Virtual Screening on Marine Natural Products for Discovering TMPRSS2 Inhibitors

Mehdi Mahmudpour1,2, Iraj Nabipour1, Mohsen Keshavarz2 and Maryam Farrokhnia1*

1The Persian Gulf Marine Biotechnology Research Center, The Persian Gulf Biomedical Sciences Research Institute, Bushehr University of Medical Sciences, Bushehr, Iran, 2The Persian Gulf Tropical Medicine Research Center, The Persian Gulf Biomedical Sciences Research Institute, Bushehr University of Medical Sciences, Bushehr, Iran

Although SARS-CoV-2 entry to cells strictly depends on angiotensin-converting enzyme 2 (ACE2), the virus also needs transmembrane serine protease 2 (TMPRSS2) for its spike protein priming. It has been shown that the entrance of SARS-CoV-2 through ACE2 can be blocked by cellular TMPRSS2 blockers. The main aim of this study was to find potential inhibitor(s) of TMPRSS2 through virtual screening against a homology model of TMPRSS2 using the library of marine natural products (MNPs). The homology modeling technique for generating a three-dimensional structure of TMPRSS2 was applied. Molecular docking, MM-GBSA and absorption, distribution, metabolism, excretion (ADME) evaluations were performed to investigate the inhibitory activity of marine natural products (MNPs) against TMPRSS2 and their pharmacokinetic properties. Camostat and nafamostat mesylate were used as the standard inhibitory molecules. Seven MNPs were able to inhibit TMPRSS2 better than the standard compounds. MNP 10 with CAS number 107503-09-3, called Watasenia β-D-Preluciferyl glucopyrasoiuronic acid, was found to be the best inhibitor of TMPRSS2 with acceptable pharmacokinetic properties. Herein, for the first time, a new marine natural product was introduced with potent inhibitory effects against TMPRSS2. MNP 10 exhibited favorable drug-like pharmacokinetic properties and it promises a novel TMPRSS2 blocker to combat SARS-CoV-2.

Keywords: SARS-CoV-2, COVID-19, tmprrss2, molecular modeling, molecular docking

HIGHLIGHTS
1) Marine natural products (MNPs) are a valuable source for anti-SARS-CoV-2 drugs.
2) MNP 10 is a potent TMPRSS2 inhibitor to combat SARS-CoV-2.
3) MNP 10 has favorable drug-like pharmacokinetic characteristics.

INTRODUCTION
The devastating pandemic caused by SARS-CoV-2 (Astuti, 2020; Guan et al., 2020; Wang et al., 2020; Zhu et al., 2020) that first broke out in Wuhan, China in late 2019, has become the most important global health and socioeconomic issue. Although there is a worldwide effort to develop an effective vaccine against SARS-CoV-2 using both established and new vaccine production technologies (WHO, 2021), no one can yet claim what kind of therapy can be absolutely efficient for the treatment...
of or protection against COVID-19. Based on the involved pathophysiologic pathways, different kinds of therapeutic modalities have been conducted in numerous clinical trials with conflicting results.

SARS-CoV-2 is a spherical shaped virus with a diameter of about 60–140 nm with some pleomorphism belonging to Coronaviride family. The enveloped virus has distinctive spikes (Zhu et al., 2020). Its genome is around 29.8 kilobase with a single-stranded positive-sense RNA (Lu et al., 2020a; Chan et al., 2020; Zhou et al., 2020) encoding 12 putative structural and non-structural proteins; of which spike(S), envelope (E), membrane (M) and nuclecapsid (N) proteins are structural. The S protein complex has two subunits, the S1 subunit contains a single peptide, a receptor-binding domain (RBD) which mediates attachment of virion to host cell surface receptors, and an N-terminal domain (NTD). The S2 subunit mediates fusion between the viral and host cellular membranes which facilitates virus genome entry into the host cell (Gui et al., 2017; Kirchdoerfer et al., 2018; Song et al., 2018; Wan et al., 2020).

Although it has been discovered that the SARS-CoV-2 entry to cells strictly depends on Angiotensin-Converting Enzyme 2 (ACE2) (Li et al., 2003; Li and De Clercq, 2020), it has been shown in several studies that SARS-CoV-2 also needs transmembrane serine protease 2 (TMPRSS2) for S protein priming (Hoffmann et al., 2020a; Hoffmann et al., 2020b; Stopsack et al., 2020). Hoffmann et al. showed that the entrance of SARS-CoV-2 through ACE2 can be blocked using cellular TMPRSS2 blockers. The SARS-CoV-2 spike protein contains several arginine residues with a high cleavability action at the communication point in the S1/S2 cleavage site. It has been suggested that this zoonotic-origin cleavage site sequence is required for SARS-CoV-2 entrance into human cells. These findings are consistent with previous observations from several clinically relevant viruses such as MERS, other Coronaviruses, and Influenza A virus (Kim et al., 2006; Matsuyama et al., 2010; Glowacka et al., 2011; Shulla et al., 2011; Kawase et al., 2012; Gierer et al., 2013; Zhou et al., 2015; Shen et al., 2017; Iwata-Yoshikawa et al., 2019; Kleine-Weber et al., 2019). Moreover, Heurich et al. showed that TMPRSS2 and other potentially related proteases cleave the ACE2 and SARS-S protein leading to the SARS-CoV entry and fusion of the virus S protein with the host cell membrane, respectively (Heurich et al., 2014). Therefore blocking S protein priming by specific serine protease might have the potential to control SARS-CoV-2 infection. There are some therapeutic agents like bromehexine (Lucas et al., 2014), camostat mesylate (Shirato et al., 2013), and nafamostat mesylate (Yamamoto et al., 2016) that have been elucidated to be a good inhibitor of TMPRSS2.

Regarding TMPRSS2 inhibitors as potent candidates for anti SARS-CoV-2 infection, it may be promising to investigate natural resources to discover novel components with anti-TMPRSS2 activities. The extreme and unusual environment of the ocean has extraordinary organisms with astonishing properties, which can reveal new horizons for treatment in modern medicine including marine-derived secondary metabolites with evident anti-inflammatory, antitumor, antimicrobial, antiviral, antimalarial, and antioxidant activities (Molinski et al., 2009; Gogineni et al., 2015; Blunt et al., 2018; Riccio et al., 2020; Yi et al., 2020; Carroll et al., 2021). Based on the extraordinary self-defense capacities of marine organisms and the occurrence in them of some deadly viral infections, these organisms might be regarded as a source of novel antiviral agents, which may be able to combat a SARS-CoV-2 infection. Hence, marine-derived natural compounds should be considered in our efforts to overcome the challenges of COVID-19 treatments.

A promising method to investigate viral entry and proliferation is to apply a computer-aided active site directed inhibition study. To the best of our knowledge, there are no studies that have screened the marine natural products (MNPs) libraries specifically to find a blocker of TMPRSS2. However, several studies have screened the libraries of other natural products, particularly plant-derived compounds, to discover the potential inhibitors of TMPRSS2 (Chikhale et al., 2020; Da Silva Antonio et al., 2020; Idris et al., 2020; Rahman et al., 2020; Singh et al., 2020; Vivek-Ananth et al., 2020; Hu et al., 2021).

This study aimed to find potential inhibitor(s) of TMPRSS2 through virtual screening against a homology model of TMPRSS2 using the library of marine natural products (MNPs).
MATERIALS AND METHODS

Homology Modeling

The essential step in the study of the structural and functional aspects of any protein is to have its suitable crystal structure. Unfortunately, the three-dimensional (3D) structure of TMPRSS2 had not been found at the time of the current study. Hence, in this case, the only option was to generate a 3D coordinate of TMPRSS2 by comparative prediction approach. Herein, the online server SWISS-MODEL (Guex et al., 2009;
Bertoni et al., 2017; Waterhouse et al., 2018) (https://swissmodel.expasy.org/) was used to build the 3D structure of TMPRSS2. The amino acid sequence of human transmembrane protease serine 2, from the Universal Protein Resource “UniProtKB” (accession no: O15393) isoform-2,492 amino acids long (https://www.uniprot.org/uniprot/O15393), was selected for homology modeling that was performed by a template-based method. Then, the RAMPAGE server was used to validate the 3D modeled structure (http://mordred.bioc.cam.ac.uk/~%7B%7Drapper/rampage.php).

Pharmacophore-Based Virtual Screening

The pharmacophore model was created using the Pharmit server (http://pharmit.csb.pitt.edu/). First, pharmacophore features were automatically extracted from the co-crystallized inhibitor of Serine protease Hepsin (PDB: 5CE1.}

| TABLE 1 | Homology modeling results and validation for predicting 3D structure of TMPRSS2. |
|----------|---------------------------------|
| Server/tool | Parameter | Score |
| Swiss-model | QMEAN z score | –1.47 |
| | GMQE | 0.48 |
| | Seq Similarity | 51% |
| | Coverage | 71% |
| | Seq identity | 33.82% |
| RAMPAGE(Ramachandran plot) | Favored region | 92.7% |
| | Allowed region | 6.7% |
| | Outlier region | 2% |
| MolProbity | Molprobity Score | 1.89 |
| | Ramachandran Favored | 92.20% |
| | Ramachandran Outliers | 1.16% |
| | Rotamer Outliers | 1.35% |
| | C-Beta Deviations | 7% |

| TABLE 2 | Predicted binding site of homology model of TMPRSS2. |
|----------|---------------------------------|
| Predicted binding site | C-score | E-Value | Involved residues | Docking energy of representative ligand-template complex | Predicted binding residues |
| CDD-search | — | 7.39e-100 | ILE293-GLN524 | — | HIS333 to SER478 Catlytic site: HIS333, ASP472 and SER478 Substrate binding site: ASP472, SER478, GLY499 |
| COACH-D | 0.99 | — | ILE293-GLN524 | –6.0 | HIS333, LYS379, ASP472-GLY476, SER478, THR496-CYS502 and GLY509 |

**FIGURE 2** (A) The predicted binding site by CPACH-D server. The protein template is a serine protease of the coagulation system (PDB: 5jb8) and ligand is [(4S, 5S)-2-[[[(2S)-2-amino-4-carboxybutanoyl]amino]acetyl]amino]-6-chloro-5-hydroxyhexyl]-(diaminomethylidene)azanium with pubchem CID: 137347860. (B) The binding sites predicted by Sitemap. Site1: Sitescore 0.972, Dscore 1.001, size 87, and volume 150. Site 2: Sitescore 0.968, Dscore 1.028, size 91, and volume 228. Site 3: Sitescore 0.994, Dscore 1.044, size 108, and volume 331.
A), and camostat mesylate was also used as another ligand for the pharmacophore modeling to obtain a more realistic model for homology structure of TMPRSS2 as input receptor. Hence, the pharmacophore model used for TMPRSS2 virtual screening was generated based on the predicted binding interactions of TMPRSS2 with these two inhibitors. The Pharmit parameters for 3D-pharmacophore research were changed according to these pharmacophore parameters. Then the MNP library, which contains 164,952 conformers from 14,064 molecules, was searched on this model. The hit compounds with an RMSD ≥ 4 Å and minimized affinity ≥ -6 were discarded. The remaining poses were minimized using functions of Pharmit. Finally, the pharmacophore-based minimized entries were further docked with the target protein to identify the lead compounds with the best docking scores.

**Active Site Identification and Preparation of TMPRSS2 for Docking**

The conserved domain (CD) search was done on the Fasta sequence of the homology model of TMPRSS2 using the NCBI’s conserved domain database (CDD/SPARCLE: https://www.ncbi.nlm.nih.gov/cdd/) (Lu et al., 2020b). Then, it was analyzed and its cleavage, active, and substrate binding site residues were predicted. The COACH-D server (https://yanglab.nankai.edu.cn/COACH-D/) (Yang et al., 2013; Wu et al., 2018) was also applied to predict TMPRSS2 putative ligand-binding sites. Then, in a model-template active site comparative study, the active and substrate binding residues of the (5CE1.A) template were similarly obtained.

Finally, the homology structure of TMPRSS2 was imported into Maestro Protein Wizard and sitemap analysis was done by SiteMap (Halgren, 2007; Halgren, 2009; Schrödinger, 2015b). This software was used with default settings, in which the top five possible binding sites by a minimum of 15 points were identified while cropping site maps set at 4 Å from the nearest site point. SiteMap used a more restrictive definition of hydrophobicity by standard grid. A SiteScore value above 0.80 is indicative of high druggability and promising drug-binding sites and is used in conjunction with Dscores, which serve as a measure of hydrophobicity. Dscore or druggability score penalizes increasing hydrophilicity and is thus used as a druggability measure for a pocket. In general, Dscore < 0.83 is considered as “undruggable,” 0.83–0.98 as “difficult to drug” and > 0.98 as “druggable” (Halgren, 2007; Halgren, 2009; Vidler et al., 2012).

**Molecular Docking**

**Ligand Preparation**

One conformation was generated per compound which was followed by geometry optimization with PM3 (Stewart, 1991), a semi-empirical method using Hyperchem release 7 for windows (HyperCube Inc., 2002). The geometry-optimized structures were retrieved in MOL type for further analysis with the LigPrep application which has been implemented in the Schrödinger 2015-2 suite of software (Schrödinger, 2015a). The ionization state was specified at pH = 7.00±2.0 using Epik (Epik, 2015) based on Hammett and Taft methodologies (Sastry et al., 2013). The desalt option was the same as the program default. All 32 possible conformations were produced for each compound at pH 7.00 in the OPLS3 (Harder et al., 2016). The obtained ligands were then used in the docking calculations.
TABLE 3 | CAS number, 2D and 3D presentations of investigated marine natural products.

| # | CAS#     | 2D structure | 3D structure |
|---|----------|--------------|--------------|
| 1 | 26605-16-3 | ![2D structure](image1) | ![3D structure](image2) |
| 2 | 125127-57-3 | ![2D structure](image3) | ![3D structure](image4) |
| 3 | 174286-21-6 | ![2D structure](image5) | ![3D structure](image6) |
| 4 | 454470-88-3 | ![2D structure](image7) | ![3D structure](image8) |
| 5 | 143572-73-0 | ![2D structure](image9) | ![3D structure](image10) |
| 6 | 170894-36-7 | ![2D structure](image11) | ![3D structure](image12) |
| 7 | 81720-10-7 | ![2D structure](image13) | ![3D structure](image14) |

(Continued on following page)
Generation of the Grid
As adequate generation of the grid is a key step in the prediction of a ligand binding to a receptor, the 3D boundary for ligand binding was produced by Glide, version 10.2 of Mastero, Schrödinger (Glide, 2015). First, the protein preparation wizard was used with the following settings: 1) The original hydrogens removal and subsequent addition of hydrogens. 2) The atomic charges and bond orders were assigned. 3) The N and C termini were capped. 4) The disulfide bonds were generated between sulfur atoms (within 3.2 Å). 5) Epik was applied to generate possible protonation states at neutral pH. 6) The H-bonds were assigned, optimized by PROPKA (Olsson et al., 2011; Sondergaard et al., 2011) at pH 7.0, and then the structure was minimized with the OPLS3 force field. Glide was used to generate the grid on the catalytic domain of the receptor. The grid box size was set to 32*32*32 Å.

Molecular Interaction and Docking Studies
After grid generation, ligand docking was done according to the protocols in Glide version 10.2. The homology model of TMPRSS2

| # | CAS#     | 2D structure | 3D structure |
|---|----------|--------------|--------------|
| 8 | 6184-17-4| ![2D structure](image8.png) | ![3D structure](image8.png) |
| 9 | 174286-17-0 | ![2D structure](image9.png) | ![3D structure](image9.png) |
| 10 | 107503-09-3 | ![2D structure](image10.png) | ![3D structure](image10.png) |
| 11 | 59985-26-1 | ![2D structure](image11.png) | ![3D structure](image11.png) |
was used as the receptor, and the different internally produced conformations by the software were passed across some filters such as Euler angles, grid-based force field evaluation, and energy minimization by Monte Carlo. Finally, docking score is an important parameter for evaluating the conformations, and in this study, the output of standard precision (SP) docking was put forward in extra precision (XP) docking. The docked compounds were ranked based on their docking scores.

**Pose Rescoring With Molecular Mechanics Generalized Born Surface Area**

The binding energies of all docking poses were calculated using the molecular mechanics generalized Born surface area (MM-GBSA) approach implemented in the Prime program in the Schrödinger software suite (Prime, 2015). This approach employs a single minimized protein-ligand structure, and so is used as an efficient approach to rapidly refine and rescoring docking results. A variable dielectric solvent model VSGB 2.0 (Li et al., 2011) was used, this solvent model contains several empirical corrections for modeling the directionality of hydrogen bond and π-stacking interactions. MM-GBSA has been shown to give good binding free energies for a wide range of protein-ligand complexes (Mulakala and Viswanadhan, 2013). It is also widely used to evaluate docking poses, to determine the stability of ligand-target complex for predicting binding affinity in drug design (Genheden and Ryde, 2011; Wang et al., 2019).

**Absorption, Distribution, Metabolism, and Excretion and Drug-likeness Analysis**

It is known that nearly 40% of drug candidates fail in clinical trials because of poor absorption, distribution, metabolism, and excretion (ADME). Hence, it is very crucial to recognize these problematic candidates at an early stage to avoid wasted time and resources. Accurate ADME prediction is based on full 3D molecular structures. Qikprop offers a set of several predictors including central nervous system (CNS) penetration, predicted apparent Caco-2 cell permeability across the gut-blood barrier in nm/sec (QPPCaco), apparent MDCK cell permeability (QPPMDCK), human oral absorption, Lipinski’s rule of five, and predicted maximum transdermal transport rate (JM). Another option of Qikprop (Qikprop, 2015) is to rank compounds based on how drug-like they are. In the current study, the ADME, drug-likeness, and medicinal chemistry parameters of these 11 compounds were predicted by QikProp (Qikprop, 2015).

**RESULTS AND DISCUSSION**

**Homology Modeling and Evaluation of Model Quality**

Since the crystal structure of TMPRSS2 was unavailable at the time of this study, the 3D structure of TMPRSS2 was predicted using the online server SWISS-MODEL, as shown in Figure 1A. Its global model quality estimate (GMQE) score was 0.48, this score estimates the quality of the expected output model by a particular template. Its QMEAN Z-score was -1.47, with sequence coverage of 71%, sequence identity of 33.82%, and sequence similarity of 50% in comparison with the template (PDB ID: 5CE1.A). It has been shown that when the sequence similarity with the template is more than 30%, the obtained model can be considered reliable and suitable for further study (Xiang, 2006). Benkert et al. (2011) showed that QMEAN Z-score is an estimation of the degree of native likeness of the model, and a value close to 0 (and not lower than -4) could be an acceptable agreement criteria for the experimental structure of similar size Table 1.

By validating the obtained results of SWISS-MODEL and cross-checking in RAMPAGE, it was observed that there were 319 (92.7%) residues were in the favored region, 23 (6.7%) residues in the allowed region, and 2 (0.6%) residues in the outlier region. Moreover, there were no steric clashes or deviations in bond length or bond angle compared to the protein structure report (Supplementary Figure S1 and Supplementary Figure S2). These results indicated that the obtained model may have the correct geometry. The 3D arrangement of the model is shown in Figure 1B. The alignment of the template (PDB ID: 5CE1.A) and the target protein is shown in Figure 1C. The summary of obtained results is presented in Table 1.

**Catalytic Site of the Homology Model of TMPRSS2**

According to a conserved domain database (CDD) search, the protein classification of TMPRSS2 was Trypsin-like serine proteases with E-value 7.39e-100, and the residues of ILE293-GLN524 were involved in the characteristic domain of this protein. Based on the CDD algorithm, six amino acid residues have been identified as especially important in the active site of TMPRSS2, and one residue for its cleavage site. The active site included residue HIS333 to SER478 where (HIS333, ASP382, and SER478) were the three important amino acids at the catalytic site whereas (ASP472, SER497, and GLY499) residues were found to be the substrate binding site (Table 2). Moreover, the COACH-D results supported those findings. According to the COACH-D best prediction results, the protein template was a serine protease of the coagulation system (PDB ID: 5J8B) with a confidence score of 0.99.

The corresponding docking energy of the template for its representative ligand (with pubchem CID: 137347860) was -6.0 kcal/mol. HIS333, LYS379, ASP472-GLY476, SER478, THR496-CYS502, and GLY509 were identified as the predicted binding residues, of which the CDD search showed HIS333, ASP472, SER478, SER497, and GLY499, as the same catalytic domain residues.

From the five binding sites which were calculated by sitemap, two sites had a Dscore above 1, and their locations were completely in agreement with the CDD search and COACH-D results (Site 1: Sitescore 0.972, Dscore 1.001, size 87, and volume 150. Site 2: Sitescore 0.968, Dscore 1.026, size 91, and volume 228). The Dscore of site 3 was 1.044 with Sitescore 0.994 but its position is not in agreement with the catalytic site of Trypsin-like serine proteases active domain which is located in the base of its S1 pocket where it contains Asp472, and therefore is predicted to cleave after lysine or arginine residues (Hedstrom, 2002; Wilson et al., 2005; Blay and Pei, 2019). The predicted binding sites by all
Pharmacophore Model
Since a pharmacophore states the crucial features of interactions, such as the spatial arrangement of each interaction in the close contact of ligand and the target, its accurate setup is very important in binding site pharmacophore modeling. In this study, the 3D structure of TMPRSS2 homology was used to set pharmacophore using the Pharmit server; this server provides both pharmacophore and molecular shape search options and the results are ranked by the energy. The generated pharmacophore features were selected according to the co-crystal inhibitor of the template and camostat mesylate. In this modeling, the binding-site derived pharmacophore models include three subgroups of ligand binding sites: i) two amide nitrogen atoms were added to represent hydrogen bond donors (DON), ii) four negatively charged oxygen atoms (as in a carboxyl group) were added to represent a hydrogen bond acceptor (ACC), and iii) the two isopropyl group were added to represent a hydrophobic center (HYD) (Figure 3A). According to the generated pharmacophore model, a vast library of MNP (14,064 molecules, 164,952 conformers) was filtered. A total of 25,000 hits that met the criteria were minimized, resulting in 114 conformers. A total of 11 structures were retained by using one conformer for each molecule, with an RMSD lower than 4 Å and a binding score lower than -6 (Figure 3B and Table 3).

Molecular Docking and Molecular Mechanics Generalized Born Surface Area Studies
The top 11 selected MNPs of the 114 structures which are shown in Table 3 were separately docked into the catalytic site of TMPRSS2. The results of docking of these structures, as well as camostat and nafamostat mesylate, as two standard inhibitors of TMPRSS2 (Scheme 1), are presented in Table 4.

The molecular docking analysis revealed that all the studied compounds had comparable or lower docking scores than those of the standard inhibitors. The highest docking Glide score is -8.16 in compound 10, whereas these scores were -4.52 and -3.73 in camostat and nafamostat mesylate respectively. Also, the docking scores in compounds 3, 4, 6, 7, and 11 are lower than -7. Thus, these MNPs could be considered as the most potent inhibitors for TMPRSS2. Moreover, the highest Glide energy value was -59.21 for compound 10, and its Glide emodel value is -71.34, which are almost the highest values. The Glide energies were -42.47 and -38.27 for camostat mesylate and nafamostat mesylate respectively, and Glide emodel values were -53.19 and -45.75 respectively. Therefore, from the observed theoretical superiority of these 11 MNPs compared to the standard inhibitors (camostat and nafamostat mesylate), these compounds may be encouraging for further studies. Interestingly, among these selected compounds, compound 10 had the most promising results. Compound 10 (CAS number 107503-09-3; Watasenia Preluciferyl β-D-glucopyrasoiuronic acid) is a bioluminescent substance that was derived from the

| Ligand | Docking score | Glide energy | Glide emodel | ΔGbind |
|--------|---------------|--------------|--------------|--------|
| 1      | -3.99         | -40.89       | -49.54       | -96.06 |
| 2      | -5.82         | -47.34       | -62.26       | -64.33 |
| 3      | -7.04         | -51.70       | -72.78       | -73.48 |
| 4      | -7.10         | -47.50       | -67.96       | -67.30 |
| 5      | -3.68         | -36.62       | -48.24       | -54.84 |
| 6      | -7.09         | -42.77       | -54.31       | -62.99 |
| 7      | -7.02         | -48.18       | -57.76       | -53.89 |
| 8      | -6.51         | -47.50       | -55.17       | -76.00 |
| 9      | -5.42         | -45.15       | -59.48       | -67.82 |
| 10     | -8.16         | -59.21       | -71.34       | -76.00 |
| 11     | -7.12         | -49.37       | -63.35       | -67.68 |
| Camostat | -4.52       | -42.47       | -53.19       | -57.91 |
| Nafamostat | -3.73       | -38.27       | -45.75       | -47.49 |
liver of myctophina fish, *Diaphus elucens* (Inoue et al., 1987; Blunt and Munro, 2007).

As it is presented in Table 5, the 2D template of best poses demonstrates the types of contacts formed between the ligands and target with cutoff 4.00 Å. Remarkably, close contact/interactions within the catalytic domain were detected for all MNPs, however, all the important residues of catalytic domains significantly contributed in interactions with compound 10, such as HIS333 which was involved in $\pi$-$\pi$ stacking interaction, and SER497 and GLY499 which were the important residues of substrate binding, in the close vicinity of phenol and iduronic acid moieties (Figure 4A). In addition, SER473, GLU426, LYS379, and GLU336 participated in intermolecular H-bond interaction with the active domain (Figure 4B). Taken together, compounds 3, 4, 10, and 11 had stronger interactions with TMPRSS2 than the standard ligands based on both their

| MNP | 2D presentation of best docking pose | MNP | 2D presentation of best docking pose |
|-----|------------------------------------|-----|------------------------------------|
| 1   | ![Image](image1.png)               | 2   | ![Image](image2.png)               |
| 3   | ![Image](image3.png)               | 4   | ![Image](image4.png)               |
| 5   | ![Image](image5.png)               | 6   | ![Image](image6.png)               |
| 7   | ![Image](image7.png)               | 8   | ![Image](image8.png)               |
| 9   | ![Image](image9.png)               | 11  | ![Image](image11.png)              |

Camostate

Nafamostate

**Table 5** | 2D presentation of best Glide docking pose at the catalytic domain of TMPRSS2 for investigated marine natural product.
Absorption, Distribution, Metabolism, and Excretion and Drug-likeness Analysis

In this study, the ADME https://www.sciencedirect.com/topics/pharmacology-toxicology-and-pharmaceutical-science/admeproperties of 11 MNPs were analyzed using the QikProp tool. This analysis represents the physicochemical properties of chemical compounds along with their biological functions. The resulting physicochemical and biological properties are molecular formula, molecular weight, volume, SASA, acceptor H-bond, donor H-bond groups, the number of ring atoms, Lipinski's rule of three, and the number of primary metabolites up to 7 (Di and Kerns, 2015). In addition, the predicted qualitative human oral absorption (2 = medium and 3 = high) and the predicted skin permeability (logKp values between -8.0–1.0) are considered. Finally, the ADME-compliance score drug-likeness parameter (#star) was applied for predicting the drug-likeness with the following criteria: molecular mass less than 500 Da, up to 5 hydrogen bond donors, no more than 10 hydrogen bond acceptors, and an octanol-water partition coefficient (logPo/w) no higher than 5. The rule states that a molecule or an inhibitor can be orally absorbed/active if two or more of these thresholds are not violated. However, Jorgensen's rule of three may also be used to evaluate the bioavailability of each marine natural product by estimating its solubility, permeability, and liver first-pass metabolism through the following rules: predicted aqueous solubility (logSwat) higher than -5.7 (with S in mol/dm3), predicted apparent Caco-2 cell rate permeability (BIPcaco-2) higher than 22 nm/s, and the number of primary metabolites up to 7. In general, Lipinski's rule of five (Lipinski et al., 1997) is not a strict criterion for natural compounds (Lipinski, 2003) and it has been revealed that natural compounds mostly do not follow Lipinski’s rule and they tend to keep their low hydrophobicity as well as their potential of donating the intermolecular H-bonds (Ganesan, 2008).
used to evaluate the pharmacokinetic of the studied compounds, including 25 different properties within the acceptable range of 95% of the known drugs. Herein, compounds 2, 5, 9, and 10 had the fewest violations when ADME-compliance score drug-likeness parameter (#stars) was considered less than 2 (Table 6), however, the recommended value is 0–5, and compounds 6 and 8 were also considered in the acceptable range of #stars. It is concluded that these MNPs may be proper candidate drugs for TMPRSS2 inhibition.

Again, amongst the latter compounds, compound 10 showed a higher human oral absorption (23), as well as median aggregation to plasma proteins (Qlog k has serum protein binding: −0.13) and predicted aqueous solubility values (QPlog S: 3.22). Moreover, by considering the number of “stars” and the violations from the Lipsinski and Jorgensen rules, the obtained results for compound 10 indicated a high degree of reliability to be a drug candidate.

To find chemical similarity to the known drug molecules, the QikProp (Qikprop, 2015) software database identified five similar drug molecules for each entry according to its predicted descriptors. In this study, the results of two of the compounds were very promising. Accordingly, Azithromycin with a similarity of 65.70% was suggested for compound 6. It is interesting to know that in the early phase of COVID-19, azithromycin could reduce the need for hospitalization or duration of clinical recovery (Echeverría-Esnal et al., 2020; Million et al., 2020; Molina et al., 2020). There is also an opinion supporting the potential effectiveness of Azithromycin in SARS-CoV-2 infection, as well as its antiviral activity and immunomodulatory effects (Bleyzac et al., 2020). Interestingly, Amprenavir (Shen et al., 2010), as an antiretroviral protease inhibitor for HIV infection has been identified with 68.5% similarity to compound 10.

In conclusion, marine natural product 10, with high similarity to a known antiretroviral protease, might be considered as a potent inhibitor of TMPRSS2. However, molecular docking and MM-GBSA studies demonstrated that compounds 3, 4, 6, 7, 8, 9, and 11 were also able to inhibit TMPRSS2 as well. In a comparison study, these compounds showed better results than the standard TMPRSS2 inhibitors (camostat and nafamostat mesylate). Further computational, experimental, and clinical investigations are warranted to reveal their anti-SARS-CoV-2 activities.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Materials, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

MF conceived and designed the format of the manuscript. All calculations have been done by MF. All the authors, MM, IN, and MK reviewed the manuscript and all of them contributed to the critical reading and discussion of the manuscript. All authors have read and agreed to the published version of the manuscript.

FUNDING

The authors wish to express their thanks for the partially financial support of Bushehr University of medical sciences.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fchem.2021.722633/full#supplementary-material

REFERENCES

Astuti, L., and Yraffil, Y. (2020). Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2): An Overview of Viral Structure and Host Response. Diabetes Metab. Syndr. Clin. Res. Rev. 14, 407–412. doi:10.1016/j.dsx.2020.04.020
Benkert, P., Biasini, M., and Schwede, T. (2011). Toward the Estimation of the Absolute Quality of Individual Protein Structure Models. Bioinformatics 27, 343–350. doi:10.1093/bioinformatics/btq662
Bertoni, M., Kiefer, F., Biasini, M., Bordoli, L., and Schwede, T. (2017). Modeling Protein Quaternary Structure of Homo- and Hetero-Oligomers beyond Binary Interactions by Homology. Sci. Rep. 7, 10480–10495. doi:10.1038/s41598-017-09654-8
Blay, V., and Pei, D. (2019). Serine Proteases: How Did Chemists Tease Out Their Catalytic Mechanism. ChemTexts 5, 1–7. doi:10.1007/s40828-019-0093-4
Bleyzac, N., Goutelle, S., Bourguignon, L., and Tod, M. (2020). Azithromycin for COVID-19: More Than Just an Antimicrobial. Clin. Drug Invest. 40 (8), 683–686. doi:10.1007/s40261-020-00933-3
Blunt, J. W., Carroll, A. R., Copp, B. R., Davis, R. A., Keyzers, R. A., and Prinsep, M. R. (2018). Marine Natural Products. Nat. Prod. Rep. 35, 8–53. doi:10.1039/c7np00052a
Blunt, J. W., and Munro, M. H. (2007). Dictionary of marine Natural Products with CD-ROM. London, United Kingdom: Chapman and Hall/CRC Press.

Carroll, A. R., Copp, B. R., Davis, R. A., Keyzers, R. A., and Prinsep, M. R. (2021). Marine Natural Products. Nat. prod. Rep. 38 (2), 362–413. doi:10.1039/D0NP0089B
Chan, J. F.-W., Kok, K.-H., Zhu, Z., Chu, H., To, K. K.-W., Yuan, S., et al. (2020). Genomic Characterization of the 2019 Novel Human-Pathogenic Coronavirus Isolated from a Patient with Atypical Pneumonia after Visiting Wuhan. Emerging Microbes & Infections 9, 221–236. doi:10.1080/22221751.2020.1719902
Chikhale, R. V., Gupta, V. K., Eldesoky, G. E., Wabaidur, S. M., Patil, S. A., and Islam, M. A. (2020). Identification of Potential Anti-TMPRSS2 Natural Products through Homology Modelling, Virtual Screening and Molecular Dynamics Simulation Studies. J. Biomol. Struct. Dyn. 39, 6660–6675. doi:10.1080/07391102.2020.1798813
Da Silva Antonio, A., Wiedemann, L. S. M., and Veiga-Junior, V. F. (2020). Natural Products’ Role against COVID-19. RSC Adv. 10, 23379–23393. doi:10.1039/d0ra03774e
Di, L., and Kerns, E. H. (2015). Drug-Like Properties: Concepts, Structure Design and Methods from ADME to Toxicity Optimization. Cambridge: Academic Press.
Echeverría-Esnal, D., Martín-Onituleo, C., Navarrete-Ruco, M. E., De-Antonio Cuscó, M., Fernández, O., Horcajada, J. P., et al. (2020). Azithromycin in the Treatment of COVID-19: A Review. Expert Rev. Anti Infect. Ther. 19 (2), 147–165. doi:10.1080/14787210.2020.1813024
Erik (2015). version 3. 2 ed. New York, NY: Schrödinger.
Cysteine Protease Inhibitors Prevents Severe Acute Respiratory Syndrome Coronavirus Entry. J. Virol. 86, 6337–6345. doi:10.1128/jvi.00094-12
Kim, T. S., Heinlein, C., Hackman, R. C., and Nelson, P. S. (2006). Phenotypic Analysis of Mice Lacking the TMPRSS2-Encoded Protease. Mol. Cell. Biol. 26, 965–975. doi:10.1128/mcb.26.3.965-975.2006
Kirchdoerfer, R. N., Wang, N., Pallesen, J., Wrapp, D., Turner, H. L., Cottrell, C. A., et al. (2018). Stabilized Coronavirus Spikes Are Resistant to Conformational Changes Induced by Receptor Recognition or Proteolysis. Sci. Rep. 8, 15701–15711. doi:10.1038/s41598-018-34171-7
Kleine-Weber, H., Elzayat, M. T., Wang, L., Graham, B. S., Müller, M. A., Drosten, C., et al. (2019). Mutations in the Spike Protein of Middle East Respiratory Syndrome Coronavirus Transmitted in Korea Increase Resistance to Antibody-Mediated Neutralization. J. Virol. 93, e01381–18. doi:10.1128/JVI.01381-18
Li, G., and De Clercq, E. (2020). Therapeutic Options for the 2019 Novel Coronavirus (2019-nCoV). Nat. Rev. Drug Discov. 19, 149–150. doi:10.1038/s41573-020-0016-0
Li, J., Abel, R., Zhu, K., Cao, Y., Zhao, S., and Friesner, R. A. (2011). The VSG2.0 Model: A Next Generation Energy Model for High Resolution Protein Structure Modeling. Proteins 79, 2794–2812. doi:10.1002/prot.21306
Li, W., Moore, M. J., Vasilev, N., Shi, J., Wong, S. K., Berne, M. A., et al. (2003). Angiotensin-Converting Enzyme 2 Is a Functional Receptor for the SARS Coronavirus. Nature 426, 450–454. doi:10.1038/nature02145
Lipinski, C. A. (2003). Chris Lipinski Discusses Life and Chemistry after the Rule of Five. Drug Today, 8, 12–16. doi:10.1615/DrugToday.v8.iss2.04
Lipinski, C. A., Lombardo, F., Dominy, B. W., and Feeney, P. J. (1997). Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. Adv. Drug Deliv. Rev. 23, 3–25. doi:10.1016/s0169-409x(96)00423-3
Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., et al. (2020a). Genomic Characterization and Epidemiology of 2019 Novel Coronavirus: Implications for Virus Origins and Receptor Binding. The lancet 395, 565–574. doi:10.1016/s0140-6736(20)30251-8
Lu, S., Wang, J., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., et al. (2020b). CDD/SPARCLE: The Conserved Domain Database in 2020. Nucleic Acids Res. 48, D265–D268. doi:10.1093/nar/gkaa991
Lucas, J. M., Heinlein, C., Kim, T., Hernandez, S. A., Malik, M. S., True, L. D., et al. (2014). The Androgen-Regulated Protease TMPRSS2 Activates a Proteolytic cascade Involving Components of the Tumor Microenvironment and Promotes Prostate Cancer Metastasis. Cancer Discov. 4, 1303–1325. doi:10.1158/2159-8290.cd-13-1010
Madhavi Sastry, G., Adzhigirey, M., Day, T., Annabhimouj, R., and Sherman, W. (2013). Protein and Ligand Preparation: Parameters, Protocols, and Influence on Virtual Screening Enrichments. J. Comput. Aided Mol. Des. 27, 221–234. doi:10.1007/s10822-013-9644-8
Matsuyama, S., Nagata, N., Shirato, K., Kawase, M., Takeda, M., and Taguchi, F. (2010). Efficient Activation of the Severe Acute Respiratory Syndrome Coronavirus Spike protein by the Transmembrane Protease TMPRSS2. J. Virol. 84, 12658–12664. doi:10.1128/jvi.01542-10
Million, M., Lagier, J-C., Gautret, P., Colson, P., Fournier, P-E., Amranse, S., et al. (2020). Early Treatment of COVID-19 Patients with Hydroxychloroquine and Azithromycin: A Retrospective Analysis of 1061 Cases in Marseille, France. Trav. Med. Infect. Dis. 35, 101738. doi:10.1016/j.trstmtd.2020.101738
Molina, J. M., Delaugerre, C., Le Goff, J., Mela-Lima, B., Ponscarme, D., Goldwirt, L., et al. (2020). No Evidence of Rapid Antiviral Clearance or Clinical Benefit with the Combination of Hydroxychloroquine and Azithromycin in Patients with Severe COVID-19 Infection. Médecine et Maladies Infectieuses 50, 384. doi:10.1016/j.medmal.2020.03.006
Molinski, T. F., Dalisay, D. S., Lienske, S. L., and Saludes, J. P. (2009). Drug Development from Marine Natural Products. Nat. Rev. Drug Discov. 8, 69–85. doi:10.1038/nrd2487
Mukalaka, C., and Viswanadhan, V. N. (2013). Could MMGBSA Be Accurate Enough for Calculation of Absolute Protein/Ligand Binding Free Energies. J. Mol. Graphics Modell. 46, 41–51. doi:10.1016/j.jmgm.2013.09.005
Olsson, M. H. M., Sonderegard, C. R., Rostkowski, M., and Jensen, J. H. (2011). PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. J. Chem. Theor. Comput. 7, 525–537. doi:10.1021/ct100578z
Prime (2015). version 4.0. New York, NY: Schrödinger, LLC.
Qikprop (2015). version 4.4. New York, NY: Schrödinger, LLC.
Rahman, N., Basharat, Z., Youusuf, M., Castaldo, G., Rastrelli, L., and Khan, H. (2020). Virtual Screening of Natural Products against Type II Transmembrane
Serine Protease (TMPRSS2), the Priming Agent of Coronavirus 2 (SARS-CoV-2). Molecules 25, 2271. doi:10.3390/molecules25102271
Riccio, G., Ruocco, N., Matalipassi, M., Costantini, M., Zuppi, V., Coppola, D., et al. (2020). Ten-Year Research Update Review: Antiviral Activities from Marine Organisms. *Biomolecules* 10, 1007. doi:10.3390/biom10071007
Schrödinger, L. (2015a). *LigPrep, Version 3.1*. New York, NY: Schrödinger.
Schrödinger, L. (2015b). *LigPrep, Version 3.5*. New York, NY: Schrödinger, LLC.
Shen, C.-H., Wang, Y.-F., Kovalevsky, A. Y., Harrison, R. W., and Weber, I. T. (2010). Amprenavir Complexes with HIV-1 Protease and its Drug-Resistant Mutants Altering Hydrophobic Clusters. *FEBS J.* 277, 3699–3714. doi:10.1111/j.1742-4658.2010.07771.x
Shen, L. W., Mao, H. J., Wu, Y. L., Tanaka, Y., and Zhang, W. (2017). TMPRSS2: A Potential Target for Treatment of Influenza Virus and Coronavirus Infections. *Biochimie* 142, 1–10. doi:10.1016/j.biochi.2017.07.016
Shirato, K., Kawase, M., and Matsuyama, S. (2013). Middle East Respiratory Syndrome Coronavirus Infection Mediated by the Transmembrane Serine Protease TMPRSS2. *J. Virol.* 87, 12552–12561. doi:10.1128/jvi.01890-13
Shulla, A., Heald-Sargent, T., Subramanya, G., Zhao, J., Perlman, S., and Gallagher, T. (2011). A Transmembrane Serine Protease Is Linked to the Severe Acute Respiratory Syndrome Coronavirus Receptor and Activates Virus Entry. *J. Virol.* 85, 873–882. doi:10.1128/jvi.02662-10
Singh, R., Gautam, A., Chandel, S., Ghosh, A., Dey, D., Roy, S., et al. (2020). Protease Inhibitory Effect of Natural Polyphenolic Compounds on SARS-CoV-2: An In Silico Study. *Molecules* 25, 4604. doi:10.3390/molecules25204604
Søndergaard, C. H., Olsson, M. H., Rostkowski, M., and Jensen, J. H. (2010). Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pKa Values. *J. Chem. Theor. Comput.* 7, 2284–2295. doi:10.1021/ct100133y
Song, W., Gui, M., Wang, X., and Xiang, Y. (2017). TMPRSS2: A Potential Target for Treatment of Influenza Virus and Coronavirus Infections. *Biochimie* 142, 1–10.
Shen, L. W., Mao, H. J., Wu, Y. L., Tanaka, Y., and Zhang, W. (2017).TMPRSS2: A Potential Target for Treatment of Influenza Virus and Coronavirus Infections. *Biochimie* 142, 1–10.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Mahmudpour, Nabipour, Keshavarz and Farrokhnia. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.