

Article

Inhibitors of SARS-CoV-2 Main Protease from a Library of Marine Natural Products: A Virtual Screening and Molecular Modeling Study

Davide Gentile ^{1,*}, Vincenzo Patamia ¹, Angela Scala ², Maria Teresa Sciortino ², Anna Piperno ² and Antonio Rescifina ^{1,3,*}

¹ Department of Drug Sciences, University of Catania, V.le A. Doria, 95125 Catania, Italy, davide.gentile@unict.it (DG), vincenzo.patamia@unict.it (VP) and arescifina@unict.it (AR)

² Department of Chemical, Biological, Pharmaceutical, and Environmental Sciences, University of Messina, V.le F. Stagno d'Alcontres 31, 98166 Messina, Italy, ascal@unime.it (AS), mtsciortino@unime.it (MTS), apiperno@unime.it (AP)

³ Consorzio Interuniversitario Nazionale di ricerca in Metodologie e Processi Innovativi di Sintesi (CINMPS), Via E. Orabona, 4, 70125 Bari, Italy; arescifina@unict.it (AR)

* Correspondence: davide.gentile@unict.it (DG), arescifina@unict.it (AR)

Abstract: The current emergency due to the worldwide spread of the COVID-19 caused by the new SARS-CoV-2 is a great concern for global public health. Already in the past, the outbreak of severe acute respiratory syndrome (SARS) in 2003 and Middle Eastern respiratory syndrome (MERS) in 2012 demonstrates the potential of coronaviruses to cross-species borders and further underlines the importance of identifying new-targeted drugs. An ideal antiviral agent should target essential proteins involved in the lifecycle of SARS-CoV. Currently, some HIV protease inhibitors (*i.e.*, Lopinavir) are proposed for the treatment of COVID-19, although their effectiveness was not yet assessed. The main protease (M^{pro}) provides a highly validated pharmacological target for the discovery and design of inhibitors. We identified potent M^{pro} inhibitors employing computational techniques that entail the screening of a Marine Natural Product (MNP) library. MNP library was screened by hyphenated pharmacophore model, and molecular docking approaches. Molecular dynamics and re-docking further confirmed the results obtained by structure-based techniques and allowed to highlight some crucial aspects. Seventeen potential SARS-CoV-2 M^{pro} inhibitors have been identified among the natural substances of marine origin. As these compounds were extensively validated by a consensus approach and by molecular dynamics, the likelihood that at least one of these compounds could be bioactive is excellent.

Keywords: COVID-19; SARS-CoV-2; Marine natural product; Virtual screening; Docking

1. Introduction

The new coronavirus, designated as SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2), was first identified in Wuhan, China, in December 2019 [1]. SARS-CoV-2 belongs to the family of Coronaviridae, single-stranded RNA virus (+ ssRNA) that spreads widely among humans and other mammals, causing a wide range of infections from common cold symptoms to fatal diseases, such as severe respiratory syndrome [2, 3]. The mortality rate of this new virus appears to be around 2% in China, which is much lower than the mortality rate of SARS-CoV [4]. It should be considered that the majority of fatal cases are vulnerable populations with comorbidity such as immunosuppression, diabetes, or heart disease. The primary issue all over the world is the high human-to-human transmission that has led to the spread of its outbreaks in many countries [5]. More than 81,170 cases of infection and 3,270 deaths have been reported in China, and 63,900 and 6,000, respectively, in Italy since January 22, 2020 (data from <https://www.worldometers.info/coronavirus/> at 24 March 2020). Also, cases of infection from USA, Spain, Germany, Iran, France, Sud Korea, Switzerland, UK, Netherlands, Austria, Belgium, Norway, Australia, Canada, Portugal, Sweden (in

decreasing order of infection since >2,000) and others have been reported, indicating that the disease poses a potential threat to global health [1]. Unfortunately, the number of infected continues to overgrow, and no drugs have been approved to be effective. Therefore, the need to discover and develop drugs for the treatment of the Coronavirus Disease 2019 (COVID-19) is urgent. Many research teams have exploited the major protease (M^{pro}) also named chymotrypsin-like protease ($3CL^{pro}$) [6] of SARS-CoV-2 as a potential drug target to combat COVID-19. Sequence alignment revealed that the SARS-CoV-2 M^{pro} shares a 96% similarity to that of SARS-CoV [7]. For example, Jiang and colleagues identified 30 drugs and compounds as SARS-CoV-2 M^{pro} inhibitors through protein modeling and virtual screening [8], which represents rapid progress in the way of dealing with the crisis. Furthermore, one of the 30 drugs/compounds, Remdesivir, was suggested to be a potential SARS-CoV-2 inhibitor from Liu et al., which also suggested three other possible inhibitors [9]. The SARS-CoV-2 M^{pro} and SARS-CoV M^{pro} structures are quite similar, the main difference being the surface of the proteins, where 12 different amino acids are located. Both enzymes consist of three domains; the domains I (residues 1–101) and II (residues 102–184) consist of an antiparallel β -barrel, and, for enzymatic activity, the α -helical domain III is required (residues 201–301) [10–12]. Furthermore, both SARS-CoV-2 M^{pro} and SARS-CoV M^{pro} have a structure similar to that of the cysteine proteases, although the third catalytic residue is missing at their active site. Their active site includes a catalytic dyad, His41 and Cys145, and a particularly stable water molecule that forms at least three hydrogen-bonding interactions with the surrounding residues, including a catalytic histidine, which corresponds to the position of the third catalytic element [11][12]. It should also be noted that one of the different amino acids in SARS-CoV-2 M^{pro} , namely Ser46, is located on the Cys44–Pro52 loop that flanks the cavity of the active site. Another fundamental difference between SARS-CoV-2 M^{pro} and SARS-CoV M^{pro} is given by the different sizes and shapes of the external pockets found in the two systems. Surprisingly, the volume of the external pockets of the SARS-CoV-2 M^{pro} is, on average, half that of the SARS-CoV M^{pro} . Thus, considering the similarity of the two structures, it is presumable to assume that there may be large differences between the accessibility to the binding cavity and/or the accommodation of the shape of the cavity in response to an inhibitor that can be bound. Moreover, the SARS-CoV M^{pro} apo structure has shown the largest outer pocket respect to both ligand bounded M^{pro} structures, suggesting that the SARS M^{pro} binding cavity is highly flexible and significantly changes in both volume and shape after ligand binding [11][12].

These features can be exploited in the design of lead compounds with a potentially broad spectrum of activity. Generally, a suitable substrate can be converted to a good inhibitor by replacement of a part of the substrate sequence that binds directly to the active site of the protease, reversibly or irreversibly, with the chemical warhead targeting the catalytic mechanism. Peptide inhibitors were designed by attaching a chemical warhead (*i.e.*, Michael acceptors, aldehydes, ketones, etc.) to a peptide that mimics the natural substrate [13]. These inhibitors act through a two-step procedure, wherein they first bind and form a non-covalent complex with the enzyme such that the warhead is preorganized near the catalytic residue. Subsequently, the nucleophilic attack occurs at the cysteine, leading to the formation of the covalent bond. Some peptidomimetic derivatives contain Michael acceptors as warheads and are an important class of cysteine protease inhibitors. The cysteine residue undergoes 1,4-addition to the inhibitor at the Michael acceptor warhead group, and the subsequent protonation of the α -carbanion results in the irreversible inhibition of the enzyme [13]. Our work aims to perform a virtual screening against the SARS-CoV-2 M^{pro} binding site using a library of Marine Natural Products (MNP). Many MNP have been detected as having various biological activities; peptides isolated from fish as well as algal polysaccharides have been reported to have anti-cancer, anticoagulant, and inhibitor activities. Marine bacteria and fish oils contain a considerable amount of omega-3 fatty acids, whereas seaweeds and shellfish, such as crustaceans, have potent antioxidants, including carotenoids and phenolic compounds [14]. On a pharmacophore model, built by Pharmit server (<http://pharmitcsb.pitt.edu/>) [15] starting from the SARS-CoV-2 M^{pro} (PDB ID: 6LU7) and the complexed ligand N3 (PRD_002214) structure employed as input, the virtual screening on the 164,952 conformers of the 14,064 molecules contained in the MNP library was carried out. Among these, 180 molecules were docked using AutoDock Vina software. Finally, we conducted

a parallel docking study with AutoDock4 and MD simulation studies, of 10 ns each, on the 17 molecules that showed the best promising results in terms of inhibitory activity.

2. Results and Discussion

2.1 Catalytic site of the SARS-CoV-2 M^{pro}

The crystal structure of SARS-CoV-2 M^{pro} (PDB ID 6LU7) in complex with the N3 (ID PRD_002214) inhibitor was previously solved by X-ray crystallography at a high resolution of 2.16 Å [8]. In this structure, the N3 peptide inhibitor is covalently bound with the Cys145 residue (Figure 1). Moreover, the surface topology of the active site inherent both SARS-CoV-2 M^{pro} and SARS-CoV M^{pro} proteins shows some differences, consisting in more sub-cavities and a smaller volume (337.2 Å³ vs 447.7 Å³) for the first one (Figure 1). It should be noted that the volume of the crystallized ligand influences the volume of the catalytic pocket. Moreover, the high flexibility of the catalytic site plays an essential role in the volume of the M^{pro} pocket. Consequently, inhibitors of the SARS-CoV M^{pro} might present altered binding affinities for the SARS-CoV-2 M^{pro}.

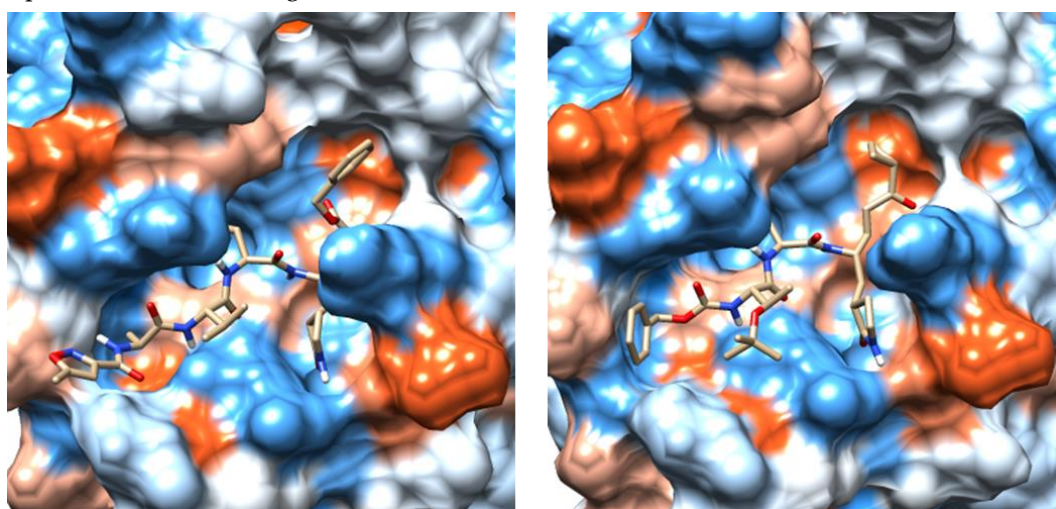


Figure 1. Comparison between the binding pockets of the crystalline structure of the N3 inhibitor linked to the SARS-CoV-2 M^{pro} (PDB ID: 6LU7) (left), and the crystal structure of TG-0205486 inhibitor bound to SARS-CoV M^{pro} (PDB ID: 2ZU5) (right).

2.2 Pharmacophore Model

A pharmacophore defines the essential features of interactions, including the spatial arrangement of each interaction in the bond between the drug and the target [15]. Structure-based pharmacophore modeling usually uses the 3D-structure of a protein while is interacting with its docked ligand. Therefore, pharmacophore is capable of incorporating more detailed information about regions available to the ligand to bind efficiently to its receptor.

In drug design, the arrangement of steric and electronic characteristics of the pharmacophore is necessary to guarantee the finest molecular interactions with the respective biological target and to inhibit its biological function [16]. The use of multiple techniques, therefore, improves the probability of identifying a diverse set of lead compounds. To design a novel SARS-CoV-2 M^{pro} inhibitor, we sought new scaffolds from MNP using a pharmacophore model of the N3/SARS-CoV-2 M^{pro} co-crystallized structure. The workflow of the virtual screening strategy for the discovery of new scaffolds binding the catalytic site of SARS-CoV-2 M^{pro} is depicted in Figure 2. The 3D pharmacophore search was performed using the Pharmit server [15], which provides both pharmacophore and molecular shape search modalities as well as a ranking of results by energy minimization.

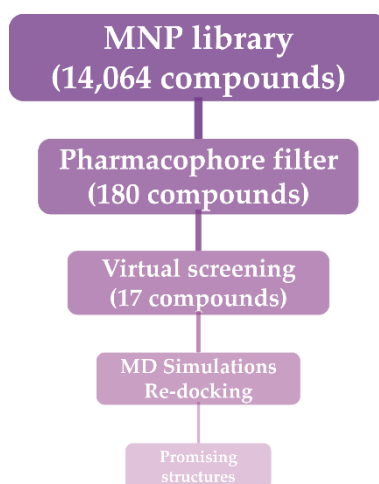


Figure 2. Workflow of the adopted molecular modeling procedure.

The proposed pharmacophore model is a binding-site-derived pharmacophore model, which includes the following pharmacophore features of ligands binding to the enzyme active site: three amide nitrogen atoms to represent hydrogen bond donors (HBD), two negatively charged oxygen atoms (as in a carboxyl group) to represent a hydrogen bond acceptor (HBA), and the isopropyl group to represent a hydrophobic center (HPC) (Figure 3).

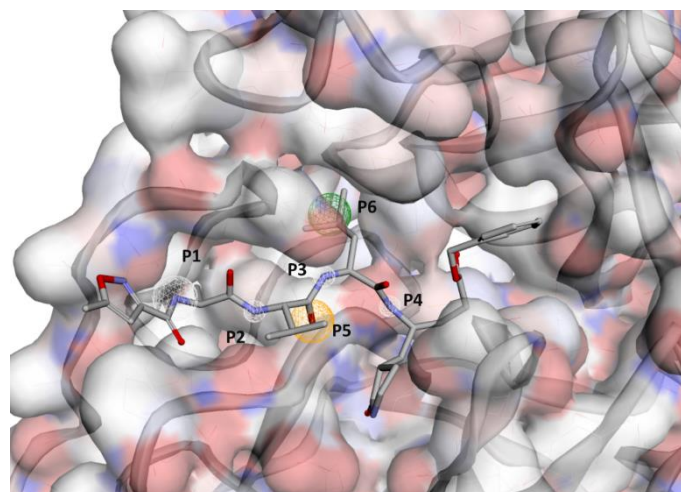


Figure 3. Pharmacophore model generated by the Pharmit server, including HBD (white spheres), negatively charged oxygen atom to represent an HBA (orange sphere) and HPC (green sphere).

The essential characteristics of the generated pharmacophore model included four HBD features, P1–P4, allocated in an opportune position to interact with residues Thr190, Glu166, Gln189, and His164, respectively; one HBA feature, P5, that is positioned to interact with Glu166 residue; and one HPC feature, P6, that is related to the interactions of the isopropyl group of the N3 inhibitor located in the hydrophobic pocket.

The generated pharmacophore model was used to filter a vast library of MNP (14,064 molecules, 164,952 conformers). From this library, 770 conformers meet the criteria of the pharmacophore filter, and after an ulterior filter that retained only one conformer for molecule with an RMSD lower than 2 Å with respect to the co-crystallized N3 ligand, 197 structures remained. Peptide structures without peptidomimetic regions or unable to act as Michael acceptors were manually discarded, as they do not meet the pharmacophore criteria of protease inhibitors. The residual 180 molecules have undergone the docking procedure.

2.3 Molecular Docking and MD Simulation

Molecular docking algorithms are often calibrated on a training set of experimental ligand-protein complexes, and the accuracy of these docking programs is often highly dependent on the used training set [19]. In this case, due to the lack of known ligands, it is essential to confirm that the docking software used for virtual screening can replicate the binding mode of a known experimental inhibitor for the enzymes studied. Although neither an effective antiviral drug nor a vaccine against COVID-19 is currently available, several reports have indicated that HIV-1 protease inhibitors, such as Lopinavir, have the potential for designing SARS-CoV-2 protease ligands [17]. In the attempt to have reference values (positive control), we decided to consider both the N3 co-crystallized ligand within the catalytic site of SARS-CoV-2 M^{Pro} (PDB ID: 6LU7) and Lopinavir as comparative standards for the molecular docking and MD simulation experiments.

Replication of the experimental binding pose by molecular docking confirmed the suitability of the docking algorithm for virtual screening. Autodock Vina software was able to accurately predict the pose of N3, with an RMSD of only 0.254 Å respect to that of the co-crystallized one.

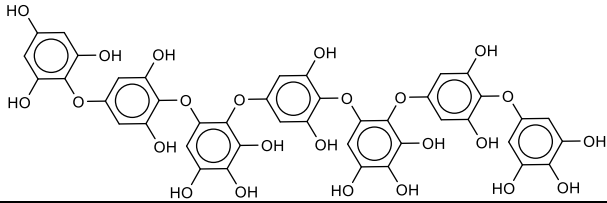
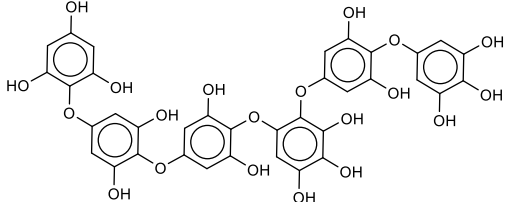
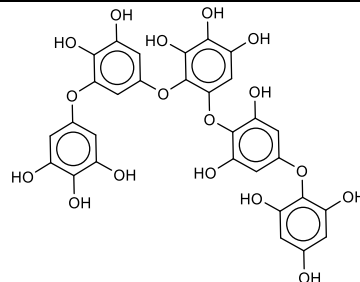
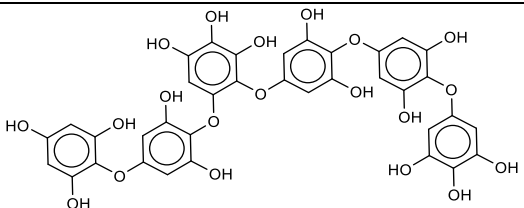
The 180 MNP structures filtered and selected according to the pharmacophore descriptors were separately docked into the SARS-CoV-2 M^{Pro} binding site. The flexible virtual screening was performed using Autodock Vina to find the most favorable binding interactions, and the calculated free binding energies were reported in Table 1 for the best 17 compounds and in Table S1 for all the others. To further validate the pharmacophore model descriptors, validate the poses and binding energies, and in-depth investigate the interactions of the new ligands within the catalytic site of the protease, we conducted a parallel docking study, with Autodock4, and MD simulations on those compounds (1–17) that showed a better affinity (Table 1). For all compounds, a re-docking was performed, using Autodock Vina, after MD simulation, taking the averaged pose of the last 3 ns.

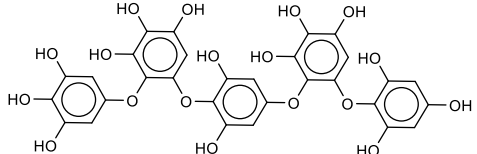
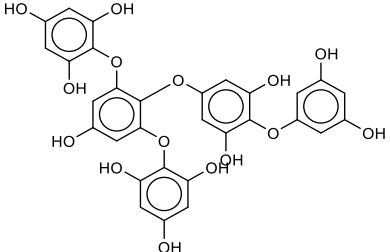
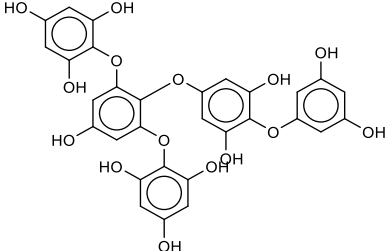
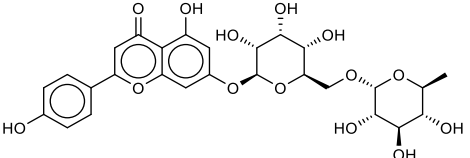
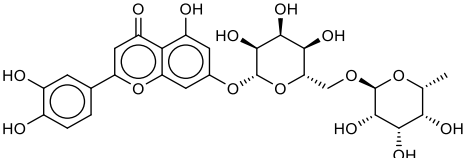
The most promising inhibitors of the SARS-CoV-2 M^{Pro} (Table 1, 1–17) are primarily represented by a class of molecules called phlorotannins, oligomers of phloroglucinol (1,3,5-trihydroxybenzene), isolated from *Sargassum spinuligerum* brown alga [18]. Although most of these phlorotannins were identified in *Sargassum spinuligerum*, other species of *Sargassum* may also contain a large number of phlorotannins, including phlorethols, fuhalols, and fucophlorethols. [19]. Algae from the *Sargassum* family are used extensively in traditional Chinese medicine [18].

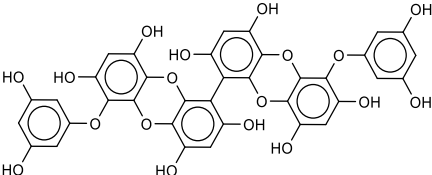
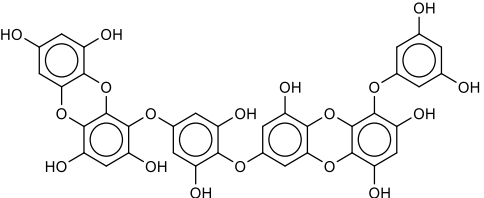
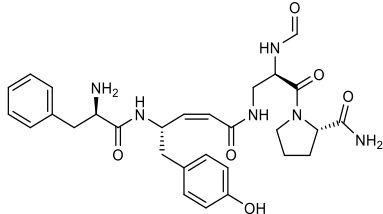
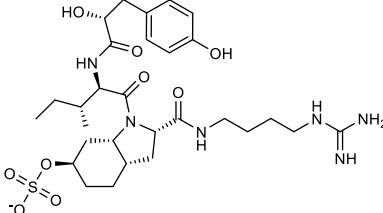
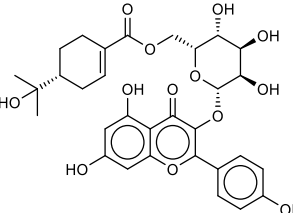
The results of the molecular docking showed that the tested compounds (1–19) had docking energies ranging from -14.6 to -10.7 kcal/mol (Table 1). Heptafuhalol A (1), showed the lowest docking energy (-14.60 kcal/mol). As shown in Figure 4, the hydroxyl groups in heptafuhalol A form an extensive network of H-bonds within the protease receptor site. The acceptor residues of hydrogen bonds are represented by Thr24, Ser46, Asn142, Glu166, and Pro168. Furthermore, π -hydrogen bonds with His41 and Gly143 residues, and hydrophobic interactions with Met49, Met65, Leu141, and Pro168, further stabilize the ligand-receptor complex. The results from the MD simulation of the heptafuhalol A/SARS-CoV-2 M^{Pro} complex are shown in Figure S1. It is interesting to note that during the MD simulation the His41 residue, belonging to the catalytic dyad, establishes an H-bond with the hydroxyl residue of the ligand, with a distance of 1.97 Å and an average D–H–A angle of 177.8° after 10 ns of simulation, with a free energy of -18.0 kcal/mol from re-docking. The RMSDs of the overall structures (SARS-Cov-2 and compound 1) with respect to the first ones at time 0 were analyzed and plotted during the 10 ns of MD simulation (Figure 4c). The overall RMSD for the protein system appeared to have reached equilibrium after 800 ps, and the stabilization of the protein-ligand complex after 3 ns, keeping the broad network of hydrogen bonds of the complex constant. The ligand is linearly arranged with three aromatic systems along the back of the protein responsible for binding to the viral proteins. Total energy, RMSDs, ligand interactions, and docking binding pose of compounds 2–17, N3, and Lopinavir were reported in Figures S2–19.

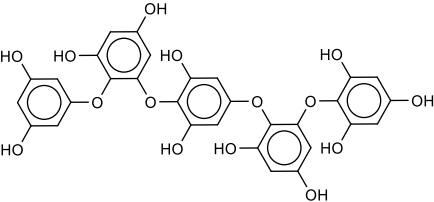
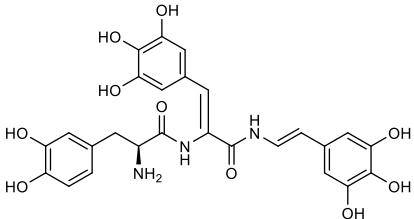
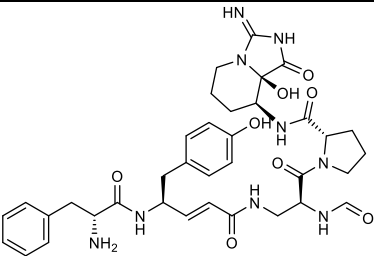
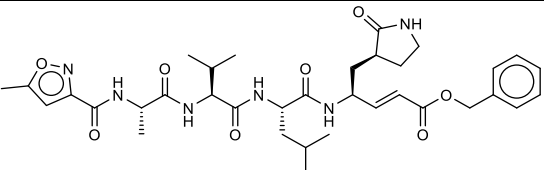
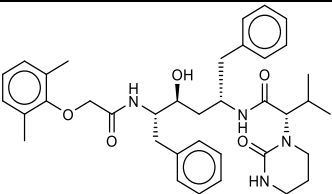
1

Table 1. Structures and calculated free binding energies (ΔG_B , in kcal/mol) of the selected MNP (compounds 1–17), N3 ligand (18), and Lopinavir (19).

Compound/Name	MNP ID	Structure	ΔG_B Vina	ΔG_B Autodock4	Average ΔG_B	ΔG_B re-docking Vina
1 Heptafuhalol A	76265-30-0		-15.4	-13.8	-14.6	-18.0
2 Phlorethopentafuhalol B	138551-15-2		-14.6	-13.9	-14.2	-13.8
3 Pseudopentafuhalol C	202211-82-3		-14.5	-13.8	-14.2	-15.8
4 Phlorethopentafuhalol A	138529-06-3		-14.0	-14.3	-14.1	-12.3

<p>5 Hydroxypentafuhalol A</p>	137809-92-8		-14.6	-11.7	-13.1	-13.7
<p>6 Pentaphlorethol B</p>			-13.9	-12.2	-13.1	-13.9
<p>7 8,8'-Bieckol</p>	89445-12-5		-13.7	-12.1	-12.9	-13.5
<p>8 Apigenin-7-O-neohesperidoside</p>	36790-49-5		-12.4	-12.3	-12.4	-12.8
<p>9 Luteolin-7-rutinoside</p>	20633-84-5		-12.1	-12.3	-12.2	-12.2

<p>10 6,6'-Bieckol</p>	88095-81-2		-12.2	-12.0	-12.1	-15.8
<p>11 Dieckol</p>	88095-77-6		-12.0	-12.1	-12.1	-12.9
<p>12 Pseudotheonamide D</p>	224577-36-0		-12.2	-10.9	-11.6	-12.2 (-14.9) ^a
<p>13 Aeruginosin 98B</p>	167228-01-5		-12.1	-10.4	-11.3	-12.5
<p>14 Resinoside B</p>	144027-79-2		-12.2	-10.2	-11.2	-11.4

15 Pentaphloretol A	164176-23-2		-12.8	-9.4	-11.1	-14.5
16 Tunichrome An2	115982-31-5		-11.5	-10.5	-11.0	-13.5
17 Pseudotheonamide C	224577-35-9		-10.5	-10.9	-10.7	-11.0 (-14.4) ^a
18 N-[(5-methylisoxazol-3-yl)carbonyl]alanyl-L-valyl-N ¹ -((1 <i>r</i> ,2 <i>z</i>)-4-(benzyloxy)-4-oxo-1-[[<i>(3r)</i> -2-oxopyrrolidin-3-yl]methyl][11]but-2-enyl)-L-leucinamide	PRD_002214		-11.9 (-15.0) ^a	-11.0 (-15.3) ^a	-11.4 (-15.1) ^a	-14.5 (-15.1) ^a
19 Lopinavir	92727		-10.3	-10.3	-10.3	-12.5

2

^a From covalent docking.

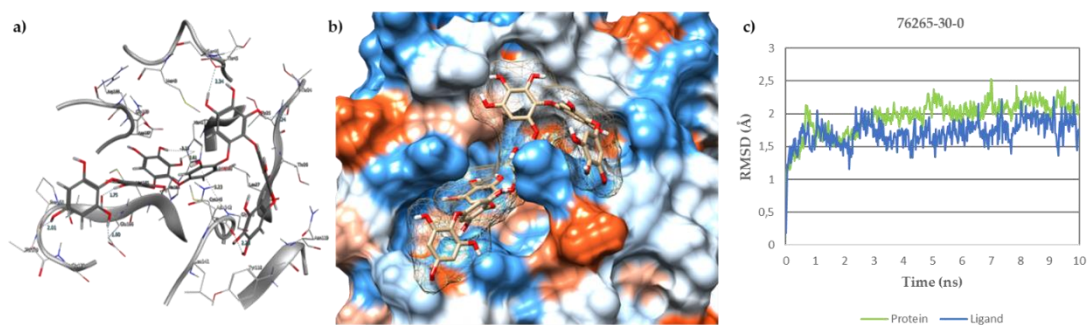


Figure 4. a) Docked pose of **1** (grey, stick model) is shown with the binding pocket residues and interacting residues with M^{Pro} (PDB ID: 6LU7). b) View of **1** inside binding pocket in hydrophobic surface representation. c) RMSD variation of the protein structure of M^{Pro} and heptafuhalol A (**1**) during the MD simulation.

Among the most active inhibitors, we find the compounds **7**, **10**, and **11**, belonging to the family of phlorotannins, isolated in the brown algae *Ecklonia cava*. The latter is an edible seaweed, which has been recognized as a rich source of bioactive derivatives, mainly phlorotannins. These phlorotannins exhibit various beneficial biological activities such as antioxidant, anticancer, antidiabetic, anti-human immunodeficiency virus, antihypertensive, matrix metalloproteinase enzyme inhibition, hyaluronidase enzyme inhibition, radioprotective, and antiallergic [20-26]. Noteworthy, dieckol (**11**) has already reported as one of the most potent SARS-CoV M^{Pro} phlorotannin inhibitors (IC₅₀ = 2.7 μM). Docking studies highlighted that interactions between dieckol and the amino acid residues in the active site of M^{Pro} are mainly constituted by an H-bonds network with calculated binding energy of -11.76 kcal/mol [27], that is comparable to the energy found by us with the SARS-CoV-2 M^{Pro}.

Although compound **7** has a better affinity (average binding energy = -12.9 kcal/mol), due to an extensive network of H-bonds, it is interesting to note that compounds **7** and **10** interact through H-bonds with the residues of the catalytic dyad His41 and Cys145 (Figure 5). In particular, 6,6'-bieckol forms an H-bond with the His41 residue with a D–A distance of 1.91 Å and a D–H–A angle of 176.0°. Dieckol shows a profile of hydrophobic interactions lower than the other two (**7** and **10**), showing interactions with the residues Leu27, Met41, Met49, Met165, Leu167, and Leu167.

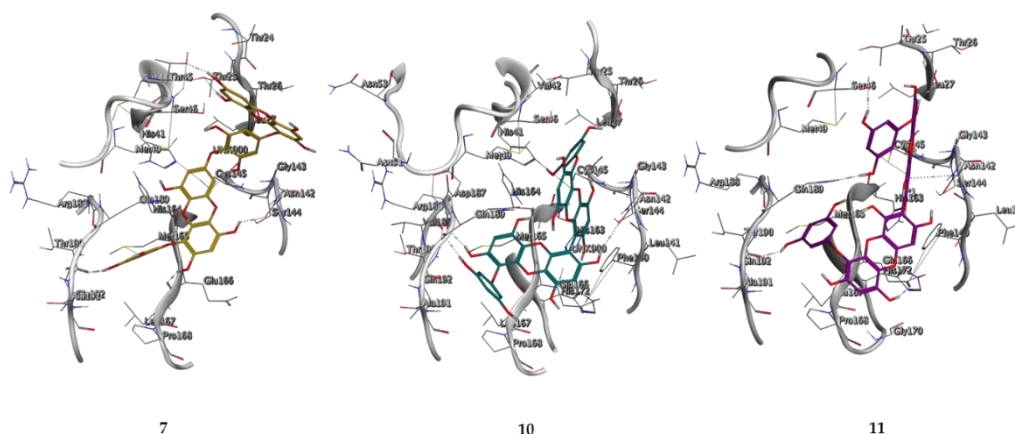


Figure 5. Interaction profile of the best-docked poses for compounds **7**, **10**, and **11**.

Among the compounds with the best-calculated activity against the SARS-CoV-2 M^{Pro}, also pseudotheonamide D (**12**) and pseudotheonamide C (**17**) emerged with a binding energy of -11.6 kcal/mol and -10.7 kcal/mol, respectively. In both cases, there is the formation of H-bonds with the residues Glu166 and Gln189, fundamental for the recognition of viral peptides within the catalytic site. Furthermore, phenyl groups occupy the small hydrophobic pocket of the enzyme interacting

with the residues Leu27, Met49, Phe140, and Leu167, so stabilizing the ligand-protein complex (Figure 6a).

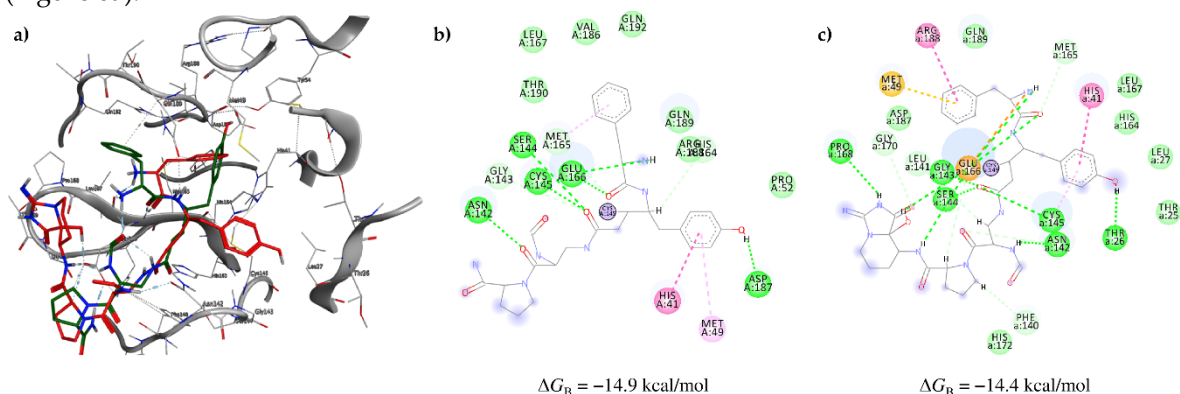


Figure 6. a) Interaction profile of the best-docked poses for compounds **12** (green, stick model) and **17** (red, stick model). Binding site interactions between SARS-CoV-2 M^{Pro} and pseudotheonamides b) **12** and c) **17**.

It is interesting to note that these compounds could represent Michael acceptor covalent inhibitors. In fact, the conjugate bond is located at a distance of about 4 Å from Cys145, suggesting that both pseudotheonamides (**12** and **17**) can form a covalent bond with Cys145. These pseudotheonamides have been isolated from the marine sponge *Theonella swinhoei* and have shown good inhibitory activity on the serine protease [28]. Consequently, we simulated the possible formation of a covalent bond between compounds **12** and **17** and the residue Cys145. A short (2 ns) MD simulation was performed in order to stabilize the new complex. The lower energy system was further minimized, and covalent docking was performed. The binding energy of **12** and **17** is very similar (-14.9 kcal/mol and -14.4 kcal/mol, respectively) with a significant increase compared to the non-covalent interaction. The two compounds adopt a similar pose within the catalytic site, establishing H-bonds with the Asn142, Ser144, and Glu166 residues, while the benzyl groups settle into the hydrophobic pockets (Figures 6b,c). Peptidomimetic derivatives containing Michael acceptor warheads are an essential class of cysteine protease inhibitors. In general, inhibitor design strategies involve the replacement of a substrate scissile amide bond with an appropriate Michael acceptor group. The cysteine residue undergoes 1,4-addition to the inhibitor at the Michael acceptor warhead group, and the subsequent protonation of the α -carbanion results in the irreversible inhibition of the enzyme [29-31].

Another class of promising M^{Pro} inhibitors has been identified in flavonoids such as Apigenin-7-*O*-neohesperidoside, Luteolin-7-rutinoside, and Resinoside. These compounds are also widespread on terrestrial plants and in food waste with good anti-tumor, anti-inflammatory, and antioxidant activity [32-36]. Among these, Apigenin-7-*O*-neohesperidoside or Rhoifolin (whose structure belongs to flavone glycoside and its aglycone is apigenin, while the neohesperidose disaccharide constitutes the glycosidic structure) has the best binding energy (-12.39 kcal/mol). The docking pose of apigenin (Figure S8) shows H-bonds between the aromatic region and residues Leu141, Glu166, and Thr190, establishing a π -stacking interaction with Gln189. In SARS-CoV M^{Pro} it has been shown that the Gln189 mutation negatively affects inhibitory activity, suggesting that this area of the protein plays a key role in the binding interaction [37].

3. Materials and Methods

3.1. Dataset of compounds

The chemical structures of the marine dataset were retrieved from Marine Natural Products (MNP, http://dockingfiles.umh.es/download/MNP_sdf_14492.zip). The full list of the 180 molecules that passed the pharmacophore filter, including the MNP ID, contacting receptor residues, and Vina

binding energy results, are available in Table 1 (compounds 1–17, and in the supplementary material (Table S1).

3.2. Pharmacophore-based virtual screening and database preparation

The 3D pharmacophore search was performed using the Pharmit server (<http://pharmit.csb.pitt.edu/>) [15]. The pharmacophore model was constructed by Pharmit by inserting the SARS-CoV-2 enzyme (PDB 6LU7) and N3 ligand (PRD_002214) structures as input. Pharmit parameters for 3D-pharmacophore research have remained unchanged, except for the hydrophobic center (isopropyl group) with a radius of 1.5 Å. This model was the basis for the virtual screening of the MNP library, which contained 14,064 molecules for a total of 164,952 conformers. The search was directed to select only one orientation for each conformation of the molecules. Compounds with an RMSD ≥ 2 Å than the N3 ligand were discarded. The remaining poses were minimized according to the functions of Pharmit.

3.3. Structures preparation and minimization

The structures of all the molecules used in this study were built using Marvin Sketch (18.24, ChemAxon Ltd., Budapest, Hungary). A first molecular mechanics energy minimization was used for 3D structures created from the SMILES; the Merck molecular force field (MMFF94) present in Marvin Sketch [38] was used. The protonation states were calculated, assuming a neutral pH. The PM3 Hamiltonian, as implemented in the MOPAC package (MOPAC2016 v. 18.151, Stewart Computational Chemistry, Colorado Springs, CO, USA) [39-41], was then used to further optimize the 3D-structures before the alignment for the docking calculations.

3.4. Molecular docking

Flexible ligand docking experiments were performed by employing AutoDock 4.2.6 and AutoDock Vina software implemented in YASARA (v. 19.5.5, YASARA Biosciences GmbH, Vienna, Austria) [42, 43], using the three-dimensional crystal structure of SARS-CoV-2 M^{pro} in complex with an inhibitor N3 PRD_002214 (PDB ID: 6LU7) obtained from the Protein Data Bank (PDB, <http://www.rcsb.org/pdb>), and the Lamarckian genetic algorithm (LGA). The covalent bond between the Cys145 residue and the crystallized ligand has been eliminated. His41 and Cys145 residues were protonated and optimized using YASARA software. The maps were generated by the program AutoGrid (4.2.6) with a spacing of 0.375 Å and dimensions that encompass all atoms extending 5 Å from the surface of the structure of the crystallized ligand. All parameters were inserted at their default settings, as previously reported [44]. In the docking tab, the macromolecule and ligand were selected, and GA parameters were set as $ga_runs = 100$, $ga_pop_size = 150$, $ga_num_evals = 25,000,000$, $ga_num_generations = 27,000$, $ga_elitism = 1$, $ga_mutation_rate = 0.02$, $ga_crossover_rate = 0.8$, $ga_crossover_mode = two\ points$, $ga_cauchy_alpha = 0.0$, $ga_cauchy_beta = 1.0$, number of generations for picking worst individual = 10.

3.5. Molecular Dynamics Simulations

The molecular dynamics simulations of the M^{pro}/ligand complexes were performed with the YASARA Structure package. A periodic cubic simulation cell with boundaries extending 8 Å [45] from the surface of the complex was employed. The box was filled with water, with a maximum sum of all water bumps of 1.0 Å, and a density of 0.997 g mL⁻¹.

The setup included an optimization of the hydrogen bonding network [46] to increase the solute stability, and a pK_a prediction to fine-tune the protonation states of protein residues at the chosen pH of 7.4 [47]. NaCl ions were added with a physiological concentration of 0.9%, with an excess of either Na or Cl to neutralize the cell. Water molecules were deleted to readjust the solvent density to 0.997 g/mL. The simulation was run using the ff14SB force field [48] for the solute, GAFF2 [49], and AM1BCC [50] for ligands and TIP3P for water. The cutoff was 10 Å for Van der Waals forces (the default used by AMBER [51], no cutoff was applied to electrostatic forces (using the Particle Mesh

Ewald algorithm, [52]). The equations of motions were integrated with multiple time steps of 2.5 fs for bonded interactions and 5.0 fs for nonbonded interactions at a temperature of 298 K and a pressure of 1 atm (NPT ensemble) using algorithms described in detail previously [53]. The final system dimensions were approximately $80 \times 80 \times 80 \text{ \AA}^3$. Short MD simulation was run on the solvent only to remove clashes. The entire system was then energy minimized using first a steepest descent minimization to remove conformational stress, followed by a simulated annealing minimization until convergence ($<0.01 \text{ kcal/mol \AA}$). Finally, 10 ns MD simulations without any restrictions were conducted, and the conformations of each system were recorded every 200 ps. After inspection of the solute RMSD as a function of simulation time, the last 3 ns averaged structures were considered for further analysis.

4. Conclusions

Natural compounds can represent a synergy to be combined with pharmacological treatments in various pathologies [54]. In some cases, the natural resource is not readily available and cannot be reproduced on a large scale (for example, marine sponges) to insert it in the global market. In other cases, the raw materials for the extraction, purification, or enrichment of bioactive compounds can be reproducible (for example, algae) for the formulation of supplements.

Here, we described the screening of a collection of MNP (14,064 compounds) in search of new, potentially SARS-CoV-2 M^{pro} inhibitors. Structure-based and ligand-based drug design approaches have developed into valuable drug discovery tools, owing to their versatility and synergistic character [44, 55, 56].

For the ligand-based evaluation, we used a pharmacophore model developed by the Pharmit server, whereas, for the structure-based evaluation, an initial docking analysis was performed, followed by a parallel docking approach on lead molecules. The poses of the 17 selected ligands have been analyzed by MD simulations. The selected compounds showed a better energy score than the drug currently used to treat COVID-19. Furthermore, it has been shown that several classes of compounds, such as phlorotannins, flavonoid, and pseudo peptides, can inhibit the SARS-CoV-2 M^{pro}, as demonstrated for the SARS-CoV M^{pro}.

Future in vitro activity assays of the ligands identified in this study will provide vital information on novel scaffolds for lead optimization.

Supplementary Materials: Supplementary materials can be found at www.mdpi.com/xxx/s1.

Author Contributions: Conceptualization, D.G. and A.R.; methodology, D.G., V.P. and A.S.; software, D.G. and A.R.; validation, D.G., V.P. and M.T.S.; formal analysis, D.G., V.P. and A.P.; investigation, D.G., A.S., and A.R.; resources, D.G. and A.R.; data curation, D.G., V.P., A.S. and A.P.; writing—original draft preparation, D.G., V.P., A.S., and A.P.; writing—review and editing, all authors; supervision, A.R.; project administration, A.R. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Hui, D. S.; E, I. A.; Madani, T. A.; Ntoumi, F.; Kock, R.; Dar, O.; Ippolito, G.; McHugh, T. D.; Memish, Z. A.; Drosten, C.; Zumla, A.; Petersen, E., The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health - The latest 2019 novel coronavirus outbreak in Wuhan, China. *Int J Infect Dis* **2020**, *91*, 264-266.
- Zumla, A.; Chan, J. F.; Azhar, E. I.; Hui, D. S.; Yuen, K. Y., Coronaviruses - drug discovery and therapeutic options. *Nat Rev Drug Discov* **2016**, *15*, 327-347.
- Song, Z.; Xu, Y.; Bao, L.; Zhang, L.; Yu, P.; Qu, Y.; Zhu, H.; Zhao, W.; Han, Y.; Qin, C., From SARS to MERS, Thrusting Coronaviruses into the Spotlight. *Viruses* **2019**, *11*.
- Coronavirus latest: WHO describes outbreak as pandemic. *Nature* **2020**.
- Munster, V. J.; Koopmans, M.; van Doremalen, N.; van Riel, D.; de Wit, E., A Novel Coronavirus Emerging in China - Key Questions for Impact Assessment. *N Engl J Med* **2020**, *382*, 692-694.
- Chen, Y.; Liu, Q.; Guo, D., Emerging coronaviruses: Genome structure, replication, and pathogenesis. *J Med Virol* **2020**, *92*, 418-423.

7. Xu, Z.; Peng, C.; Shi, Y.; Zhu, Z.; Mu, K.; Wang, X.; Zhu, W., Nelfinavir was predicted to be a potential inhibitor of 2019-nCov main protease by an integrative approach combining homology modelling, molecular docking and binding free energy calculation. *bioRxiv* **2020**, 2020.2001.2027.921627.
8. Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; Duan, Y.; Yu, J.; Wang, L.; Yang, K.; Liu, F.; Jiang, R.; Yang, X.; You, T.; Liu, X.; Yang, X.; Bai, F.; Liu, H.; Liu, X.; Guddat, L. W.; Xu, W.; Xiao, G.; Qin, C.; Shi, Z.; Jiang, H.; Rao, Z.; Yang, H., Structure of Mpro from COVID-19 virus and discovery of its inhibitors. *bioRxiv* **2020**, 2020.2002.2026.964882.
9. Morse, J. S.; Lalonde, T.; Xu, S.; Liu, W. R., Learning from the Past: Possible Urgent Prevention and Treatment Options for Severe Acute Respiratory Infections Caused by 2019-nCoV. *Chembiochem* **2020**, *21*, 730-738.
10. Bacha, U.; Barrila, J.; Velazquez-Campoy, A.; Leavitt, S. A.; Freire, E., Identification of novel inhibitors of the SARS coronavirus main protease 3CLpro. *Biochemistry-Uk* **2004**, *43*, 4906-4912.
11. Anand, K.; Ziebuhr, J.; Wadhwani, P.; Mesters, J. R.; Hilgenfeld, R., Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. *Science* **2003**, *300*, 1763-1767.
12. Bzówka, M.; Mitusińska, K.; Raczynska, A.; Samol, A.; Tuszyński, J.; Góra, A., Molecular Dynamics Simulations Indicate the COVID-19 Mpro Is Not a Viable Target for Small-Molecule Inhibitors Design. *bioRxiv* **2020**, 2020.2002.2027.968008.
13. Pillaiyar, T.; Manickam, M.; Namasivayam, V.; Hayashi, Y.; Jung, S. H., An Overview of Severe Acute Respiratory Syndrome-Coronavirus (SARS-CoV) 3CL Protease Inhibitors: Peptidomimetics and Small Molecule Chemotherapy. *J Med Chem* **2016**, *59*, 6595-6628.
14. Mayer, A. M.; Rodriguez, A. D.; Berlinck, R. G.; Fusetani, N., Marine pharmacology in 2007-8: Marine compounds with antibacterial, anticoagulant, antifungal, anti-inflammatory, antimalarial, antiprotozoal, antituberculosis, and antiviral activities; affecting the immune and nervous system, and other miscellaneous mechanisms of action. *Comp Biochem Physiol C Toxicol Pharmacol* **2011**, *153*, 191-222.
15. Sunseri, J.; Koes, D. R., Pharmit: interactive exploration of chemical space. *Nucleic Acids Res.* **2016**, *44*, W442-W448.
16. Yang, S. Y., Pharmacophore modeling and applications in drug discovery: challenges and recent advances. *Drug Discov Today* **2010**, *15*, 444-450.
17. Glombitza, K. W.; Keusgen, M., Fuhalols and Deshydroxyfuhalols from the Brown Alga *Sargassum-Spinuligerum*. *Phytochemistry* **1995**, *38*, 987-995.
18. Liu, L.; Heinrich, M.; Myers, S.; Dworjanyan, S. A., Towards a better understanding of medicinal uses of the brown seaweed *Sargassum* in Traditional Chinese Medicine: a phytochemical and pharmacological review. *J Ethnopharmacol* **2012**, *142*, 591-619.
19. Li, Y.; Fu, X.; Duan, D.; Liu, X.; Xu, J.; Gao, X., Extraction and Identification of Phlorotannins from the Brown Alga, *Sargassum fusiforme* (Harvey) Setchell. *Mar Drugs* **2017**, *15*.
20. Artan, M.; Li, Y.; Karadeniz, F.; Lee, S. H.; Kim, M. M.; Kim, S. K., Anti-HIV-1 activity of phloroglucinol derivative, 6,6'-dieckol, from *Ecklonia cava*. *Bioorg Med Chem* **2008**, *16*, 7921-7926.
21. Kong, C. S.; Kim, J. A.; Yoon, N. Y.; Kim, S. K., Induction of apoptosis by phloroglucinol derivative from *Ecklonia Cava* in MCF-7 human breast cancer cells. *Food Chem Toxicol* **2009**, *47*, 1653-1658.
22. Zhang, R.; Kang, K. A.; Piao, M. J.; Ko, D. O.; Wang, Z. H.; Lee, I. K.; Kim, B. J.; Jeong, I. Y.; Shin, T.; Park, J. W.; Lee, N. H.; Hyun, J. W., Eckol protects V79-4 lung fibroblast cells against gamma-ray radiation-induced apoptosis via the scavenging of reactive oxygen species and inhibiting of the c-Jun NH(2)-terminal kinase pathway. *Eur J Pharmacol* **2008**, *591*, 114-123.
23. Heo, S. J.; Ko, S. C.; Cha, S. H.; Kang, D. H.; Park, H. S.; Choi, Y. U.; Kim, D.; Jung, W. K.; Jeon, Y. J., Effect of phlorotannins isolated from *Ecklonia cava* on melanogenesis and their protective effect against photo-oxidative stress induced by UV-B radiation. *Toxicol In Vitro* **2009**, *23*, 1123-1130.
24. Yoon, N. Y.; Eom, T. K.; Kim, M. M.; Kim, S. K., Inhibitory effect of phlorotannins isolated from *Ecklonia cava* on mushroom tyrosinase activity and melanin formation in mouse B16F10 melanoma cells. *J Agric Food Chem* **2009**, *57*, 4124-4129.
25. Li, Y.; Lee, S. H.; Le, Q. T.; Kim, M. M.; Kim, S. K., Anti-allergic effects of phlorotannins on histamine release via binding inhibition between IgE and Fc epsilonRI. *J Agric Food Chem* **2008**, *56*, 12073-12080.
26. Myung, C. S.; Shin, H. C.; Bao, H. Y.; Yeo, S. J.; Lee, B. H.; Kang, J. S., Improvement of memory by dieckol and phlorofucofuroeckol in ethanol-treated mice: possible involvement of the inhibition of acetylcholinesterase. *Arch Pharm Res* **2005**, *28*, 691-698.

27. Park, J. Y.; Kim, J. H.; Kwon, J. M.; Kwon, H. J.; Jeong, H. J.; Kim, Y. M.; Kim, D.; Lee, W. S.; Ryu, Y. B., Dieckol, a SARS-CoV 3CL(pro) inhibitor, isolated from the edible brown algae *Ecklonia cava*. *Bioorg Med Chem* **2013**, *21*, 3730-3737.
28. Nakao, Y.; Masuda, A.; Matsunaga, S.; Fusetani, N., Pseudotheonamides, serine protease inhibitors from the marine sponge *Theonella swinhoei*. *J. Am. Chem. Soc.* **1999**, *121*, 2425-2431.
29. Ghosh, A. K.; Xi, K.; Ratia, K.; Santarsiero, B. D.; Fu, W.; Harcourt, B. H.; Rota, P. A.; Baker, S. C.; Johnson, M. E.; Mesecar, A. D., Design and synthesis of peptidomimetic severe acute respiratory syndrome chymotrypsin-like protease inhibitors. *J Med Chem* **2005**, *48*, 6767-6771.
30. Shie, J. J.; Fang, J. M.; Kuo, T. H.; Kuo, C. J.; Liang, P. H.; Huang, H. J.; Wu, Y. T.; Jan, J. T.; Cheng, Y. S.; Wong, C. H., Inhibition of the severe acute respiratory syndrome 3CL protease by peptidomimetic alpha,beta-unsaturated esters. *Bioorg Med Chem* **2005**, *13*, 5240-5252.
31. Ghosh, A. K.; Xi, K.; Grum-Tokars, V.; Xu, X.; Ratia, K.; Fu, W.; Houser, K. V.; Baker, S. C.; Johnson, M. E.; Mesecar, A. D., Structure-based design, synthesis, and biological evaluation of peptidomimetic SARS-CoV 3CLpro inhibitors. *Bioorg Med Chem Lett* **2007**, *17*, 5876-5880.
32. Moharram, F. A.; El Dib, R.; Marzouk, M. S.; El-Shenawy, S. M.; Ibrahim, H. A., New Apigenin Glycoside, Polyphenolic Constituents, Anti-inflammatory and Hepatoprotective Activities of *Gaillardia grandiflora* and *Gaillardia pulchella* Aerial Parts. *Pharmacogn Mag* **2017**, *13*, S244-S249.
33. Meng, S.; Zhu, Y.; Li, J. F.; Wang, X.; Liang, Z.; Li, S. Q.; Xu, X.; Chen, H.; Liu, B.; Zheng, X. Y.; Xie, L. P., Apigenin inhibits renal cell carcinoma cell proliferation. *Oncotarget* **2017**, *8*, 19834-19842.
34. Samet, I.; Villareal, M. O.; Motojima, H.; Han, J.; Sayadi, S.; Isoda, H., Olive leaf components apigenin 7-glucoside and luteolin 7-glucoside direct human hematopoietic stem cell differentiation towards erythroid lineage. *Differentiation* **2015**, *89*, 146-155.
35. Grosser, J. W.; Gmitter, F. G., Jr.; Tusa, N.; Recupero, G. R.; Cucinotta, P., Further evidence of a cybridization requirement for plant regeneration from citrus leaf protoplasts following somatic fusion. *Plant Cell Rep* **1996**, *15*, 672-676.
36. Napoli, E.; Siracusa, L.; Ruberto, G.; Carrubba, A.; Lazzara, S.; Speciale, A.; Cimino, F.; Saija, A.; Cristani, M., Phytochemical profiles, phototoxic and antioxidant properties of eleven *Hypericum* species - A comparative study. *Phytochemistry* **2018**, *152*, 162-173.
37. Chen, L.; Li, J.; Luo, C.; Liu, H.; Xu, W.; Chen, G.; Liew, O. W.; Zhu, W.; Puah, C. M.; Shen, X.; Jiang, H., Binding interaction of quercetin-3-beta-galactoside and its synthetic derivatives with SARS-CoV 3CL(pro): structure-activity relationship studies reveal salient pharmacophore features. *Bioorg Med Chem* **2006**, *14*, 8295-8306.
38. Barf, T.; Lehmann, F.; Hammer, K.; Haile, S.; Axen, E.; Medina, C.; Uppenberg, J.; Svensson, S.; Rondahl, L.; Lundback, T., N-Benzyl-indolo carboxylic acids: Design and synthesis of potent and selective adipocyte fatty-acid binding protein (A-FABP) inhibitors. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 1745-1748.
39. Stewart, J. J. P., Optimization of parameters for semiempirical methods IV: extension of MNDO, AM1, and PM3 to more main group elements. *J. Mol. Model.* **2004**, *10*, 155-164.
40. Alemán, C.; Luque, F. J.; Orozco, M., Suitability of the PM3-derived molecular electrostatic potentials. *J. Comput. Chem.* **1993**, *14*, 799-808.
41. Qiao, F.; Luo, L.; Peng, H.; Luo, S.; Huang, W.; Cui, J.; Li, X.; Kong, L.; Jiang, D.; Chitwood, D. J.; Peng, D., Characterization of Three Novel Fatty Acid- and Retinoid-Binding Protein Genes (*Ha-far-1*, *Ha-far-2* and *Hf-far-1*) from the Cereal Cyst Nematodes *Heterodera avenae* and *H. filipjevi*. *PLoS One* **2016**, *11*, e0160003.
42. Krieger, E.; Koraimann, G.; Vriend, G., Increasing the precision of comparative models with YASARA NOVA--a self-parameterizing force field. *Proteins* **2002**, *47*, 393-402.
43. Krieger, E.; Vriend, G., YASARA View - molecular graphics for all devices - from smartphones to workstations. *Bioinformatics* **2014**, *30*, 2981-2982.
44. Floresta, G.; Amata, E.; Barbaraci, C.; Gentile, D.; Turnaturi, R.; Marrazzo, A.; Rescifina, A., A Structure- and Ligand-Based Virtual Screening of a Database of "Small" Marine Natural Products for the Identification of "Blue" Sigma-2 Receptor Ligands. *Mar Drugs* **2018**, *16*.
45. Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P., A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **2003**, *24*, 1999-2012.
46. Krieger, E.; Dunbrack, R. L., Jr.; Hoof, R. W.; Krieger, B., Assignment of protonation states in proteins and ligands: combining pKa prediction with hydrogen bonding network optimization. *Methods Mol Biol* **2012**, *819*, 405-421.

47. Krieger, E.; Nielsen, J. E.; Spronk, C. A.; Vriend, G., Fast empirical pKa prediction by Ewald summation. *J Mol Graph Model* **2006**, *25*, 481-486.
48. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C., ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696-3713.
49. Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157-1174.
50. Jakalian, A.; Jack, D. B.; Bayly, C. I., Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* **2002**, *23*, 1623-1641.
51. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C., Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **2006**, *65*, 712-725.
52. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577-8593.
53. Krieger, E.; Vriend, G., New Ways to Boost Molecular Dynamics Simulations. *J. Comput. Chem.* **2015**, *36*, 996-1007.
54. Di Lodovico, S.; Napoli, E.; Di Campli, E.; Di Fermo, P.; Gentile, D.; Ruberto, G.; Nostro, A.; Marini, E.; Cellini, L.; Di Giulio, M., Pistacia vera L. oleoresin and levofloxacin is a synergistic combination against resistant *Helicobacter pylori* strains. *Sci Rep* **2019**, *9*, 4646.
55. Floresta, G.; Gentile, D.; Perrini, G.; Patamia, V.; Rescifina, A., Computational Tools in the Discovery of FABP4 Ligands: A Statistical and Molecular Modeling Approach. *Mar Drugs* **2019**, *17*.
56. Floresta, G.; Amata, E.; Gentile, D.; Romeo, G.; Marrazzo, A.; Pittala, V.; Salerno, L.; Rescifina, A., Fourfold Filtered Statistical/Computational Approach for the Identification of Imidazole Compounds as HO-1 Inhibitors from Natural Products. *Mar Drugs* **2019**, *17*.