




Identification of potential anti-TMPRSS2 natural products through homology modelling, virtual screening and molecular dynamics simulation studies

Rupesh V. Chikhale , Vivek K. Gupta , Gaber E. Eldesoky , Saikh M. Wabaidur , Shripad A. Patil & Md Ataul Islam


To cite this article: Rupesh V. Chikhale , Vivek K. Gupta , Gaber E. Eldesoky , Saikh M. Wabaidur , Shripad A. Patil & Md Ataul Islam (2020): Identification of potential anti-TMPRSS2 natural products through homology modelling, virtual screening and molecular dynamics simulation studies, Journal of Biomolecular Structure and Dynamics, DOI: [10.1080/07391102.2020.1798813](https://doi.org/10.1080/07391102.2020.1798813)

To link to this article: <https://doi.org/10.1080/07391102.2020.1798813>

 View supplementary material [↗](#)

 Published online: 03 Aug 2020.

 Submit your article to this journal [↗](#)



 Article views: 1266

 View related articles [↗](#)

 View Crossmark data [↗](#)



Identification of potential anti-TMPRSS2 natural products through homology modelling, virtual screening and molecular dynamics simulation studies

Rupesh V. Chikhale^a , Vivek K. Gupta^b, Gaber E. Eldesoky^c, Saikh M. Wabaidur^c, Shripad A. Patil^b and Md Ataul Islam^{d,e,f} 

^aSchool of Pharmacy, University of East Anglia, Norwich, UK; ^bDepartment of Biochemistry, ICMR-National JALMA Institute for Leprosy and Other Mycobacterial Diseases (ICMR), Agra, India; ^cDepartment of Chemistry, College of Science, King Saud University, Riyadh, Saudi Arabia; ^dDivision of Pharmacy and Optometry, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK; ^eSchool of Health Sciences, University of Kwazulu-Natal, Durban, South Africa; ^fDepartment of Chemical Pathology, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa

Communicated by Ramaswamy H. Sarma

ABSTRACT

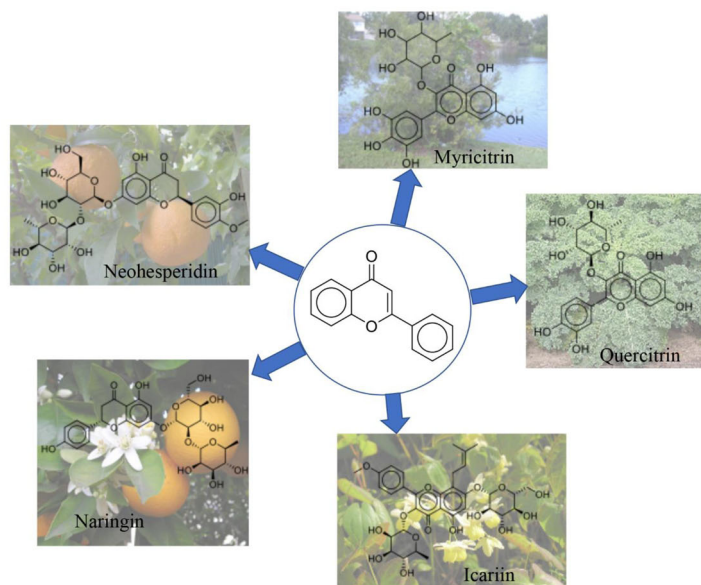
Recent outbreak of severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) has led to a pandemic of COVID-19. The absence of a therapeutic drug and vaccine is causing severe loss of life and economy worldwide. SARS-CoV and SARS-CoV-2 employ the host cellular serine protease TMPRSS2 for spike (S) protein priming for viral entry into host cells. A potential way to reduce the initial site of SARS-CoV-2 infection may be to inhibit the activity of TMPRSS2. In the current study, the three-dimensional structure of TMPRSS2 was generated by homology modelling and subsequently validated with a number of parameters. The structure-based virtual screening of Selleckchem database was performed through 'Virtual Work Flow' (VSW) to find out potential lead-like TMPRSS2 inhibitors. Camostat and bromhexine are known TMPRSS2 inhibitor drugs, hence these were used as control molecules throughout the study. Based on better dock score, binding-free energy and binding interactions compared to the control molecules, six molecules (Neohesperidin, Myricitrin, Quercitrin, Naringin, Icariin, and Ambroxol) were found to be promising against the TMPRSS2. Binding interactions analysis revealed a number of significant binding interactions with binding site amino residues of TMPRSS2. The all-atoms molecular dynamics (MD) simulation study indicated that all proposed molecules retain inside the receptor in dynamic states. The binding energy calculated from the MD simulation trajectories also favour the strong affinity of the molecules towards the TMPRSS2. Proposed molecules belong to the bioflavonoid class of phytochemicals and are reported to possess antiviral activity, our study indicates their possible potential for application in COVID-19.

ARTICLE HISTORY

Received 17 June 2020
Accepted 16 July 2020

KEYWORDS

SARS-CoV-2; COVID-19; TMPRSS2; natural products; virtual screening; molecular docking; molecular dynamics



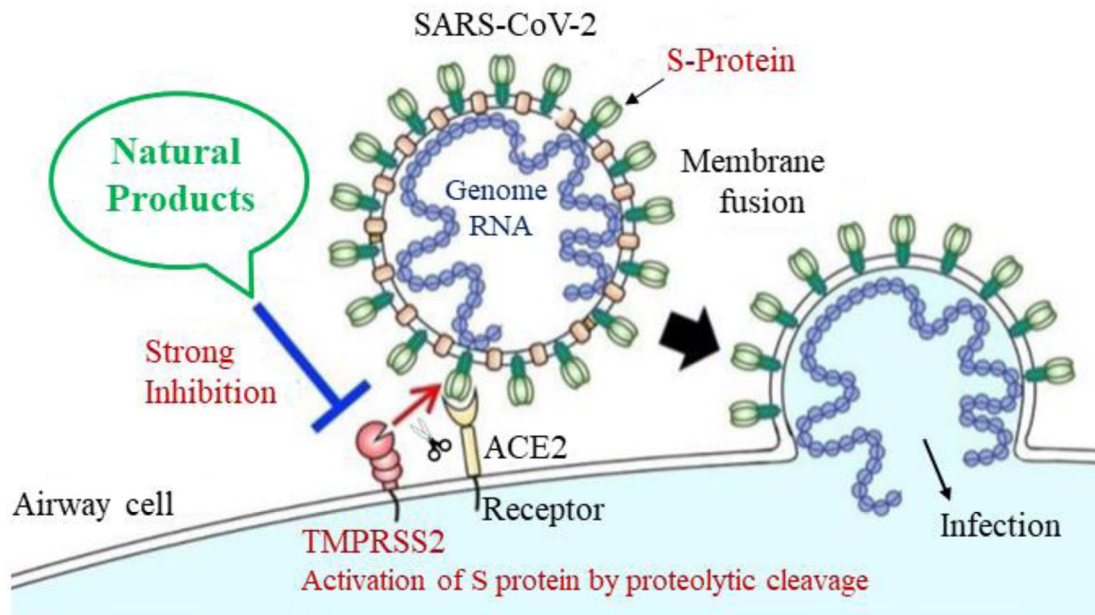


Figure 1. TMPRSS2 play important role in SARS-CoV-2 viral entry into the host cell. It does not appear to play an essential role in the host organ system and provides a degree of redundancy. Therefore, it has become the target of choice for broad therapeutic applicability against SARS-CoV2. Reproduced Image from 'Nafamostat, an existing safe drug, may inhibit entry of SARS-CoV-2' (The Institute of Medical Science, The University of Tokyo).

Introduction

The recent spread of the novel, pathogenic Severe Acute Respiratory Syndrome-Corona Virus-2 (SARS-CoV-2) from China to the entire world has taken the shape of a global health emergency (Lu et al., 2020). The World Health Organisation (WHO) has declared coronavirus disease 19 (COVID-19) a pandemic, characterised by rapid human-to-human transmission, high morbidity, and mortality rate (Zhou et al., 2020). Countries across the globe have reported confirmed COVID-19 related infections and deaths (WHO, 2020). So far about 3,50,000 deaths and more than 5 million confirmed cases have reported worldwide. Moreover, the deaths and new infections are increasing exponentially day by day. Unfortunately, there are no effective drugs against the SARS-CoV-2 virus to date which creates an extremely concerning situation to the entire world. The antiviral drugs remdesivir and favipiravir are considered as a therapeutic option in the treatment of COVID-19. Scientific communities across academia and industry are impatiently searching for effective drugs from the already FDA approved drugs or the drugs in clinical studies for repurposing.

Many epidemiological studies have shown that the transmission characteristics of 2019-nCoV appear to be similar to severe acute respiratory syndrome-associated coronavirus (SARS-CoV). The spike (S) protein of coronaviruses facilitates viral entry into target cells. Notably, the amino acid sequence identity between the Wuhan-CoV and SARS-CoV Spike proteins is 76.47% (Wang et al., 2020). The SARS-CoV-2 host cell entry has two essential stages; first, the SARS-CoV-2 contacts the angiotensin-converting enzyme 2 (ACE2) through its spike's receptor binding domain (RBD) and second, the transmembrane protein serine 2 (TMPRSS2) cleaves viral S protein causing an irreversible conformational change leading to

fusion of viral S protein and the host cell membrane (Figure 1). The S protein cleavage, called priming allows the virus and host cell membranes to fuse, a process driven by the S2 subunit (Glowacka et al., 2011; Hoffmann et al., 2020; Matsuyama et al., 2010; Shulla et al., 2011).

The transmembrane protease serine 2 (TMPRSS2) is a type II transmembrane serine protease (TTSP) which plays a critical role in SARS-CoV. The Middle East respiratory syndrome-associated coronavirus (MERS-CoV) and Asian H7N9 influenza virus along with some H1N1 subtypes influenza-A viruses infections are indicating that targeting TMPRSS2 could be a novel antiviral strategy to treat coronaviruses and some low pathogenic influenza virus infections (Hatesuer et al., 2013; Zhang & Liu, 2020). TMPRSS2 is an androgen-regulated type II transmembrane serine protease widely expressed in epithelial cells of the respiratory, gastrointestinal, and urogenital tract with high expression levels in the prostate and colon (Bertram et al., 2012; Bugge et al., 2009; Chen et al., 2010; Lucas et al., 2014). The lack of TMPRSS2 in the airways reduces the severity of lung pathology in the animal model after infection by SARS-CoV and MERS-CoV (Zhu et al., 2020).

It has been known that androgen can positively regulate TMPRSS2 expression (Lin et al., 1999), hence the levels of TMPRSS2 in adult males are found higher than females or children (Vetter et al., 2020). Recently, Wang et al. (2020) hypothesised that TMPRSS2 expression may positively correlate with severity in COVID-19 (Yamamoto et al., 2016). It has been also reported that H7N9 influenza (Chen et al., 2013; Jernigan & Cox, 2015) and COVID-19 (Gupta et al., 2020) affect the proportion of male patients more than double that of female patients. Candidate targets such as receptors and proteases for virus entry and membrane fusion now considered as potential targets (Hoffmann et al., 2020). Successful

inhibition of TMPRSS-2 may be a potential way to reduce the initial site of SARS-CoV-2 infection. The above strategy might be a safe therapeutic approach and also represents a promising candidate for potential interventions against COVID-19. It is already approved and accepted that a TMPRSS2 inhibitor in clinical use can block the entry and might constitute a successful treatment option for the COVID-19 (Hoffmann et al., 2020). Hence, TMPRSS2 could be an important target to design potential inhibitors with a unique benefit to block this host protease.

Camostat mesylate is a well-known TMPRSS2 inhibitor, widely used in Japan and approved for the treatment of acute symptoms of chronic pancreatitis (CP) (Kawase et al., 2012; Zhou et al., 2015). It was also found to have antiviral activity against SARS-CoV by inhibiting TMPRSS2 (Shulla et al., 2011). H1N1 influenza virus-infected human tracheal epithelial cells treated with camostat mesylate significantly reduced viral load (Bahgat et al., 2011; Lee et al., 1996). Recently, Nafamostat mesylate was reported to have antiviral activity against the Middle East Respiratory Syndrome Corona Virus (MERS-CoV) infection *in vitro* via inhibiting the activity of TMPRSS2, and reduced viral entry more effectively than camostat (Yamamoto et al., 2016). Bromhexine hydrochloride (BHH) is a Food and Drug Administration (FDA) approved drug and it is a semisynthetic derivative from the India shrub *Adhatoda vasica*. BHH is reported to effectively attenuate prostate cancer metastasis in mice by specific inhibition of TMPRSS2 (Ko et al., 2015). Guaifenesin is reported as a mucoactive drug that acts by relaxing mucus in the airways and making coughs more productive. Moreover, Guaifenesin is legally marketed medicine in the US and it is used to help in wet cough and chest congestion that develops due to the common cold (Albrecht et al., 2017).

Natural products are evolutionarily optimised as lead-like compounds and to date, they are the greatest sources of therapeutic drugs for a wide range of diseases (Newman & Cragg, 2012). Ayurveda and Traditional Chinese Medicines (TCM) have been used in the treatment of various chronic and infectious diseases (Dobos & Tao, 2011; Mathpati et al., 2020). Such extensive potential of natural products can be used to find possible TMPRSS2 inhibitors for therapeutic application in COVID-19. To explore this potential, we planned to perform screening of natural product database against the homology model of TMPRSS2. Further, the chemical space was reduced by molecular docking-based binding energy and long-range classical molecular dynamics (MD) simulations. Lastly, the binding-free energy is calculated through molecular mechanics-generalised Born surface area (MM-GBSA) approach to decide final phytocompounds (Genheden & Ryde, 2015). The credential of the current work substantiated by the finding of five promising phytocompounds as TMPRSS2 inhibitors.

Materials and methods

Homology modelling of TMPRSS2

To understand the structural and functional aspects of any protein molecule, a three dimensional (3D) structure is

essential. Unfortunately, the crystal structure of TMPRSS2 protease has not been resolved to date. In this situation, the comparative structure prediction approach is the only option to generate the 3D coordinates of the TMPRSS2 protease molecule. A total of 492 amino acid residues long sequence of TMPRSS2 protease obtained from the Universal Protein Resource 'UniProtKB' database (Accession Number: O15393) and homology modelling performed by template-based method on the Swiss-Model (Biasini et al., 2014). The Swiss-Model is an extremely popular and critically acclaimed web-based modelling tool. The web interface is designed in such a way that users can easily build and evaluate the protein homology models at many different levels of complexity. This tool is fully automated with low user intervention required. A BLAST (Johnson et al., 2008) search was performed followed by HHBLITS (Remmert et al., 2012) search leading to 1854 applicable templates. Among all the retrieved templates, the Serine Protease hepsin (Herter et al., 2005) having PDB ID: 1Z8G was found to be the best template with a 35.21% identity and best possible coverage. Moreover, the sequence similarity of the query sequence with the template was found to be 39.00%. By considering the query sequence and selected template, Swiss-Model developed the 3D structure of TMPRSS2. Before use in the molecular docking study, the developed 3D structure of TMPRSS2 was validated through a number of validation protocols to check its reliability and consistency. The model was first analysed for its stereochemical quality using the MOE software (Molecular Operating Environment [MOE], 2016) package. It provided an analysis of steric contacts within the protein and high-accuracy Ramachandran Plot (RC-plot). The second set of validation was performed on the model using the MolProbity (Lovell et al., 2003) and QMEAN (Benkert et al., 2009) server to determine the qualitative and energetic analysis. The QMEAN predicts the 'degree of nativeness' of the structural features found in the modelled protein on a global scale (Benkert et al., 2011). This parameter suggests that whether predicted protein structure through homology modelling approach is comparable to what one would expect from experimental structures of similar size (Benkert et al., 2011). It is composed of three distance dependent interaction statistical potential terms. In addition to the above, another two terms are also involved which describe the agreement of the predicted and observed secondary structure, and solvent accessibility. The QMEAN score in the range of 0 to -4 explains good agreement between the predicted protein structure and experimentally derived structures of similar size (SwissModel Help, 2020; Waterhouse et al., 2018). The QMEAN value around 0 suggests high quality of homology modelled protein structure (SwissModel Help, 2020). The structural and energetic validation was performed using the molecular dynamic simulation for the modelled TMPRSS2.

Molecular dynamics simulations of modelled TMPRSS2

To explore the structural and energetic status as well as the steric refinement of the developed TMPRSS2 structure, an all-

atom MD simulation of 100 ns was performed using AMBER18 (Song et al., 2019) software package. The Tmprss2 homology model immersed in truncated octahedron of TIP3P (Mark & Nilsson, 2001) water giving a total of 24,515 water molecules in the system. The sufficient number of Na⁺ and Cl⁻ counter ions added to neutralise the system and achieve an ionic strength of 0.1 M to mimic the physiological pH. For the protein topology generation, the ff14SB force field (Maier et al., 2015) was used. The entire experiment of MD simulation was executed on Nvidia V100-SXM2-16GB Graphic Processing Unit using the PMEMD.CUDA (Peramo, 2016) module installed on the Computational Shared Facility (CSF3), University of Manchester, UK. Simulations were run at 300 K using the Langevin thermostat with a collision frequency of 2 ps⁻¹, at 1 atm using a Monte Carlo barostat with volume exchange attempts every 100 fs. A 2-fs integration step was employed. Covalent bonds involving hydrogen constrained using SHAKE (Andersen, 1983) algorithm. A cut off of 8 Å was used for short-range nonbonded interactions whilst long-range electrostatics were treated using the particle mesh Ewald method. Equilibration consisted of rounds of NVT and NPT equilibration for 10 ns in total. Production MD run was performed for 100 ns. Root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), and other interactions were analysed using CPPTRAJ (Roe & Cheatham, 2013) over full trajectory, taking configuration in every 4 ps.

Identification of active site and grid preparation

On successful completion of MD simulation, a representative frame of Tmprss2 was extracted using the CPPTRAJ module of Amber18 and further used for active site identification. Identification of the binding site of any protein molecule is a crucial step for successful inhibition or activation of the protein. Before the molecular docking, the active site of the Tmprss2 was explored through the SiteMap module of the Schrodinger software package (Halgren, 2007). The SiteMap module is an excellent tool for identifying the active site in the protein receptor. This tool starts the searching of sites on the protein surface by linking the site-points. In the next step, it involves forming maps having hydrophobicity or hydrophilicity regions. Finally, it calculates the energy property of each site and scored accordingly (Halgren, 2007). Four top most active sites predicted by the Sitemap were analysed concerning the RMSF data obtained from the MD trajectory of residues in the predicted druggable site. The site formed by residues His18, Gln21, Glu23, Asn24, Pro25, Val49, Pro50, Gln51, Tyr52, Ala53, Pro54, Arg55, Gln59, Val65, Gln68, Pro69, Val96, Gly97, Ala98, Ala99, Ala101, Met371, Met372, Leu373, Gln374, Glu376, Gln377, Leu378, Thr447, Lys449, Asn450, Asn451, Ile452, and Trp454 was selected for the molecular docking study. Before the molecular docking study, the Tmprss2 was prepared using the Protein Preparation Wizard of Maestro. The Glide-grid module (Friesner et al., 2006) was used to generate the receptor grid enclosing the residues mentioned above to form a grid box.

Small molecular dataset

The structures of 5 therapeutically used drugs viz., Ambroxol (CAS: 18683-91-5), Bromhexine (CAS: 3572-43-8), Guaifenesin (CAS: 93-14-1), Camostat (CAS: 59721-29-8) and Nafamostat (CAS: 82956-11-4) were obtained from the ChEMBL Database (Gaulton et al., 2012). A set of 2230 compounds that belongs to the natural product dataset was obtained from the Selleckchem database (<https://www.selleckchem.com/screening/natural-product-library.html>) (Selleckchem, 2014). The library contains diverse chemical scaffolds derived from natural sources and with different biological properties. Compounds in the Selleckchem database are belong to diverse kinds of scaffolds ranging in alkaloids, phenols, terpenes and terpenoids, flavones, and flavonoids. All compounds are commercially available for high-throughput screening and other biological studies. To use these molecules for virtual screening, the entire dataset along with five drug molecules were prepared using the 'LigPrep' (Schrodinger, 2018) of Schrodinger molecular modelling suite. The 3D conformation of all the ligands was generated. The whole process was performed at a pH of 7.0 ± 2.0 with each molecule restricted to produce a maximum of 32 stereoisomers. The molecular chirality was retained and ionisation states were achieved using the 'Epik' function of the LigPrep. The geometric optimisation of ligands was done through the OPLS3 force field (Schrodinger, 2018).

Molecular docking and virtual screening workflow

Five standard drug molecules as mentioned previously were considered for the extra precision (XP) molecular docking using the Glide module of the Schrodinger suite (Bhowmick, Alissa, et al., 2020). The van der Waals radii and scaling factor were set to 0.80 and 0.15 respectively to soften the potential of nonpolar parts on drug molecules. No restraints were applied to the ligands during the entire docking protocol. Post-docking minimisation was allowed to provide minimised docked structures with a maximum of five best poses per ligand and to include in the docking output file. The drug molecules were allowed to be flexible during the docking process. The RMSD, docking score, Glide score, and binding energy were recorded for each molecule (Kerzare et al., 2016). The Virtual Screening Workflow (VSW) available on the Schrodinger suite is a grid-based ligand docking with energetics module (Friedrich et al., 2017; Jangam et al., 2019). The VSW was used for the rigorous screening, docking, and validation studies of the natural product dataset. The VSW protocol consists of three consecutive molecular docking steps, such as glide-high-throughput virtual screening (HTVS), standard precision (SP), and extra precision (XP) docking. The entire execution of the VSW process was performed in the CHPC server, Cape Town, South Africa (<https://www.chpc.ac.za/index.php/resources/lengau-cluster>). The prepared dataset of molecules was given as input in the VSW. The ligand cut off selected as 50% at each stage of HTVS, SP and XP. The good scoring poses retained and further subjected to the calculation of binding-free energy by the molecular

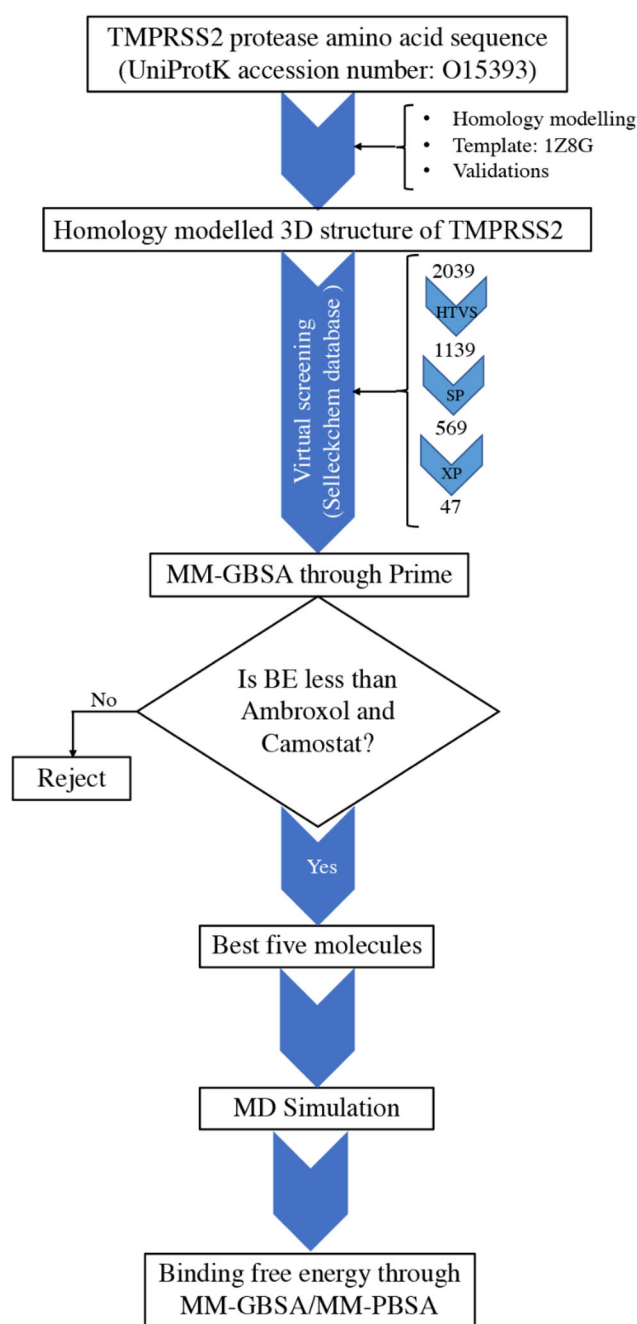


Figure 2. Flow diagram of homology modelling-based screening of TMPRSS2 inhibitors.

mechanics generalised born surface analysis (MM-GBSA) (Genheden & Ryde, 2015) and molecular mechanics Poisson–Boltzmann surface area (MM-PBSA) (Kumari et al., 2014) methods. The Glide-XP score, docked pose of the ligands and MM-GBSA binding energy was obtained from the VSW exercise were taken into consideration for further studies on the selected protein–ligand complexes.

MD simulation and MM-GBSA/MM-PBSA analysis

Final molecules obtained after successful screening through VSW were considered for MD simulation study in complex with TMPRSS2. In addition to these, above mentioned five standard drug molecules in complex with TMPRSS2 were

also studied for the all-atom unbiased molecular dynamics simulations on Amber18 software as mentioned previously. The protein–ligand complexes were immersed in truncated octahedron of TIP3P water and the required number of salt ions were added to neutralise the complex and then bringing the ionic concentration to 0.1N. Rest of the procedure remains the same for both the simulations. The protein–ligand complex RMSD, ligand RMSD and RMSF for protein residues was analysed using the CPPTRAJ module of Amber18 over full trajectory, taking configuration every 4 ps. The complete 100 ns MD simulation trajectories were subjected to calculate the binding-free energy through MM-GBSA and MM-PBSA analysis covering all the 10,000 frames. The results were tabulated, all energies are reported in Kcal/mol.

Results and discussion

In order to find out promising TMPRSS2 inhibitors, a number of *in-silico* approaches including homology modelling, virtual screening, molecular docking, MD simulation and binding-free energy calculation were performed. The flow diagram of the work is given in Figure 2.

Generation of the 3D structure of TMPRSS2

The crystal structure of TMPRSS2 is yet to be solved, hence the 3D coordinates of the same were generated through the homology modelling approach. The homology modelling was performed by a template-based method using the Swiss-Model Online Platform (Biasini et al., 2014). Swiss-Model uses ProMod3 as its engine. ProMod3 extracts structural information from an aligned template structure in Cartesian space. The sequence alignment of the query and target sequence identifies the insertions and deletions in the compared sequences. These problems are resolved by searching for the fragments in the structural database and the missing loops or chains are modelled and inserted. In the absence of a suitable fragment, a conformational space search is performed using Monte Carlo sampling (Koch, 2018). ProMod3 uses the 2010 backbone dependent rotamer library to solve the non-conserved side-chains. The rotamer configurations are estimated using the TreePack algorithm and energy minimisation is performed by SCWRL4 energy function (Waterhouse et al., 2018). It used the OpenMM library to perform computations and CHARMM27 forcefield for parametrisation. The target sequence of TMPRSS2 in FASTA format was collected from the Universal Protein Resource ‘UniProtKB’ database (Accession Number O15393). Several possible templates were retrieved by the Swiss-Modeller with the maximum identity and statistical *E* value. After critical analyses, finally, the crystal structure of serine protease hepsin (PDB ID: 1Z8G) (Herter et al., 2005) was adopted as a potential template. The maximum identity and similarity of the template were found to be 35.21 and 39.00%, respectively with the target sequence. The 3D-structure of the template molecule is given in Figure S1 (Supplementary Data). Three important parameters such as fewest restraint violations, lowest probability density

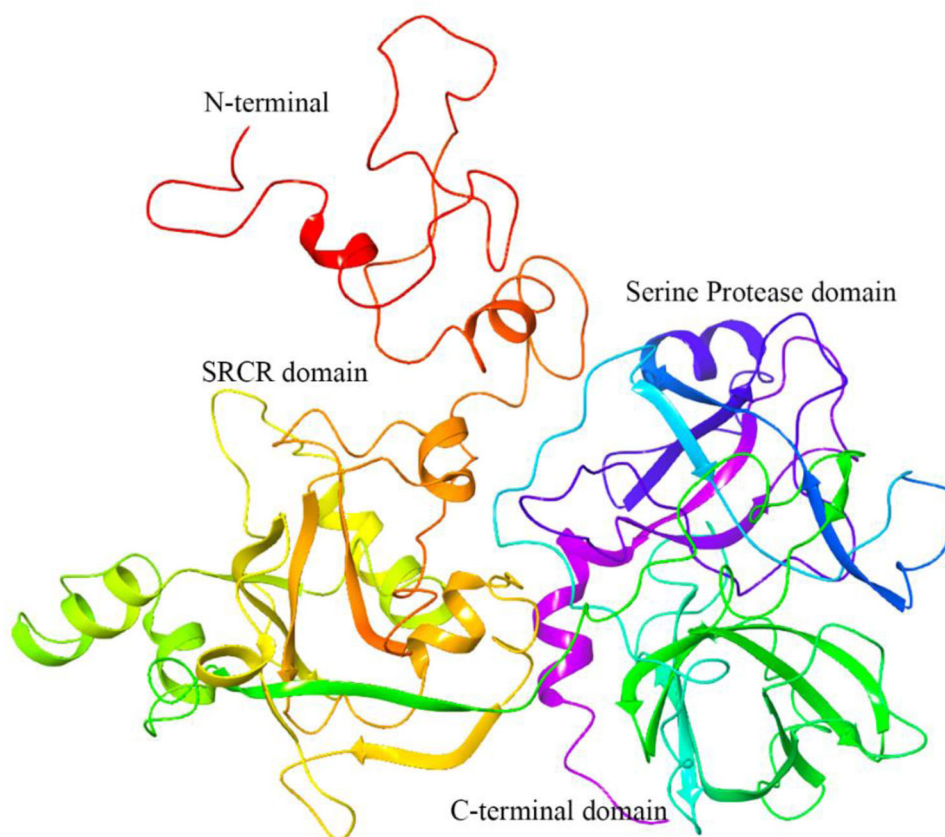


Figure 3. A three-dimensional model of TMPRSS2 developed through homology modelling.

function, and acceptable geometry were checked to select the best model. Based on the above criteria, finally, best 3D coordinates of generated TMPRSS2 were selected and it is given in Figure 3. The final selected model of TMPRSS2 was superimposed on the template molecule and the RMSD was found to be 1.255 Å which is within the acceptable range (permissible range i.e. ≤ 2 Å). The superimposed structure between the final selected model and template is given in Figure S2 (Supplementary Data). The low RMSD value undoubtedly explained the close homology and reliability of the model.

Validation and active site prediction

It is extremely important to validate any 3D protein structure generated from the template-based comparative modelling technique. First of all, the structure was considered for the refinement in the aqueous solvent and removal of the steric restraint through all-atoms MD simulation for the 100 ns period at a constant temperature of 300 K. The RMSD in the heavy atom positions converged to 6 Å after 40 ns and maintained its predicated overall protein fold for the rest of the simulation period (Figure S3, Supplementary Data). The final RMSF of individual amino acid shows high averages for residues between 150 and 200 with RMSF between 6 and 7 Å which could be attributed to the interdomain region of the TMPRSS2 (Figure S4, Supplementary Data). Further, the modelled TMPRSS2 structure was analysed by a series of validation methods to confirm its reliability and consistency. The

model was subjected to analysis using various tools including PDBsum (Laskowski, 2009), MOE, PROSA (Wiederstein & Sippl, 2007), QMEAN and LDDT (Mariani et al., 2011). The TMPRSS2 was first analysed for its stereochemical quality using the PDBsum and MOE software package. The model has the following structural features; 49 β -turns, 16 α -helices, 15 strands, 8 β -hairpins, 5 β -bulges, 4 β -sheets, 4 helix-helix interactions, and 12 γ -turn (Figure 3). The analysis of steric problems within the protein and high-accuracy Ramachandran Plot (RC-plot) was obtained and it is given in Figure S5 (Supplementary Data). The ϕ/ψ angles were found well within the allowed region with the majority of residues (90.61%) present in the favoured and allowed region. The molprobit score and clash score of all atoms were found to be 1.60 and 2.83, respectively. The rotamer outlier value of 0.24% accounting for one outlier rotamer. The Molprobit score and RC plot assessment suggest good structural integrity of the model. To understand the energetics of the model, a Qualitative Model Energy analysis (QMEAN) was performed through the QMEANDisCo tool. The QMEAN score was found to be -2.45 , which is considered good for a homology model. The TMPRSS2 model possesses two distinctive domains with a connecting region, a common feature of the TTSP family. The above distinctive domains can be recognised as the more regularly packed serine protease domain (Domain 1) and the less organised scavenger receptor cysteine-rich (SRCR) region (Domain 2). The residues from this interdomain connecting region form a groove that was detected as the most suitable site for docking studies out of

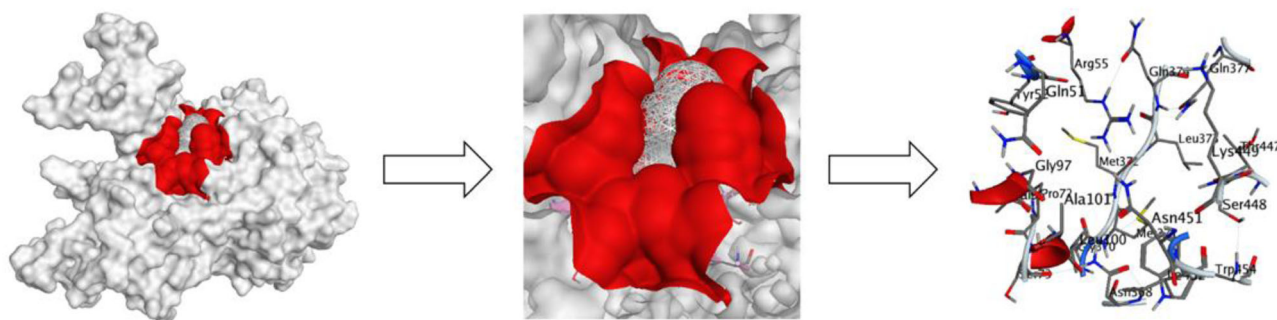


Figure 4. Position of binding pocket in the TMPRSS2.

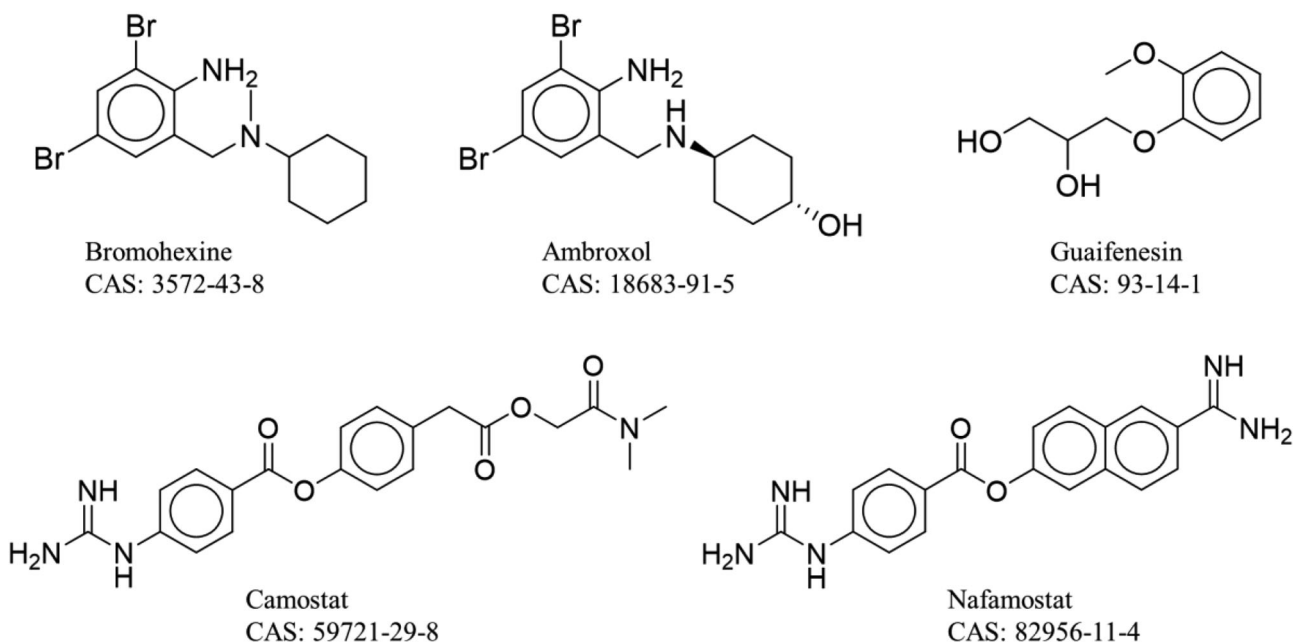


Figure 5. Two-dimensional representation of five standard drug molecules.

several possible active sites. This is unusual as in most of the serine proteases the binding site is usually around domain 1 (Somoza et al., 2003). Before molecular docking, it is crucial to find the active site in the protein molecule. The active site in the selected homology model was explored through two different methods such as Site-map in Schrodinger suite and Site-finder in MOE suite. Both approaches suggested that the position of the action in the interdomain cleft. On detailed analysis, the important amino residues confining the binding site were found to be Hie18, Gln21, Glu23, Asn24, Pro25, Val49, Pro50, Gln51, Tyr52, Ala53, Pro54, Arg55, Gln59, Val65, Gln68, Pro69, Val96, Gly97, Ala98, Ala99, Ala101, Met371, Met372, Leu373, Gln374, Glu376, Gln377, Leu378, Thr447, Lys449, Asn450, Asn451, Ile452, and Trp454. The position of the binding site cavity in 3D orientation is given in Figure 4.

Virtual screening

The availability of 3D coordinates of any target can be used in structure-based virtual screening for promising molecules. It has become popular in the scientific community and widely used in drug discovery research. The TMPRSS2

structure developed from the comparative structure prediction method was further used to screen a set of 2230 molecules belong to the natural products obtained from the Selleckchem database. It is essential to use any control molecule(s) to compare the outcomes of the pharmacoinformatics study. In this regard, a total of five standard molecules; Ambroxol (CAS: 18683-91-5; Approved drug), Bromhexine (CAS: 3572-43-8; Approved drug), Guaifenesin (CAS: 93-14-1; Approved drug), Camostat mesylate (CAS: 59721-29-8; Phase-I drug) and Nafamostat (CAS: 82956-11-4; Phase-III) were initially explored through the molecular docking and MD simulation study in the TMPRSS2. Two-dimensional (2D) representation of all five standard molecules is given in Figure 5.

The binding interactions of Bromhexine, Guaifenesin, and Nafamostat are given in Figure S6 (Supplementary Data), whereas, binding interactions of Camostat and Ambroxol are discussed in the following section. The Glide-XP score, Prime-MM-GBSA binding-free energy, and binding-free energy obtained from the MD simulation trajectory were critically explored and compared to each other (Table S1, Supplementary Data). Finally, based on the above parameters and known pharmacological properties, Camostat and Ambroxol were considered as control molecules in the

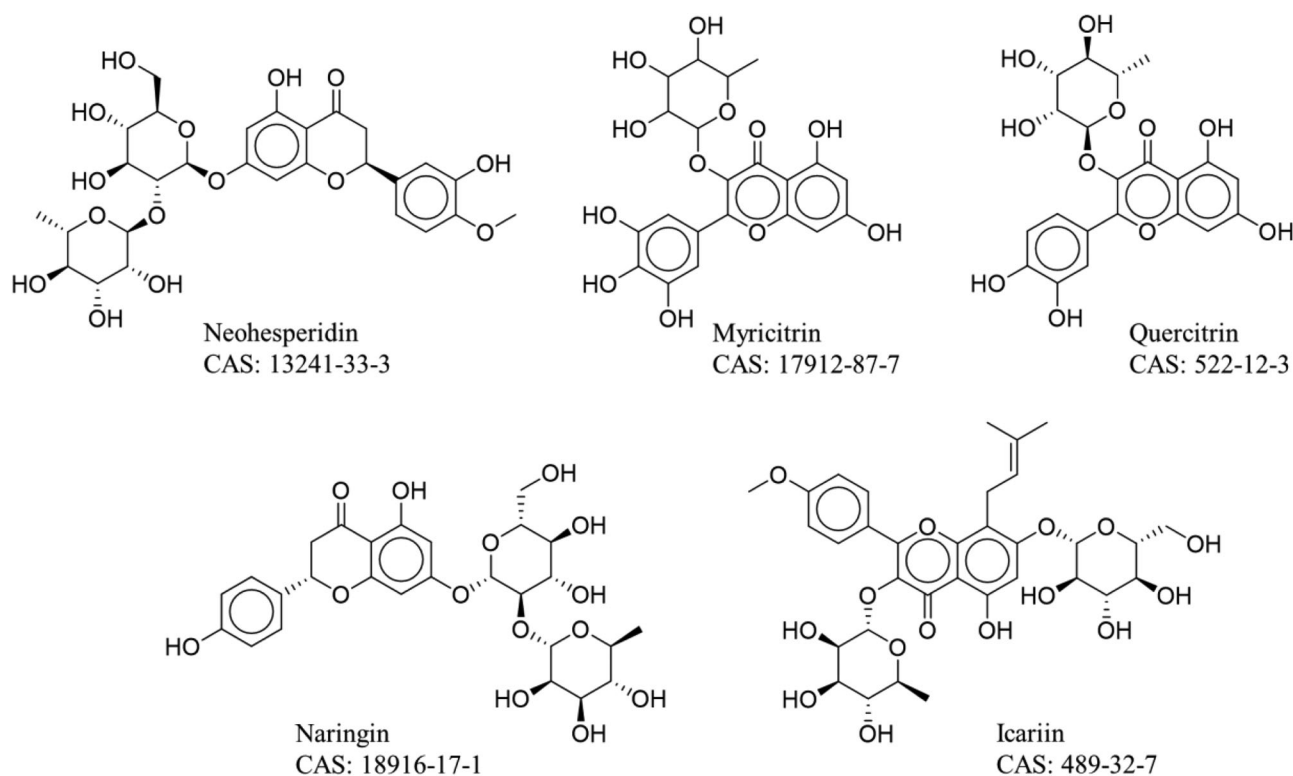


Figure 6. Two-dimensional representation of final proposed molecules.

remaining study. The Glide XP-score in molecular docking, all parameters in MD simulation study, and binding interaction profile from the post-MD simulations of all proposed compounds were compared with the above two control molecules.

The curated Selleckchem database was used for screening through a multi-step molecular docking study implemented in the VSW module of Maestro. Initially, the entire curated database was docked through the Glide-HTVS and best 50% docked molecules retained for the next step. Molecules remained in the Glide-HTVS step were used as input of the Glide-SP procedure. In this step also best 50% docked molecules were kept for further analysis. In the final step of molecular docking (glide-XP), the best 25% molecules were considered to calculate the binding-free energy through the Prime-MM-GBSA approach. The XP-score and Prime-MM-GBSA binding energy of Camostat and Ambroxol was found to be -6.231 and -30.123 Kcal, and -7.212 and -34.654 Kcal/mol, respectively. Hence, the molecules having higher XP score and binding energy compared to Ambroxol were considered for the next level of analysis. After following the above-mentioned procedure, finally, a total of 15 molecules satisfied the above criteria and used for the MD simulation study. The binding-free energy through the MM-GBSA approach from the MD simulation trajectory was calculated for all 15 molecules along with Camostat and Ambroxol. Among the five know drugs studied, Ambroxol had the highest binding energy of -44.48 Kcal/mol, and Camostat had medium binding energy of -25.00 Kcal/mol. The above values were used further as a threshold to reduce the chemical space. Finally, by fulfilling the above all criteria a total of five molecules; Neohesperidine, Myricitrin, Quercitrin, Naringin,

and Icariin were selected out of the top ten phytochemicals were selected as crucial inhibitors of TMPRSS2. The 2D representation of the final five phytochemicals is given in Figure 6.

Binding interactions analyses

The binding interaction profile of the final five phytochemicals along with two standard drug molecules was analysed through the protein–ligand interaction profiler (PLIP) (Figure 7) and the binding interactions for the rest of the molecules are mentioned in the supplementary information (Figures S7 and S8, Supplementary Data). The GScore and DockScore of final proposed molecules along with standard drug molecules are given in Table 1. Several potential hydrogen bonds (HB) and non-bonding interactions were observed between the phytochemicals and binding site amino residues of TMPRSS2 protein. Hesperetin 7-neohesperidoside, also known as neohesperidin, belongs to the class of organic compounds known as flavonoid-7-o-glycosides. These are phenolic compounds containing a flavonoid moiety which is O-glycosidically linked to carbohydrate moiety at the C7-position. Glucosyl hesperidin prevents influenza virus replication *in vitro* by inhibition of viral sialidase (Saha et al., 2009). Neohesperidin docks nicely into the active site of TMPRSS2, its pyran ring with three hydroxyl groups were found to be critical to interact with Arg55, Gly97, and Asn51 through 2, 2, and 1 HB interactions respectively. Another pyrin ring with two hydroxyl and methoxy functional groups in Neohesperidin established HB interactions with Gln 377, Thr477, Asn451, and Ile452 by 1, 1, 1 and 2, respectively.

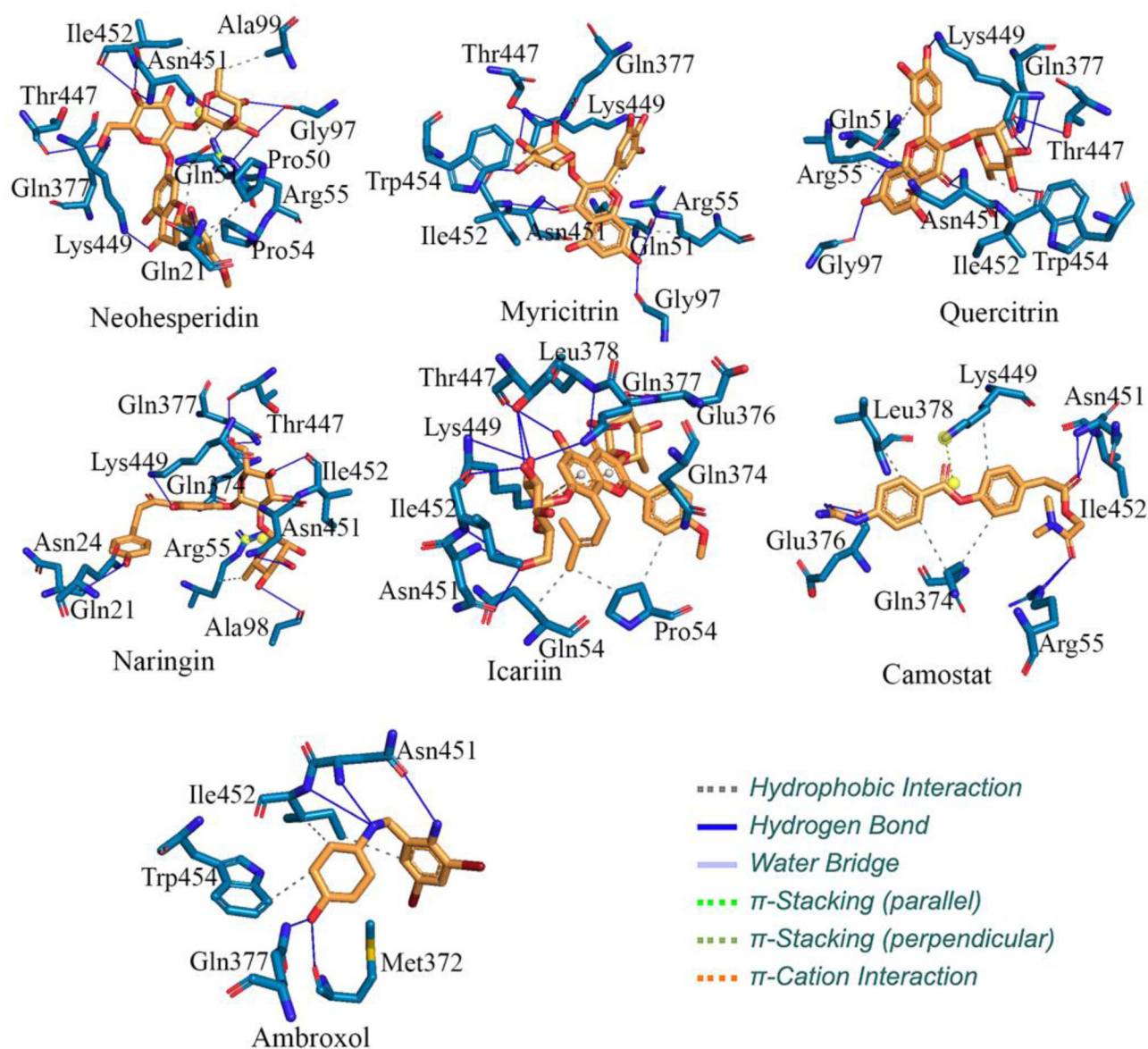


Figure 7. Binding interactions of final phytochemicals with TMPRSS2.

Table 1. Molecular docking score and Glide score obtained from docking studies, Prime-MM-GBSA and binding-free energy components for the protein–ligand complexes calculated by MM-GBSA and MM-PBSA analysis for the selected compounds (all energies are in Kcal/mol with standard deviation in parenthesis).

Molecule name	Virtual screening workflow (VSW)			Molecular dynamics simulations (MDS)	
	GScore	DockScore	Prime MM-GBSA (ΔG_{bind})	MM-GBSA (ΔG_{bind})	MM-PBSA (ΔG_{bind})
Natural products					
Neohesperidin	-12.77	-12.77	-66.53	-72.84 (7.73)	-15.83 (8.82)
Myricitrin	-11.52	-11.55	-51.70	-49.52 (3.56)	-7.28 (4.83)
Quercitrin	-10.78	-10.81	-55.01	-50.57 (4.73)	-14.61 (6.90)
Naringin	-10.73	-11.25	-37.11	-50.29 (5.28)	7.45 (7.57)
Icariin	-10.10	-10.11	-56.83	-60.34 (4.26)	-9.52 (8.83)
Standard drugs					
Camostat	-6.23	-4.62	-25.00	-25.00 (2.67)	11.99 (4.44)
Ambroxol	-7.21	-3.61	-44.48	-44.48 (4.94)	-11.36 (7.59)

ΔG_{bind} =binding-free energy.

Beyond above, several hydrophobic and π -stacking binding interactions were observed between the Neohesperidin and Pro50, Gln51, Ala99, and Ile452 amino residues of TMPRSS2 protein. Myricetin is a common plant-derived flavonoid and is well recognised for its nutraceuticals value. It

is one of the key ingredients of various foods and beverages (Semwal et al., 2016). The compound exhibits a wide range of activities that include strong antioxidant, anti-cancer, antidiabetic, and anti-inflammatory activities. Myricetin-3-O-rhamnoside showed antiviral activity and

mechanism of action against H1N1 influenza virus (Motlhatlego et al., 2018). The binding interactions between Myricitrin and TMPRSS2 were explored. In Figure 7, it can be seen that Glu377, Thr447, Lys449, and Ile452 were revealed as crucial amino residues to form strong HB interactions with Myricitrin.

One of the hydroxyl groups of pyrogallol in the molecular system potentially formed HB interaction with Lys449. The oxo and hydroxyl functional groups present in the chromone ring in Myricitrin were found to form HB interactions with Asn451, Ile452, and, Arg55 and Gly97, respectively. Few of the above amino residues were also observed to participate in non-HB interactions. Quercetin-3-rhamnoside (Q3R) possessed antiviral activity against the influenza A/WSN/33 virus *in vitro*. Mice orally treated with Q3R (6.25 mg/kg per dose) at 2 h before and once daily for 6 days after influenza virus infection showed significant decreases in weight loss, and decreased mortality (Choi et al., 2012; Wong et al., 2017). Quercetin, dihydroquercetin, and quercitrin are flavonoids (flavanols) found in various vegetarian foods including onions and leaf cabbage (Micronutrient Information Center, 2017). The molecular docking interaction between Quercitrin and TMPRSS2 reveals the pyrrole ring of Quercitrin as the most significant part of the molecules because it participates in forming a number HB and one hydrophobic interaction. Amino residues, Gln377, Thr447, Lys449, and Ile 452 formed 1, 1, 2, and 1 HB interactions respectively with Quercitrin. The Methyl group attached in the same ring was crucial to establish hydrophobic interaction with Trp454. Hydroxyl group attached to the independent phenyl ring clashed with Lys 449 through HB and the same phenyl ring also formed a hydrophobic interaction with Gln51. Moreover, the oxy group present in the fused double ring system was found to be important to form HB interactions with Asn451 and Ile452. Naringin is flavanone-7-O-glycoside found widely in the citrus fruits, especially in the grapefruit (Alam et al., 2014). It is reported to impairs the *in vitro* infection of human cells by Zika virus and the dengue virus (Cataneo et al., 2019; Frabasile et al., 2017). Naringin was docked inside the receptor cavity of TMPRSS2. A significant number of crucial amino acids, namely Gln21, Asn24, Ala98, Gln374, Gln377, Thr447, Asn451, and Ile452 established strong HB interactions with Naringin. Moreover, Arg55 shows two critical hydrophobic binding interactions with Naringin. Icariin is prenylated flavonol glycoside, it is widely found in the genus *Epimedium*. These classes of plants are widely used in Chinese traditional medicine as an aphrodisiac (Liu et al., 2006). Icariin and its phosphorylated derivatives were reported for their antiviral activity against duck hepatitis virus A (Xiong et al., 2015). The molecular docking interactions of Icariin with TMPRSS2 show that the hydroxyl and hydroxymethyl groups present in the tetrahydroxy pyran ring in the Icariin molecule make hydrogen bonds with the residues, Gln377, Thr447, Lys449, Asn451, and Ile452. In addition to the above, Glu376 and Leu378 also formed HB interactions with Icariin. Moreover, some non-hydrogen bond interactions were also observed between Icariin and binding site amino residues of TMPRSS2 protein. Ambroxol is an active metabolite of bromhexine and

superior to bromhexine is in its pharmacokinetics, efficacy, and side effect profile (NCT02914366, 2016; Weiser, 2008). Ambroxol has been used in the prevention of neonatal respiratory distress syndrome with no reported maternal or foetal/neonatal side effects. Ambroxol stimulated the release of suppressors of influenza-virus multiplication (Yang et al., 2002). Camostat is a known inhibitor of TMPRSS2 as mentioned earlier. Detailed binding interactions of Camostat and Ambroxol with TMPRSS2 were critically examined.

Hydrogen bond interaction with residue Asn451 was common for both drugs. Binding site amino residues, Arg55, and Glu376 were found crucial to interact with Camostat via HB interaction, while Met372, Glu377, and Ile452 established HB contacts with Ambroxol. Moreover, Gln374, Leu378 and Lys449, and, Ile452 and Trp454 formed hydrophobic interactions with Camostat and Ambroxol respectively. Overall, some binding site amino residues formed strong HB and hydrophobic interactions with the final selected phytochemicals. Such an observed high number of binding interactions plays a crucial role to retain the stability between TMPRSS2 and phytochemicals. The binding mode of each of phytochemicals in the TMPRSS2 receptor cavity was also explored and given in Figure S9 (Supplementary Data). It can be seen that all the final molecules fitted perfectly in the active site with optimal orientation.

Kumar et al. (2020) recently considered the homologous structure of TMPRSS2 from the Swiss-Model Repository. Further, they have adopted three natural small molecules, Withanone and Withaferin-A from *Ashwagandha*, and caffeic acid phenethyl ester (CAPE) from honey bee propolis for molecular docking followed by MD simulation studies. Camostat was used as a control molecule. In the current study, the molecular docking study revealed a number of ligand and binding amino acid residues those are common with the molecular docking outcomes by Kumar et al. It is also important to note that both Withanone and Withaferin-A, as well as Camostat, were perfectly fitted inside the active site of TMPRSS2. In our study, all proposed phytochemical along with both standard molecules, Camostat and Ambroxol were also fitted in the TMPRSS2. Moreover, MD simulation study was performed by Kumar et al. and shown the relative stability of the protein–ligand complex. In the current study, we have also successfully shown relative stability of the TMPRSS2-phytochemical complexes and strong binding affinity towards the active site.

Molecular dynamics simulation

MD simulation analysis is a crucial and effective study to explore the dynamic behaviour of the molecules. To explore the nature of phytochemicals inside the TMPRSS2 in dynamic states and the stability of protein–ligand complex, all-atoms 100 ns MD simulation was carried out. RMSD of the protein backbone, RMSF, ligand RMSD and binding-free energy were explored. The protein backbone RMSD of TMPRSS2 bound with phytochemicals was calculated from the MD simulation trajectories and given in Figure 8. The RMSD parameter is one of the crucial parameters which can explain the overall

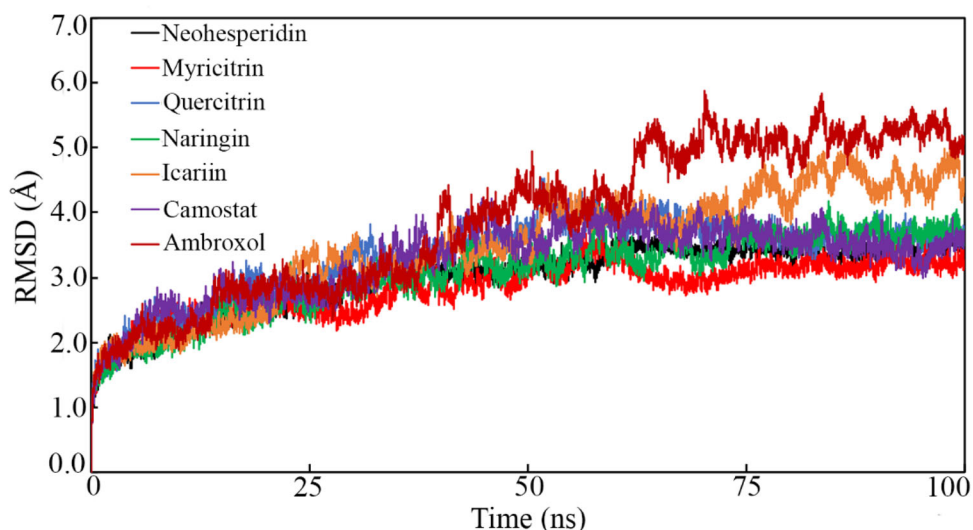


Figure 8. RMSD of TMPRSS2 backbone bound with phytochemicals.

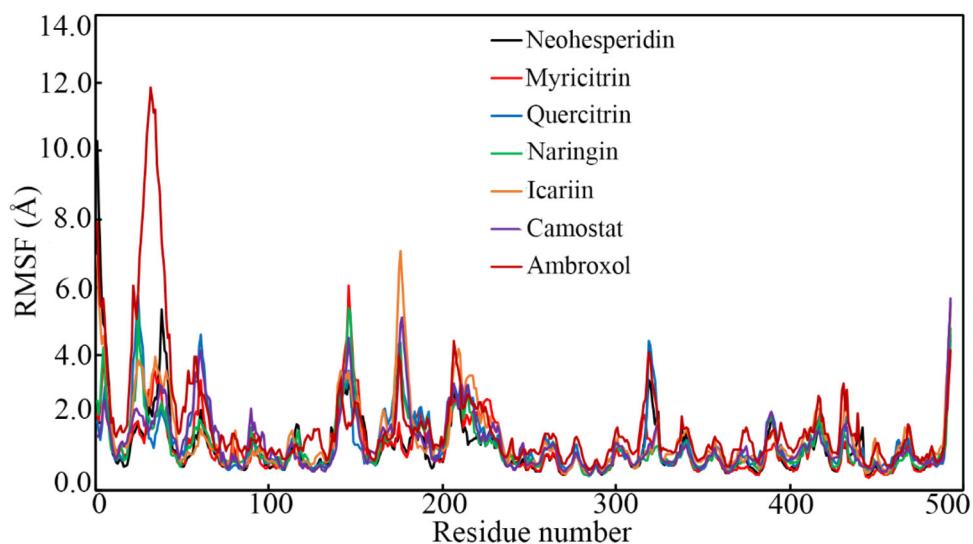


Figure 9. RMSF values of TMPRSS2 amino acids bound with proposed phytochemicals.

stability and insights into its structural conformation changes of the protein–ligand complex from the MD simulation trajectories. The high stability of the protein–ligand complex can be deduced by obtaining the lower RMSD value of the protein backbone. The backbone RMSD with low value is perfectly acceptable but larger variations are also acceptable as it might indicate that the protein is probably undergoing large or some sort of conformational change during the simulation. It can be seen that except TMPRSS2 backbone bound with Icariin and Ambroxol all backbone RMSD remained consistent throughout the MD simulation. For all complexes, the RMSD was gradually increased from starting and equilibrated after about 55 ns. In the case of TMPRSS2 backbone bound with Icariin and Ambroxol, it was observed that after about 60 ns the RMSD values increases but immediately equilibrated around 4 and 5 Å, respectively. The average RMSD value of each complex gives an idea about the relative stability. Average RMSD of TMPRSS2 backbone bound with Neohesperidin, Myricitrin, Quercitrin, Naringin, Icariin, Camostat, and Ambroxol was found to be 3.010,

2.844, 3.337, 3.070, 3.509, 3.265 and 3.909 Å, respectively. It is important to note that the average RMSD of all complexes were found within the range of 2.800 to 4.00 Å. Also, TMPRSS2 backbone RMSD of final compounds was found to be less than Ambroxol bound TMPRSS2 backbone. Therefore, from the above data and observation, it was clear that complexes between TMPRSS2 and proposed molecules remained stable in the entire molecular dynamics simulation.

Individual amino residues play a crucial role to retain the stability of the protein–ligand complex in the MD simulation study. The fluctuation of each amino acid can be calculated through the RMSF which measures the mean divergence of each amino acid from the reference position. The absolute magnitude of the RMSF value indicates high or low flexibility. On successful completion of the MD simulation, the RMSF of TMPRSS2 amino acids was calculated and it is given in Figure 9. Figure 9 explains the pattern of changes in RMSF of amino acids bound with selected phytochemicals were almost similar with some exception. The average RMSF values were found to be 1.269, 1.244, 1.414, 1.296, 1.528, 1.373,

and 1.945 Å for the complex with Neohesperidin, Myricitrin, Quercitrin, Naringin, Icariin, Camostat, and Ambroxol correspondingly. Low mean RMSF value undoubtedly explained the stability of the amino acid during MD simulation.

The deviation of the ligand from the native structure was explored by the ligand RMSD analysis. The ligand RMSD of each frame was obtained and plotted against the time simulation (Figure S10, [Supplementary Data](#)). It can be seen that all the final proposed phytochemicals were equilibrated with low RMSD values. It is also crucial to note that both drug molecules, Camostat and Ambroxol show higher deviation compared to their native conformation and equilibrated with higher RMSD value. The lower RMSD of phytochemicals compared to standard drugs shows higher affinity and better binding to the target protein.

The hydrogen bond analysis for all the seven protein–ligand complexes was performed by CPPTRAJ in Amber18. The hydrogen bonds at the beginning and end of the simulation were observed and recorded by calculating the actual distance between the receptor–donor pairs in the protein–ligand complexes. The hydrogen bond interactions formed by residues Gly97, Gln377, Asn451 and Ile452 were found to remain intact and between 2.5 and 3.8 Å most instants (Figure S10, [Supplementary Data](#)). The hydrogen bond analysis suggests high affinity between the ligands and the residues in the active site of the TMPRSS2 receptor.

To check the binding affinity of the proposed molecules towards TMPRSS2 the entire trajectories from MD simulation were considered to calculate the binding-free energy (ΔG_{bind}) through the MM-GBSA approach. The MM-GBSA and MM-PBSA are widely used approaches by the scientific community to calculate ΔG_{bind} from the MD simulation trajectories and also can be considered almost accurate in comparison to molecular docking-based binding energy (Bhowmick, AlFaris, et al., 2020; Genheden & Ryde, 2015; Hou et al., 2011). The ΔG_{bind} along with other binding energy contributing components such as Van Der Waals interactions energy (ΔG_{vdW}), electrostatics energy (ΔG_{ele}), polar solvation energy ($\Delta G_{\text{ele,sol}}$), and nonpolar solvation energy ($\Delta G_{\text{nonpol,sol}}$), these were calculated. The contribution of the solvation-free energies in the binding-free energy is crucial for the binding process. This parameter explains the loss of water interactions for a small molecule moves from solution to the gas phase in a similar way as a ligand moving from solvent into a binding site (Cournia et al., 2017). The ΔG_{bind} of final five molecules calculated using MM-GBSA and MM-PBSA are given in [Table 1](#). The value of the individual contributing component of ΔG_{bind} obtained using MM-GBSA and MM-PBSA are given in [Tables S1 and S2](#) ([Supplementary Data](#)), respectively. The high negative ΔG_{bind} value of any molecule illustrated the more affinity towards the receptor. From [Table 1](#), it can be seen that ΔG_{bind} calculated through MM-GBSA method was found to be -72.840 , -49.520 , -50.570 , -50.290 , -60.340 , -25.000 and -44.480 Kcal/mol for Neohesperidin, Myricitrin, Quercitrin, Naringin, Icariin, Camostat, and Ambroxol, respectively. The ΔG_{bind} using MM-PBSA approach was found to be -15.83 , -7.28 , -14.61 , 7.45 and -9.52 Kcal/mol for Neohesperidin, Myricitrin, Quercitrin,

Naringin and Icariin, respectively. All proposed phytochemicals exhibited a higher affinity towards the TMPRSS2 in comparison to Camostat and Ambroxol. Moreover, the main contributing components of total binding-free energy were found to be ΔG_{vdW} and ΔG_{ele} which indicated the possibilities of forming several hydrophobic and electrostatic interactions between the TMPRSS2 and phytochemicals. Hence, the above observation demonstrates that all proposed molecules might inhibit the TMPRSS2.

It is quite interesting and essential to explore the binding interactions between protein and ligand after successful completion of the MD simulation study. This analysis can give an insight into the binding interactions and orientation of the molecules relative to the molecular docking study. The last frame of the protein–ligand complex of each molecule was extracted from the MD trajectory. Subsequently, the binding interactions profile was explored through PLIP and given in [Figure 10](#). After deep exploration, it was found that all proposed molecules and standard drugs were successfully retained inside the receptor cavity of TMPRSS2 with different orientation and a small alteration of binding interactions. Most of the interacting amino acids in molecular docking studies were found to retain interaction in post-MD complexes. It is also worth noting that very few binding interactions were lost but also new contacts established during the simulation. Overall, proposed natural compounds were found excellently efficient enough to keep hold at the active site of TMPRSS2 and also retain some potential binding interactions in the dynamic states during the all-atoms MD simulation.

Targeting the host proteases in viral infections is a relatively new strategy for developing antiviral drugs that may overcome the emergence of resistance. Effective antiviral drugs against pathogenic coronaviruses and other emerging viruses are desirable to enable a rapid response to pandemic threats. TMPRSS2 is one of the host factors that has been shown essential for proteolytic activation and consequently spread and pathogenesis of coronaviruses. Considering its important role in the viral entry of SARS-CoV-2 on binding with ACE-2, it is a promising drug target for the treatment of these viral infections. The final compounds from this study belong to the class of flavonoids also known as bioflavonoids. The bioflavonoids occur widely in food and herbs used in Ayurvedic and Traditional Chinese Medicine (TCM) for the treatment of infectious and non-infectious diseases since ancient times. This work also indicates that nutritious food including green vegetables, herbs and fruits would help in improving our immunity to fight the COVID-19 owing to the presence of these phytochemicals.

Conclusion

To identify a few crucial COVID-19 therapeutics from the natural products, homology modelling-based pharmacoinformatic approaches were used on TMPRSS2. The lack of available TMPRSS2 crystal structure encouraged us to generate the 3D coordinates which were obtained through homology modelling followed by extensive virtual screening of natural products. The natural product database was screened

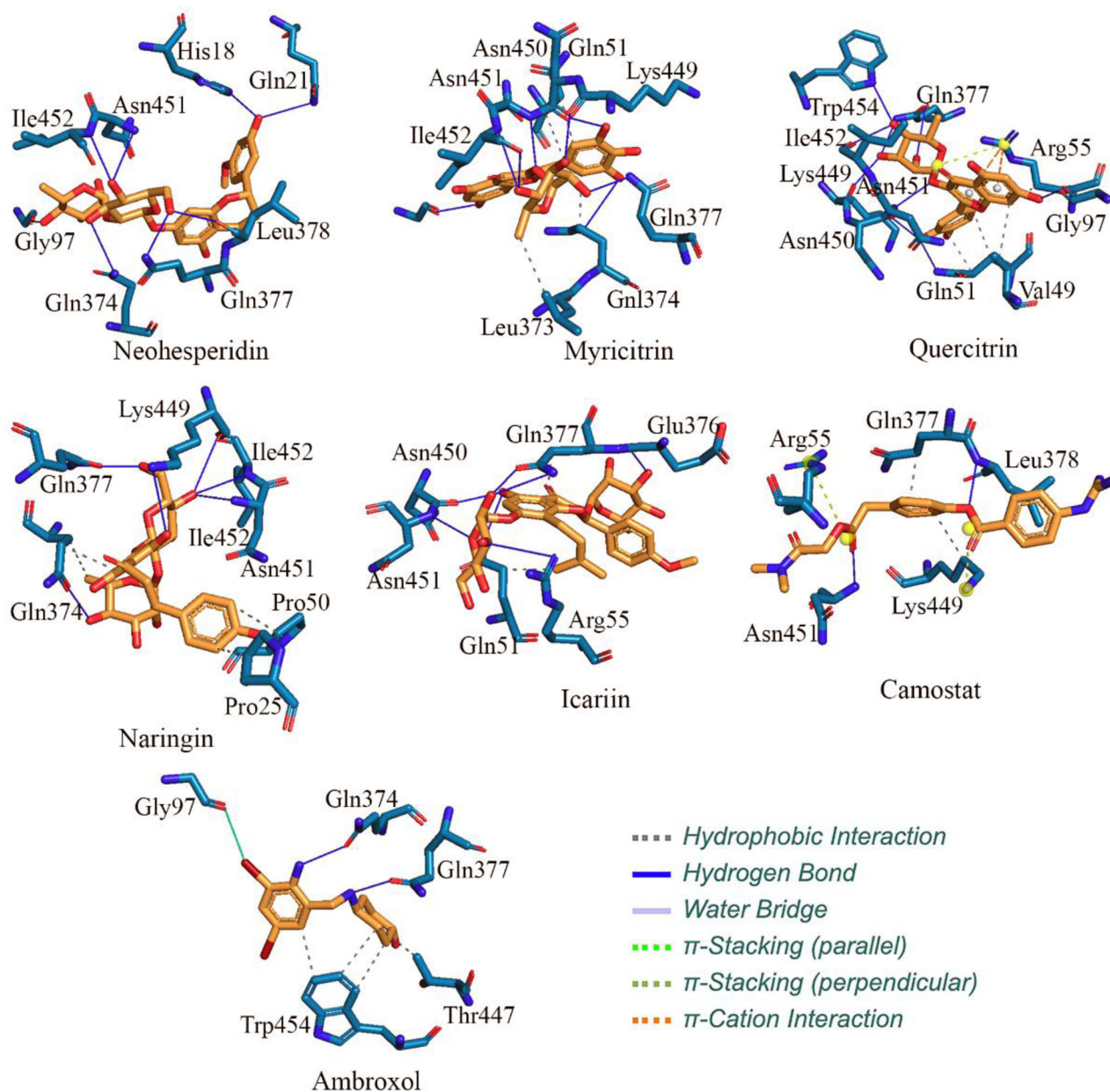


Figure 10. Post-MD simulation binding interactions analysis between phytochemicals and TMPRSS2.

at three levels of molecular docking followed by molecular dynamics simulation of ten final natural compounds along with five standard drug molecules. The free energy calculation for these fifteen molecules was performed and further analysis was carried out. The top five natural products Neohesperidin, Myricitrin, Quercitrin, Naringin, Icariin, and two therapeutically used drugs, Camostat, and Ambroxol were discussed in detail. The molecular docking, MD simulation, and binding-free energy of each molecule were derived through the MM-GBSA approach shows that all compounds possess a strong affinity towards the TMPRSS2. These natural compounds are phytochemicals belongs to the class of bioflavonoids and three of them Neohesperidin, Myricitrin, Quercitrin are known for their antiviral activity. This study cements the previous reports and provides a possible mechanism of action of these compounds for antiviral activity. The

inhibition of TMPRSS2 would inhibit the SAR-CoV-2 viral entry into the human host cells. Hence, it can be concluded that proposed molecules through the pharmacoinformatic approaches might be crucial biomolecules in the successful management of COVID-19, subject to experimental validation.

Disclosure statement

The authors declare that there is no competing interest.

Funding

The authors are grateful to the Researchers Supporting Project No. (RSP-2020/161), King Saud University, Riyadh, Saudi Arabia.

Computational resources

The CHPC (www.chpc.ac.za), Cape Town, South Africa is thankfully acknowledged for computational resources and tools.

ORCID

Rupesh V. Chikhale  <http://orcid.org/0000-0001-5622-3981>

Md Ataul Islam  <http://orcid.org/0000-0001-6286-6262>

References

- Alam, M. A., Subhan, N., Rahman, M. M., Uddin, S. J., Reza, H. M., & Sarker, S. D. (2014). Effect of citrus flavonoids, naringin and naringenin, on metabolic syndrome and their mechanisms of action. *Advances in Nutrition (Bethesda, MD)*, 5(4), 404–417. <https://doi.org/10.3945/an.113.005603>
- Albrecht, H. H., Dicipinigaitis, P. V., & Guenin, E. P. (2017). Role of guaifenesin in the management of chronic bronchitis and upper respiratory tract infections. *Multidisciplinary Respiratory Medicine*, 12, 31. <https://doi.org/10.1186/s40248-017-0113-4>
- Andersen, H. C. (1983). Rattle: A “velocity” version of the shake algorithm for molecular dynamics calculations. *Journal of Computational Physics*, 52(1), 24–34. [https://doi.org/10.1016/0021-9991\(83\)90014-1](https://doi.org/10.1016/0021-9991(83)90014-1)
- Bahgat, M. M., Błazejewska, P., & Schughart, K. (2011). Inhibition of lung serine proteases in mice: A potentially new approach to control influenza infection. *Virology Journal*, 8, 27. <https://doi.org/10.1186/1743-422X-8-27>
- Benkert, P., Biasini, M., & Schwede, T. (2011). Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics (Oxford, England)*, 27(3), 343–350. <https://doi.org/10.1093/bioinformatics/btq662>
- Benkert, P., Künzli, M., & Schwede, T. (2009). QMEAN server for protein model quality estimation. *Nucleic Acids Research*, 37(suppl_2), W510–W514. <https://doi.org/10.1093/nar/gkp322>
- 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>
- Bhowmick, S., AlFaris, N. A., ALTamimi, J. Z., ALOthman, Z. A., Aldayel, T. S., Wabaidur, S. M., & Islam, M. A. (2020). Screening and analysis of bioactive food compounds for modulating the CDK2 protein for cell cycle arrest: Multi-cheminformatics approaches for anticancer therapeutics. *Journal of Molecular Structure*, 1216, 128316. <https://doi.org/10.1016/j.molstruc.2020.128316>
- Bhowmick, S., Alissa, S. A., Wabaidur, S. M., Chikhale, R. V., & Islam, M. A. (2020). Structure-guided screening of chemical database to identify NS3–NS2B inhibitors for effective therapeutic application in dengue infection. *Journal of Molecular Recognition*, 33(7), e2838. <https://doi.org/10.1002/jmr.2838>
- Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T. G., Bertoni, M., Bordoli, L., & Schwede, T. (2014). SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research*, 42(W1), W252–W258. <https://doi.org/10.1093/nar/gku340>
- Bugge, T. H., Antalis, T. M., & Wu, Q. (2009). Type II transmembrane serine proteases. *Journal of Biological Chemistry*, 284(35), 23177–23181. <https://doi.org/10.1074/jbc.R109.021006>
- Cataneo, A. H. D., Kuczera, D., Koishi, A. C., Zanluca, C., Silveira, G. F., Arruda, T. B., de, Suzukawa, A. A., Bortot, L. O., Dias-Baruffi, M., Verri, W. A., Robert, A. W., Stimamiglio, M. A., Duarte dos Santos, C. N., Wowk, P. F., & Bordignon, J. (2019). The citrus flavonoid naringenin impairs the *in vitro* infection of human cells by Zika virus. *Scientific Reports*, 9(1), 16348. <https://doi.org/10.1038/s41598-019-52626-3>
- Chen, X., Yang, Z., Lu, Y., Xu, Q., Wang, Q., & Chen, L. (2013). Clinical features and factors associated with outcomes of patients infected with a novel influenza A (H7N9) virus: A preliminary study. *PLoS One*, 8(9), e73362. <https://doi.org/10.1371/journal.pone.0073362>
- Chen, Y. W., Lee, M. S., Lucht, A., Chou, F. P., Huang, W., Havighurst, T. C., Kim, K., Wang, J. K., Antalis, T. M., Johnson, M. D., & Lin, C. Y. (2010). TMPRSS2, a serine protease expressed in the prostate on the apical surface of luminal epithelial cells and released into semen in prostasomes, is misregulated in prostate cancer cells. *American Journal of Pathology*, 175(6), 2986–2996. <https://doi.org/10.2353/ajpath.2010.090665>
- Choi, H. J., Song, J. H., & Kwon, D. H. (2012). Quercetin 3-rhamnoside exerts antiinfluenza A virus activity in mice. *Phytotherapy Research*, 26(3), 462–464. <https://doi.org/10.1002/ptr.3529>
- Cournia, Z., Allen, B., & Sherman, W. (2017). Relative binding free energy calculations in drug discovery: Recent advances and practical considerations. *Journal of Chemical Information & Modeling*, 57(12), 2911–2937. <https://doi.org/10.1021/acs.jcim.7b00564>
- Dobos, G., & Tao, I. (2011). The model of Western Integrative Medicine: The role of Chinese medicine. *Chinese Journal of Integrative Medicine*, 17(1), 11–20. <https://doi.org/10.1007/s11655-011-0601-x>
- Frabasile, S., Koishi, A. C., Kuczera, D., Silveira, G. F., Verri, W. A., Dos Santos, C. N. D., & Bordignon, J. (2017). The citrus flavanone naringenin impairs dengue virus replication in human cells. *Scientific Reports*, 7, 41864. <https://doi.org/10.1038/srep41864>
- Friedrich, N. O., De Bruyn Kops, C., Flachsenberg, F., Sommer, K., Rarey, M., & Kirchmair, J. (2017). Benchmarking commercial conformer ensemble generators. *Journal of Chemical Information & Modeling*, 57(11), 2719–2728. <https://doi.org/10.1021/acs.jcim.7b00505>
- 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>
- Gaulton, A., Bellis, L. J., Bento, A. P., Chambers, J., Davies, M., Hersey, A., Light, Y., McGlinchey, S., Michalovich, D., Al-Lazikani, B., & Overington, J. P. (2012). ChEMBL: A large-scale bioactivity database for drug discovery. *Nucleic Acids Research*, 40(D1), D1100–D1107. <https://doi.org/10.1093/nar/gkr777>
- Genheden, S., & Ryde, U. (2015). The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opinion on Drug Discovery*, 10(5), 449–461. <https://doi.org/10.1517/17460441.2015.1032936>
- Glowacka, I., Bertram, S., Muller, M. A., Allen, P., Soilleux, E., Pfeifferle, S., Steffen, I., Tsegaye, T. S., He, Y., Gnirss, K., Niemyer, 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>
- Gupta, N., Praharij, I., Bhatnagar, T., Thangaraj, J. W. V., Giri, S., Chauhan, H., Kulkarni, S., Murhekar, M., Singh, S., Gangakhedkar, R. R., & Bhargava, B. & ICMR COVID Team. (2020). Severe acute respiratory illness surveillance for coronavirus disease 2019, India, 2020. *Indian Journal of Medical Research*, 151(1 & 2), 236–240. https://doi.org/10.4103/ijmr.IJMR_1035_20
- 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>
- Hatesuer, B., Bertram, S., Mehnert, N., Bahgat, M. M., Nelson, P. S., Pöhlmann, S., & Schughart, K. (2013). Tmprss2 is essential for Influenza H1N1 virus pathogenesis in mice. *PLoS Pathogens*, 9(12), e1003774. <https://doi.org/10.1371/journal.ppat.1003774>
- Herter, S., Piper, D. E., Aaron, W., Gabriele, T., Cutler, G., Cao, P., Bhatt, A. S., Choe, Y., Craik, C. S., Walker, N., Meininger, D., Hoey, T., & Austin, R. J. (2005). Hepatocyte growth factor is a preferred *in vitro* substrate for human hepsin, a membrane-anchored serine protease implicated in prostate and ovarian cancers. *The Biochemical Journal*, 390(1), 125–136. <https://doi.org/10.1042/BJ20041955>
- 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.e8. <https://doi.org/10.1016/j.cell.2020.02.052>
- Hou, T., Wang, J., Li, Y., & Wang, W. (2011). Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *Journal of Chemical Information & Modeling*, 51(1), 69–82. <https://doi.org/10.1021/ci100275a>
- Jangam, C. S., Bhowmick, S., Chorge, R. D., Bharatrao, L. D., Patil, P. C., Chikhale, R. V., AlFaris, N. A., ALTamimi, J., Zaidan, Wabaidur, S. M., & Islam, M. A. (2019). Pharmacoinformatics-based identification of anti-bacterial catalase-peroxidase enzyme inhibitors. *Computational Biology & Chemistry*, 83, 107136. <https://doi.org/10.1016/j.compbiolchem.2019.107136>
- Jernigan, D. B., & Cox, N. J. (2015). H7N9: Preparing for the unexpected in Influenza. *Annual Review of Medicine*, 66(1), 361–371. <https://doi.org/10.1146/annurev-med-010714-112311>
- Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., & Madden, T. L. (2008). NCBI BLAST: A better web interface. *Nucleic Acids Research*, 36(Web Server), W5–W9. <https://doi.org/10.1093/nar/gkn201>
- 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>
- Kerzare, D., Chikhale, R., Bansode, R., Amnerkar, N., Karodia, N., Paradkar, A., & Khedekar, P. (2016). Design, synthesis, pharmacological evaluation and molecular docking studies of substituted oxadiazolyl-2-oxoindolinylidene propane hydrazide derivatives. *Journal of the Brazilian Chemical Society*, 27(11), 1198–2010. <https://doi.org/10.5935/0103-5053.20160090>
- Ko, C. J., Huang, C. C., Lin, H. Y., Juan, C. P., Lan, S. W., Shyu, H. Y., Wu, S. R., Hsiao, P. W., Huang, H. P., Shun, C. T., & Lee, M. S. (2015). Androgen-induced TMPRSS2 activates matriptase and promotes extracellular matrix degradation, prostate cancer cell invasion, tumor growth, and metastasis. *Cancer Research*, 75(14), 2949–2960. <https://doi.org/10.1158/0008-5472.CAN-14-3297>
- Koch, K. R. (2018). Monte Carlo methods. *GEM: International Journal on Geomathematics*, 9(1), 117–143. <https://doi.org/10.1007/s13137-017-0101-z>
- Kumar, V., Dhanjal, J. K., Bhargava, P., Kaul, 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 & Dynamics*. <https://doi.org/10.1080/07391102.2020.1775704>
- Kumari, R., Kumar, R., & Lynn, A. Open Source Drug Discovery Consortium (2014). G-MMPBSA: A GROMACS tool for high-throughput MM-PBSA calculations. *Journal of Chemical Information & Modeling*, 54(7), 1951–1962. <https://doi.org/10.1021/ci500020m>
- Laskowski, R. A. (2009). PDBsum new things. *Nucleic Acids Research*, 37(Database), D355–D359. <https://doi.org/10.1093/nar/gkn860>
- Lee, M. G., Kim, K. H., Park, K. Y., & Kim, J. S. (1996). Evaluation of anti-influenza effects of camostat in mice infected with non-adapted human influenza viruses. *Archives of Virology*, 141(10), 1979–1989. <https://doi.org/10.1007/BF01718208>
- 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.
- Liu, J. J., Li, S. P., & Wang, Y. T. (2006). Optimization for quantitative determination of four flavonoids in Epimedium by capillary zone electrophoresis coupled with diode array detection using central composite design. *Journal of Chromatography: A*, 1103(2), 344–349. <https://doi.org/10.1016/j.chroma.2005.11.036>
- Lovell, S. C., Davis, I. W., Arendall, W. B., de Bakker, P. I. W., Word, J. M., Prisant, M. G., Richardson, J. S., Richardson, D. C. (2003). Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins*, 50(3), 437–450.
- Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., Wang, W., Song, H., Huang, B., Zhu, N., Bi, Y., Ma, X., Zhan, F., Wang, L., Hu, T., Zhou, H., Hu, Z., Zhou, W., Zhao, L., ... Tan, W. (2020). Genomic characterisation and epidemiology of 2019 novel coronavirus: Implications for virus origins and receptor binding. *Lancet*, 395(10224), 565–574. [https://doi.org/10.1016/S0140-6736\(20\)30251-8](https://doi.org/10.1016/S0140-6736(20)30251-8)
- 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>
- Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., & Simmerling, C. (2015). ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *Journal of Chemical Theory & Computation*, 11(8), 3696–3713. <https://doi.org/10.1021/acs.jctc.5b00255>
- Mariani, V., Kiefer, F., Schmidt, T., Haas, J., & Schwede, T. (2011). Assessment of template based protein structure predictions in CASP9. *Proteins: Structure, Function, & Bioinformatics*, 79(S10), 37–58. <https://doi.org/10.1002/prot.23177>
- Mark, P., & Nilsson, L. (2001). Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *Journal of Physical Chemistry A*, 105(43), 9954–9960. <https://doi.org/10.1021/jp003020w>
- Mathpati, M. M., Albert, S., & Porter, J. D. H. (2020). Ayurveda and medicalisation today: The loss of important knowledge and practice in health? *Journal of Ayurveda & Integrative Medicine*, 11(1), 89–94. <https://doi.org/10.1016/j.jaim.2018.06.004>
- Matsuyama, S., Nagata, N., Shirato, K., Kawase, M., Takeda, M., & Taguchi, F. (2010). Efficient activation of the severe acute respiratory syndrome coronavirus spike protein by the transmembrane protease TMPRSS2. *Journal of Virology*, 84(24), 12658–12664. <https://doi.org/10.1128/JVI.01542-10>
- Micronutrient Information Center (2017). *Indole-3-Carbinol* | Linus Pauling Institute | Oregon State University. Linus Pauling Institute.
- Molecular Operating Environment (MOE) (2016). 2013.08; Molecular Operating Environment (MOE), 2013.08; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7. Molecular Operating Environment (MOE), 2013.08; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2013.
- Motlhatlego, K. E., Mehrbod, P., Fotouhi, F., Abdalla, M. A., Eloff, J. N., & McGaw, L. J. (2018). Antiviral activity against influenza H1N1 virus of myricetin-3-O-rhamnoside isolated from *Newtonia buchananii* and its mechanism of action. *South African Journal of Botany*, 115, 301. <https://doi.org/10.1016/j.sajb.2018.02.091>
- NCT02914366 (2016). Amroxol as a treatment for Parkinson's disease dementia. <https://ClinicalTrials.gov/Show/NCT02914366>.
- Newman, D. J., & Cragg, G. M. (2012). Natural products as sources of new drugs over the 30 years from 1981 to 2010. *Journal of Natural Products*, 75(3), 311–335. <https://doi.org/10.1021/np200906s>
- Peramo, A. (2016). Solvated and generalised Born calculations differences using GPU CUDA and multi-CPU simulations of an antifreeze protein with AMBER. *Molecular Simulation*, 42(15), 1263–1273. <https://doi.org/10.1080/08927022.2016.1183000>
- Remmert, M., Biegert, A., Hauser, A., & Söding, J. (2012). HHblits: Lightning-fast iterative protein sequence searching by HMM–HMM alignment. *Nature Methods*, 9(2), 173–175. <https://doi.org/10.1038/nmeth.1818>
- Roe, D. R., & Cheatham, T. E. (2013). PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *Journal of Chemical Theory & Computation*, 9(7), 3084–3095. <https://doi.org/10.1021/ct400341p>
- Saha, R. K., Takahashi, T., & Suzuki, T. (2009). Glucosyl hesperidin prevents influenza A virus replication *in vitro* by inhibition of viral

- sialidase. *Biological & Pharmaceutical Bulletin*, 32(7), 1188–1192. <https://doi.org/10.1248/bpb.32.1188>
- Schrödinger (2018). LigPrep | Schrödinger. In Schrödinger Release 2018-1.
- Selleckchem. (2014). Inhibitors of signaling pathways. *Catalogue*.
- Semwal, D. K., Semwal, R. B., Combrinck, S., & Viljoen, A. (2016). Myricetin: A dietary molecule with diverse biological activities. *Nutrients*, 8(2), 90. <https://doi.org/10.3390/nu8020090>
- 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>
- Somoza, J. R., Ho, J. D., Luong, C., Ghatge, M., Sprengeler, P. A., Mortara, K., Shrader, W. D., Sperandio, D., Chan, H., McGrath, M. E., & Katz, B. A. (2003). The structure of the extracellular region of human hepsin reveals a serine protease domain and a novel scavenger receptor cysteine-rich (SRCR) domain. *Structure*, 11(9), 1123–1131. [https://doi.org/10.1016/S0969-2126\(03\)00148-5](https://doi.org/10.1016/S0969-2126(03)00148-5)
- Song, L. F., Lee, T. S., Zhu, C., York, D. M., & Merz, K. M. (2019). Using AMBER18 for relative free energy calculations. *Journal of Chemical Information & Modeling*, 59(7), 3128–3135. <https://doi.org/10.1021/acs.jcim.9b00105>
- SwissModel Help. (2020). <https://swissmodel.expasy.org/docs/help>.
- Vetter, P., Eckerle, I., & Kaiser, L. (2020). Covid-19: A puzzle with many missing pieces. *BMJ*, 368, m627. <https://doi.org/10.1136/bmj.m627>
- Wang, X., Dhindsa, R. S., Povysil, G., Zoghbi, A., Joshua, E., Hostyk, J. A., & Goldstein, D. B. (2020). Transcriptional inhibition of host viral entry proteins as a therapeutic strategy for SARS-CoV-2, Preprints. <https://doi.org/10.20944/preprints202003.0360.v1>
- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., De Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., & Schwede, T. (2018). SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research*, 46(W1), W296–W303. <https://doi.org/10.1093/nar/gky427>
- Weiser, T. (2008). Ambroxol: A CNS drug? *CNS Neuroscience & Therapeutics*. <https://doi.org/10.1111/j.1755-5949.2007.00032.x>
- WHO (2020). Coronavirus disease 2019 (COVID-19) Situation Report-54. In *WHO Bulletin*. <https://doi.org/10.1001/jama.2020.2633>
- Wiederstein, M., & Sippl, M. J. (2007). ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*, 35(Web Server), W407–W410. <https://doi.org/10.1093/nar/gkm290>
- Wong, G., He, S., Siragam, V., Bi, Y., Mbikay, M., Chretien, M., & Qiu, X. (2017). Antiviral activity of quercetin-3-β-D-glucoside against Zika virus infection. *Virologica Sinica*, 32(6), 545–547. <https://doi.org/10.1007/s12250-017-4057-9>
- Xiong, W., Ma, X., Wu, Y., Chen, Y., Zeng, L., Liu, J., Sun, W., Wang, D., & Hu, Y. (2015). Determine the structure of phosphorylated modification of icariin and its antiviral activity against duck hepatitis virus A. *BMC Veterinary Research*, 11(1), 205. <https://doi.org/10.1186/s12917-015-0459-9>
- 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 & Chemotherapy*, 60(11), 6532–6539. <https://doi.org/10.1128/AAC.01043-16>
- Yang, B., Yao, D. F., Ohuchi, M., Ide, M., Yano, M., Okumura, Y., & Kido, H. (2002). Ambroxol suppresses influenza-virus proliferation in the mouse airway by increasing antiviral factor levels. *European Respiratory Journal*, 19(5), 952–958. <https://doi.org/10.1183/09031936.02.00253302>
- Zhang, L., & Liu, Y. (2020). Potential interventions for novel coronavirus in China: A systematic review. *Journal of Medical Virology*, 92(5), 479–490. <https://doi.org/10.1002/jmv.25707>
- Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., Si, H.-R., Zhu, Y., Li, B., Huang, C.-L., Chen, H.-D., Chen, J., Luo, Y., Guo, H., Jiang, R.-D., Liu, M.-Q., Chen, Y., Shen, X.-R., Wang, X., ... Shi, Z.-L. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, 579(7798), 270–273. <https://doi.org/10.1038/s41586-020-2012-7>
- Zhou, Y., Vedantham, P., Lu, K., Agudelo, J., Carrion, R., Nunneley, J. W., Barnard, D., Pöhlmann, S., McKerrow, J. H., Renslo, A. R., & Simmons, G. (2015). Protease inhibitors targeting coronavirus and filovirus entry. *Antiviral Research*, 116, 76–84. <https://doi.org/10.1016/j.antiviral.2015.01.011>
- Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., Niu, P., Zhan, F., Ma, X., Wang, D., Xu, W., Wu, G., Gao, G. F., & Tan, W. (2020). A novel coronavirus from patients with pneumonia in China, 2019. *The New England Journal of Medicine*, 382(8), 727–733. <https://doi.org/10.1056/NEJMoa2001017>