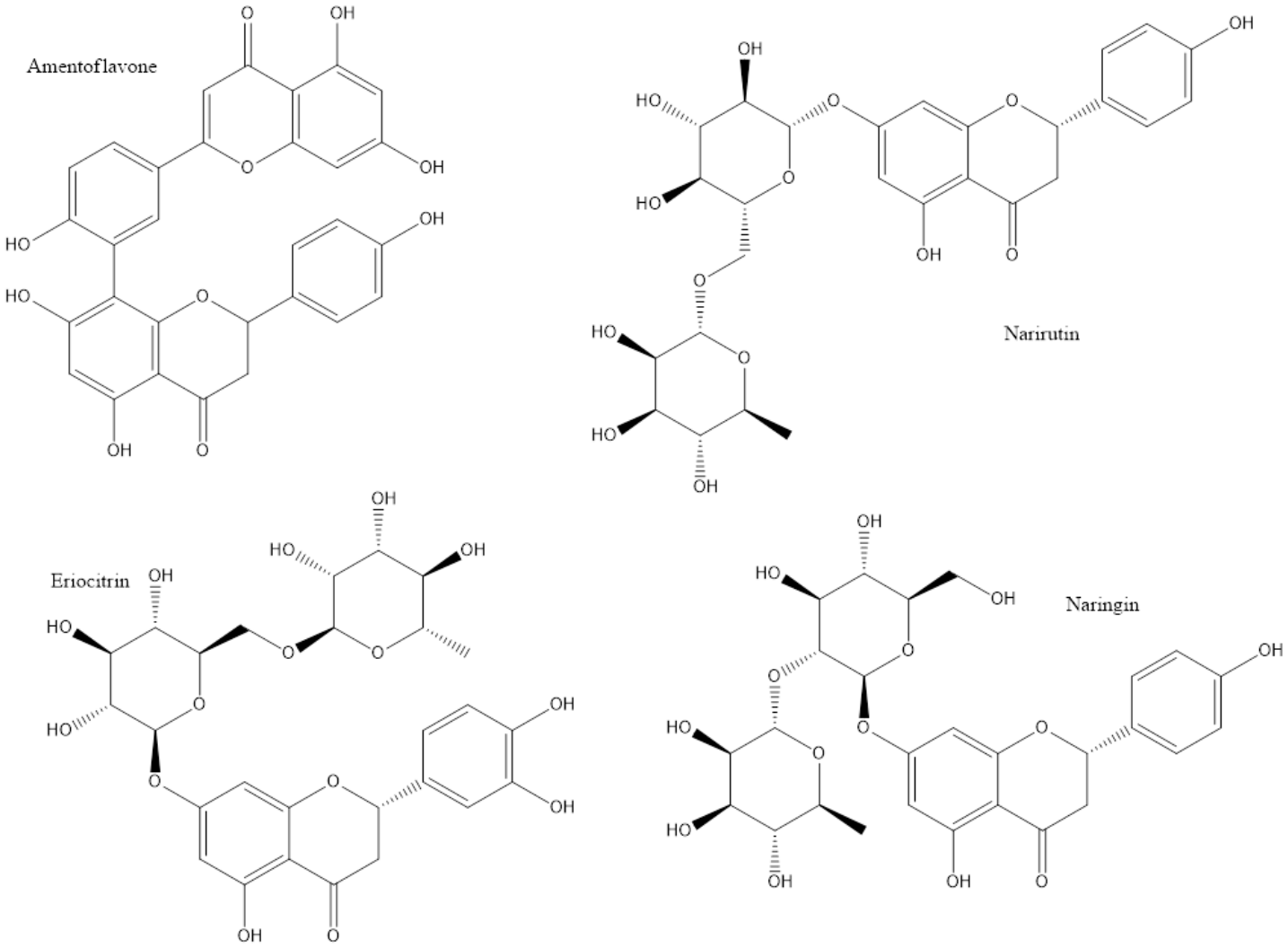

Figure 2

Representation of selected flavonoids

Figure 3

A 3D representation of best docking poses of selected ligands at the active site of TMPRSS2***.***

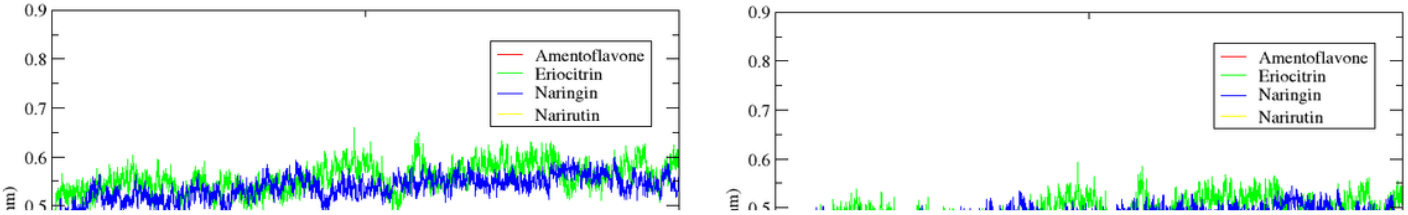

Figure 4

RMSD plots of (a) protein ligand complexes and (b) protein backbone in these complexes

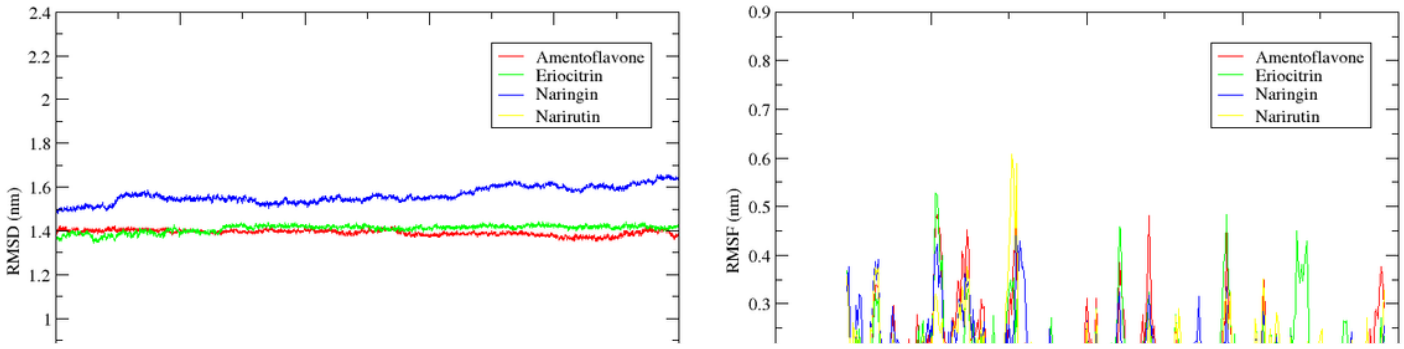

Figure 5

(a) RMSD plots of complex with respect to average structure and (b) RMSF of residues of TMPRSS2

Figure 6

Protein-ligand hydrogen bonding interaction profile of protein ligand complexes from MD trajectory

Figure 7

Protein-ligand hydrophobic interaction profile of protein ligand complexes from MD trajectory

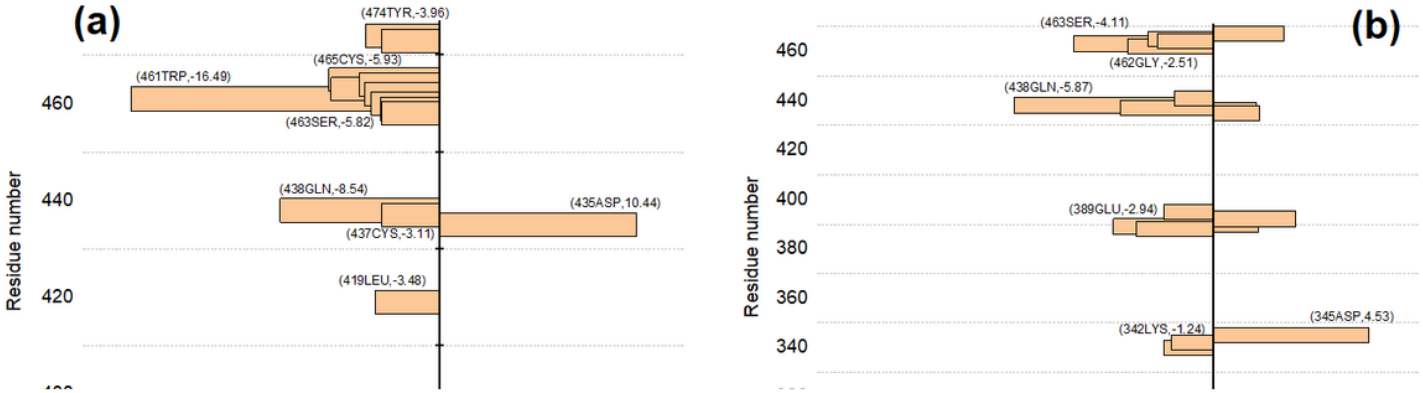

Figure 8

MM-PBSA Binding Energy decomposition of TMPRSS2 *residues in a) Amentoflavone complex and* ***b)****Naringin complex*

Figure 9

MM-PBSA Binding Energy decomposition of TMPRSS2 *residues in a) Narirutin complex and* ***b)*** *Eriocitrin complex*

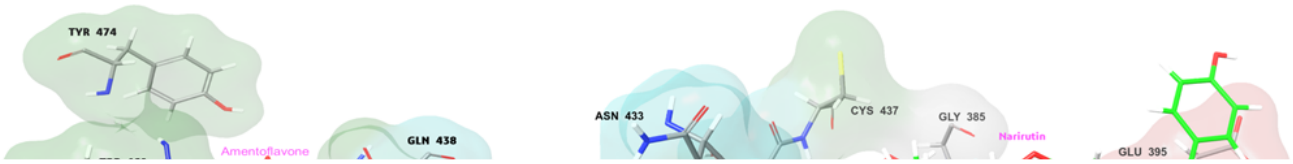

Figure 10

a: Hot spot residues ***in Amentoflavone complex***

b: Hot spot residues ***in Narirutin complex***

c: Hot spot residues ***in Eriocitrin complex***

d: Hot spot residues ***in Naringin complex***

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupportingData1.pdf