King Saud University

Arabian Journal of Chemistry

www.ksu.edu.sa
www.sciencedirect.com

ORIGINAL ARTICLE

Structure- and Ligand-Based *in silico* Studies towards the Repurposing of Marine Bioactive Compounds to Target SARS-CoV-2

Marwa A.A. Fayed [a], **Mohammed Farrag El-Behairy** [b], **Inas A. Abdallah** [c], **Hend Mohamed Abdel-Bar** [d], **Hanan Elimam** [e], **Ahmed Mostafa** [f], **Yassmin Moatasim** [f], **Khaled A.M. Abouzid** [b,*], **Yaseen A.M.M. Elshaier** [b,*]

[a] *Department of Pharmacognosy, Faculty of Pharmacy, University of Sadat City, Menoufia 32897, Egypt*
[b] *Department of Organic and Medicinal Chemistry, Faculty of Pharmacy, University of Sadat City, Menoufia 32897, Egypt*
[c] *Department of Analytical Chemistry, Faculty of Pharmacy, University of Sadat City, Menoufia 32897, Egypt*
[d] *Department of Pharmaceutics, Faculty of Pharmacy, University of Sadat City, Menoufia 32897, Egypt*
[e] *Department of Biochemistry, Faculty of Pharmacy, University of Sadat City, Menoufia 32897, Egypt*
[f] *Center of Scientific Excellence for Influenza Viruses, National Research Centre, Giza 12622, Egypt*

Abstract This work was a structured virtual screening for marine bioactive compounds with reported antiviral activities which were subjected to structure-based studies against SARS-CoV-2 co-crystallized proteins. The molecular docking of marine bioactive compounds against the main protease (M^{pro}, PDB ID: 6lu7 and 6y2f), the spike glycoprotein (PDB ID: 6vsb), and the RNA polymerase (PDB ID: 6m71) of SARS-CoV-2 was performed. Ligand-based approach with the inclusion of rapid overlay chemical structures (ROCS) was also addressed in order to examine the probability of these marine compounds sharing relevance and druggability with the reported drugs. Among the examined marine library, the highest scores in different virtual screening aspects were displayed by compounds with flavonoids core, acyl indole, and pyrrole carboxamide alkaloids. Moreover, a complete overlay with the co-crystallized ligands of M^{pro} was revealed by sceptrin and debromo-sceptrin. Thalassoilin (A-B) which was found in the Red Sea exhibited the highest binding and similarity outcomes among all target proteins. These data highlight the importance of marine

* Corresponding authors.
 E-mail addresses: khaled.abouzid@fop.usc.edu.eg (K.A.M. Abouzid), yaseen.elshaier@fop.usc.edu.eg (Y.A.M.M. Elshaier).
Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

natural metabolites in regard to further studies for discovering new drugs to combat the COVID-19 pandemic.

1. Introduction

The coronavirus disease 2019 (COVID-19) is known to affect the respiratory system, causing severe acute respiratory syndrome; it was declared by the WHO in March 2020 as a global pandemic. The severity and spreading of the disease are attributed to different types of human coronaviruses (Lai et al., 2020). A wide range of symptoms can be caused by these types, similar to those of the infection by influenza such as in the case of the Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS) (Chen et al., 2020). The SARS-CoV-2 was identified to be a single-stranded positive-sense RNA virus with a size of ~30 kb that belongs to the genus Beta-coronaviruses with a crown due to the presence of spike glycoproteins on the viral envelope (Huang et al., 2020). It can infect the throat, the lower respiratory tract and cause pneumonia in humans, although it seems that the symptoms are milder than those of the other SARS viruses (Huang et al., 2020). To date, more than 59.81 million cases in total have been confirmed worldwide along with 1.41 million death, according to the World Meter (Worldometer, 2020).

Each SARS-CoV-2 protein has its relevant functional role, active site and subsequently, its own druggable target. These proteins as non-structural proteins (nsp) e.g. nsp 3, the papain-like protease (PLpro, nsp3 domain), main protease (M^{pro}, nsp5), the RNA polymerase complex (nsp12-nsp7-nsp8), and other targets (Cavasottoa et al., 2021).

The SARS-CoV-2 main protease (M^{pro}) is a cysteine protease engaged in many cleavage steps in the precursor polyproteins, and so it has a critical role in the viral life cycle.

The active form of M^{pro} is a homodimer containing two protomers composed of three domains each: domain I, domain II, and domain III, domains II and III are connected by a long loop region (Jin et al., 2020a, 2020b; Johnson et al., 2010).

The binding of SARS-CoV-2 to angiotensin-converting enzyme 2 (ACE-2) receptors in type II pneumocytes in the lungs triggers the cascade of an inflammatory response in the lower respiratory tract. SARS-CoV-2 is very similar to SARS-CoV-1 regarding the biochemical interactions and pathogenesis (Li et al., 2020). The spike glycoprotein is composed of a transmembrane trimetric glycoprotein protruding from the viral surface; it determines the diversity of coronaviruses and host tropism (Li et al., 2003). Moreover, SARS-CoV-2 spike glycoproteins also contain a variable receptor-binding domain (RBD) in the S1 functional subunit, which binds to the ACE-2 receptor and subsequently to the type 2 transmembrane protease serine 2 (TMPRSS2). The interaction complex is cleaved by TMPRSS2, leading to ACE-2 and the activation of the spike glycoprotein (Fig. 1), facilitating viral entry into the target cell (Glowacka et al., 2011; Heurich et al., 2014). Inside the host cell, viral RNA hijacks the host cell machinery, inducing it to produce RNA and proteins that contribute to the assembly of new viral particles.

Natural bioactive products are considered to be one of the main and most diverse sources of a variety of therapeutic agents, providing more than 50% of the available drugs which are in clinical use around the world (Boopathy and Kathiresan, 2010). Natural compounds with marine origin are of great importance as a source of novel and potentially life-saving bioactive secondary metabolites, since the marine environment represents 95% of the biosphere (Jimeno et al., 2004). There are more than 40,000 different phytoplankton species, including 680 species of marine algae belonging to Rhodophyta, Phaeophyta, and Chlorophyta, commonly known as red, brown, and green seaweed, respectively, in addition to 71 mangrove plant species that were previously documented in the global marine biotope. They provide essential metabolites, including fatty acids, ionic trace minerals, vitamins, enzymes, bioflavonoids, amino acids, as well as other nutrients (Boopathy and Kathiresan, 2010), which possess several beneficial biological effects, such as anti-cancer, antibacterial, and antiviral activity even against resistant strains (Boopathy and Kathiresan, 2010). Various aquatic organisms were previously reported to provide a variety of unique lead compounds with antiviral activity (Che, 1991). The drug repurposing approach provides a quick tool to overcome the coronavirus pandemic. Most drugs subjected to repurposing with the aim of overcoming the COVID-19 outbreak are commercially available and their dosage and toxicity in humans are well known with a history of decades of clinical use in some cases (Ciliberto et al., 2020). Many valuable efforts in drug repurposing approaches have been recognized against SARS-CoV-2 utilizing cheminformatics and bioinformatics tool (Meyer-Almes, 2020; Touret et al., 2020; Shyr et al., 2020; Cavasotto and Di Filippo, 2021; Ferraz et al., 2020; Battisti et al., 2020).

Recently, Claudio N. C. et al. addressed three drug discovery strategies in the frame structure and druggability studies of the SARS-CoV-2 proteome (Cavasottoa et al., 2021). These include interfering proteins responsible for genome stability as non-structural proteins (nsp 13), blocking proofreading proteins as nsp 14, and targeting nsp15 responsible for viral replication and also the accumulation of cytokine-producing macrophages (Cavasotto et al., 2020). The repositioning of antiviral drugs that are already used against SARS-CoV-2 proteins, including remdesivir (an RNA polymerase inhibitor), umifenovir (a membrane fusion inhibitor targeting viral entry), and lopinavir/ritonavir (a drug combination targeting viral proteases) was reported and currently being considered in different combinations in a Phase IV clinical trial, as shown in Fig. 1 (Senanayake, 2020). Therefore, the drug repurposing approach could be a promising tool for the discovery of effective therapy against this novel coronavirus. This oriented us to search for a COVID-19 remedy among the known marine metabolites through in-silico screening for their activity against the COVID-19 protease, polymerase, as well as spike

a)
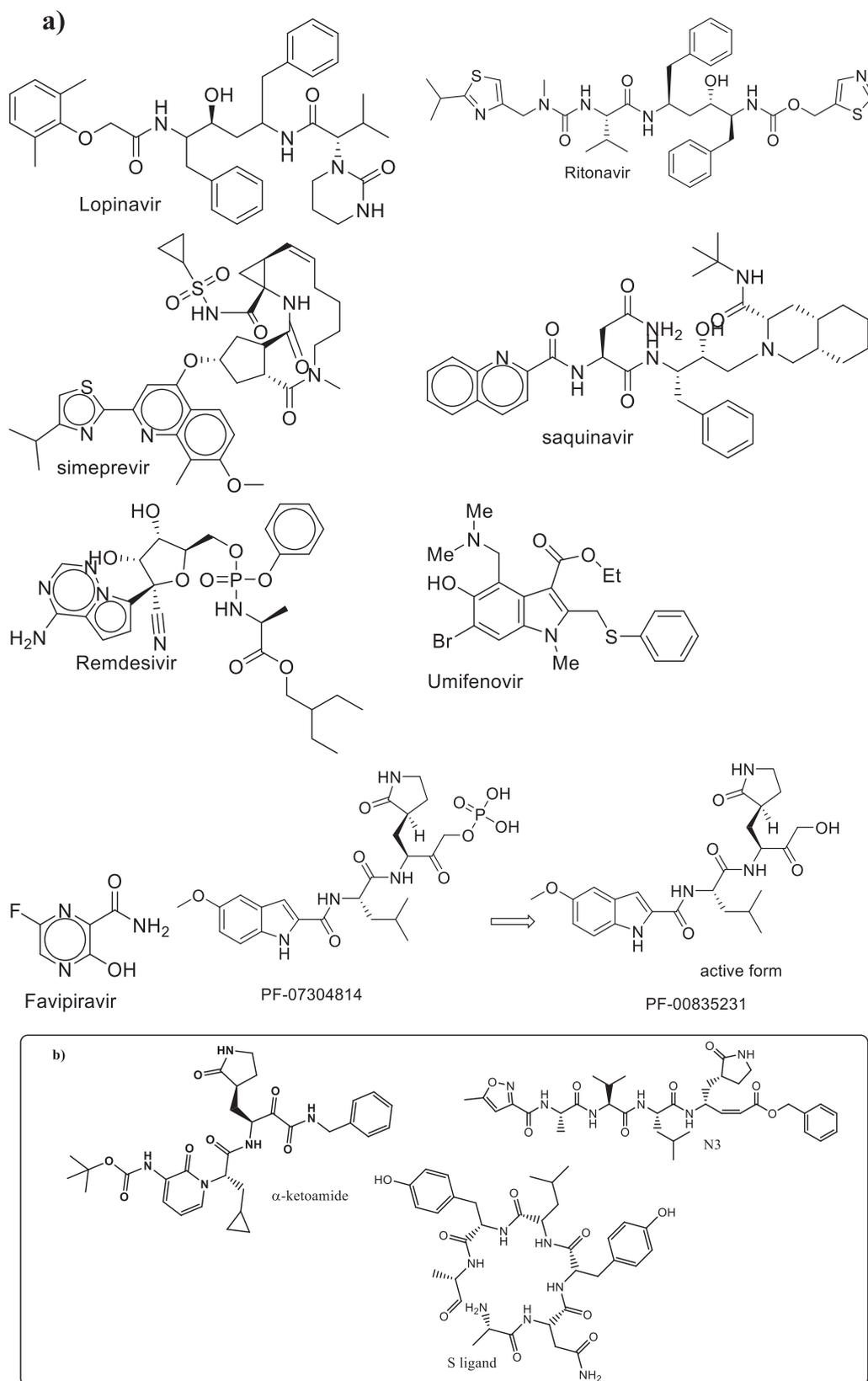
Fig. 1   a) Chemical structure of reported drugs used in the treatment of COVID-19; b) standard ligands for SARS-CoV-2 proteins.

glycoproteins and comparing their activity to already used anti-coronavirus drugs (Ibrahim et al., 2020a, 2020b, 2020c, 2021).

2. Results and discussion

Rationale of the study (Fig. 2):

This study included screening for marine bioactive compounds with reported antiviral activities (Table 1, supplementary data). Their chemical structures, activities, origins, and the geographical sources were described in Table 1. Subsequently, the structure-based approach for repurposing these compounds against SARS CoV-2 co-crystallized proteins was addressed. The departure point started with docking marine bioactive molecules with main protease M^{pro}, nsp5, which is a cysteine protease containing two protomers composed of three domains. Crystal structures of SARS-CoV-2 M^{pro} were reported in the protein data bank (PDB ID: 6lu7 and 6y2f). Then docking with The SARS-CoV-2 spike-protein (PDB ID: 6vsb) which is a large homotrimeric multidomain glycoprotein. Finally, the attention was directed to dock marine compounds with RNA polymerase (nsp12-nsp7-nsp8 (PDB ID: 6m71) (Cavasottoa et al., 2021). Full binding modes and poses were discussed for the highest-scoring compounds. Our attention was afterwards directed to the ligand-based drug repurposing tactic, which included shape alignment, rapid overlay chemical structures (ROCS), and structure-property relationship to examine the druggability of the prioritized mar-

ine compounds. These bioactive compounds were made publicly accessible for facilitating the conduction of further studies and optimization by the scientific community (Culletta et al., 2020).

2.1. Structure based approach via molecular docking study

The study was performed using the Openeye scientific software (Allam, 2020) (academic license 2020). The compounds were ranked and sorted according to their consensus score values. Docking results of marine compounds against M^{pro} (PDB ID: 6lu7 (Jin et al., 2020) and 6y2f (Zhang et al., 2020), spike glycoprotein (PDB ID: 6vsb (Wrapp et al., 2020), and RNA polymerase (PDB ID: 6m71 (Gao et al., 2020) were subsequently filtered based on their obtained consensus score values. The lowest consensus score value indicated the strongest binding to the receptor. The compounds shown in Fig. 3 displayed an interaction to the targeted receptors that was stronger than that of standard ligands (N3, α- ketoamide for M^{pro}, S ligand 1 for spike protein, and remdesivir for RNA polymerase, respectively, shown in Fig. 1). These compounds were subjected to the acquisition of further details, as shown in Tables 2–5. The top hits compounds, as shown in Fig. 3, were classified based on their chemical structures into 1) hydrocarbon-based (enyne, conjugated ene, or alkanes) compounds, including polyacetylenetriol, aplidiasphingosine, calyceramide B, and acarnidine C with different functionalities, such as alcohol, amine, amide, or guanidine; 2) indole or carboline alkaloids,
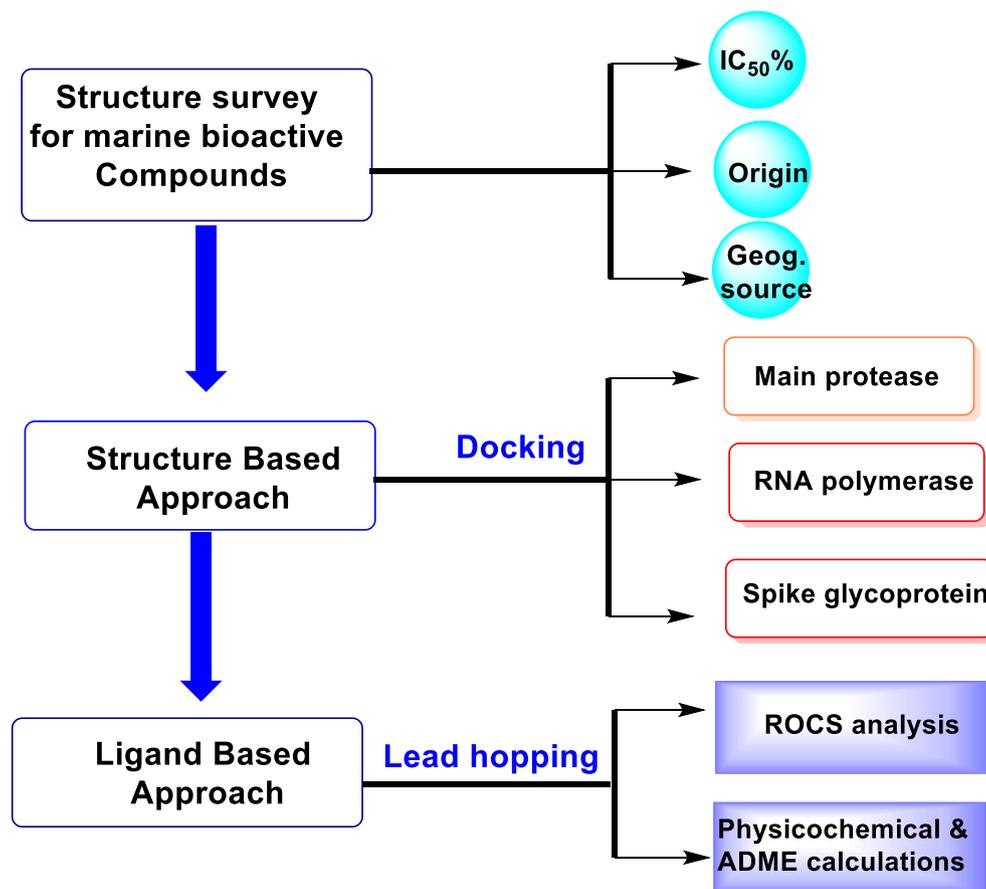
Fig. 2 Flow chart for the rationale of the study.
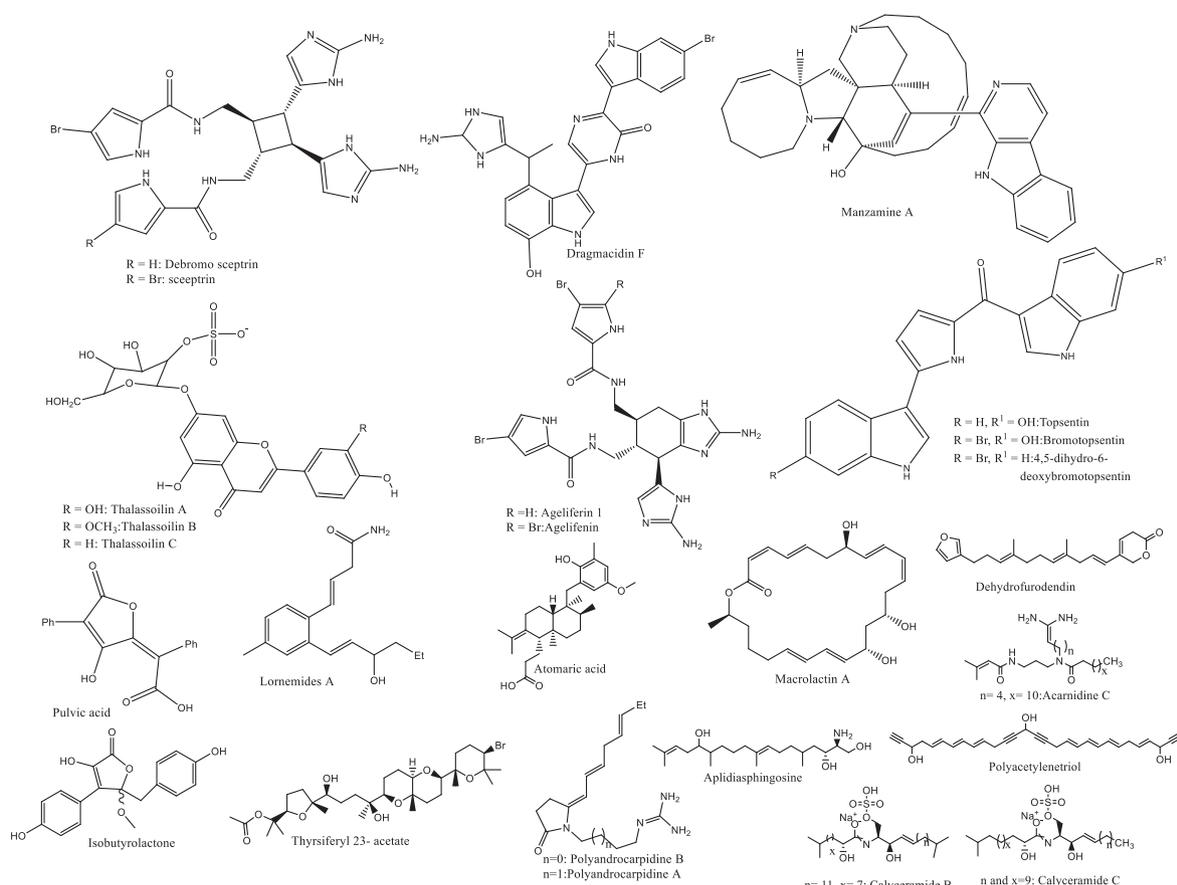
Fig. 3   Structure of marine compounds showed consensus scores more than co-crystalized ligands.

namely topsentin, bromotopsentin, 4,5-dihydro-6-deoxybromo topsentin, trypilepyrazinol, dragmacidin F, and manzamine A; 3) flavonoids, namely thalassoilin B, thalassoilin C, and thalassoilin A; 4) steroidal nucleolus compounds, such as 3β-hydro xyergosta-8,14,24(28)-trien-7-one and weinbersterol B; 5) hydroxyl lactone derivatives, such as aspernolide A and neocitreoviridin; and 6) macrolides, such as macrolactine. Among each nucleolus, weak binding was displayed by some compounds to all receptors, indicating that there is a structure in silico activity relationship that guided us to develop a suitable model for designing new inhibitors for SARS-CoV-2 proteins.

2.1.1. Molecular docking of antiviral marine compounds with M^{pro} (nsp5) of SARS-CoV-2

2.1.1.1. Docking with M^{pro} of SARS-CoV-2 (PDB ID 6lu7).
Marine compounds were found to exhibit binding affinity more than standard ligands, as represented in Fig. 3. The compounds were sorted based on their interaction with two M^{pro} SARS-CoV-2 proteins, as indicated in Tables 2 and 3. Propargyl alcohol (polyacetylenetriol) exhibited strong interaction toward receptor of M^{pro} SARSCoV-2 proteins, occupying the receptor of PDB: ID 6lu7 domains with a hydrophobic-hydrophobic interaction, as shown in Fig. 4a. Among the compounds that contained an indole moiety, deoxybromotopsentin and bromotopsentin exhibited a hydrogen bond (HB) formation with Gln89, as shown in Fig. 4b. To illustrate the weak binding of topsentin in comparison with previous indoles, a different binding mode and pose were adopted, (supplementary data).

Sceptrin was found to form HBs with key amino acids, including Glu166 and Gln189 as well as two HBs with Asn142. Moreover, sceptrin was observed to overlay with co-crystallized ligands through a strong binding interaction, as shown in Fig. 4c. Thalassoilin A and thalassoilin B exhibited the same binding mode and pose through the formation of HB with Gly143, although both displayed different modes of interaction in case of thalassoilin C, as shown in Fig. 4d. Snapshot of these three flavonoids indicated high similarity in case of ligand pose. Interestingly, moderate in vitro antiviral HCV protease activity with an IC50 value of 16 μM was displayed by these flavonoids (Hawas and Abou El-Kassem, 2017). This result suggests that the total extract of this metabolite has the potential to be further investigated. Agelifenin exhibited weak bonding and no overlay with standard ligands in comparison to agelifenin 1, which was assumed due to extra bromination (supplementary data).

2.1.1.2. Docking with M^{pro} of SARS-CoV-2 (PDB ID:6y2f).
Some compounds still exhibited a stronger binding mode with PDB ID 6y2f than in case of 6lu7. Debromo sceptrin showed unique and strong interaction with the receptor through the formation of a binding mode and poses like the co-crystallized ligand, as shown in Fig. 5a. Debromo sceptrin (pink) and sceptrin (blue) showed no co-overlay, as shown in Fig. 5b. Moreover, a complete overlay was identified for thalassoilin B and thalassoilin A, while thalassoilin C exhibited different binding modes and pose, as shown in Fig. 5c. In con-

Table 2 Consensus scores for marine bioactive compounds with M^pro PDB ID 6lu7.

	Name	Consensus score	Key interactions
1	Polyacetylenetriol	56	Hydrophobic-hydrophobic interactions
2	4,5-dihydro-6-deoxybromotopsentin	63	HB with Arg188
3	Sceptrin	72	HBs with Glu166, Gln189, and two HBs with Asn142. overlay with ligand
4	bromotopsentin	72	HB (strong) with Arg188
5	Topsentin	74	Weak HB with His41
6	Manzamine A	80	Hydrophobic-hydrophobic interactions
7	22-O-(N-Me-L-valyl)-21-epiaflaquinolone B	84	HBs with Glu166, Gln189
8	lornemide	92	Two HBs with peptide Gln190-Thr192
9	Ageliferin 1	95	HBs with Glu166 (two HB), His164, Gly143. overlay with ligand
10	Aplidiasphingosine	102	HBs with Glu166, Gln189
11	isobutyrolactone	115	HB with Gly143
12	Thalassoilin B	115	HB with Gly143
13	Thalassoilin C	125	HB with Arg188. Hydrophobic-hydrophobic interactions
14	2-(4-hydroxybenzoyl)quinazolin4(3H)-one	127	HBs with Glu166, and two HBs with Ser144
15	Pulvic acid	135	HB with Glu166
16	Alterporriol Q	136	HB with Cys145
17	caulerpin	137	Hydrophobic-hydrophobic interactions
18	Agelifenin	139	HBs with Gln189; HB with Ser144
19	Dragmacidin F	157	HB with His164.
20	TAN-931	162	HBs with Glu166 and His41.
20	Thalassoilin A	168	HB with Gly143
21	N3 (ligand)		HBs with Ser144, Glu:166, Gln189, Ala19.

Table 3 Consensus scores for marine bioactive compounds with M^pro PDB ID 6y2f.

No.	Name	Consensus score	Key interactions
1	atomaric acid	63	3HBs with peptide Gln192, Thr190, and Gln189
2	bromotopsentin	66	Hydrophobic-hydrophobic interactions
3	debromo sceptrin	74	Similar to co crystalized ligand. HBs with Glu166, Llu167; Met165, Thr25.
4	Polyacetylenetriol	89	Hydrophobic-hydrophobic interactions
5	Thalassoilin B	90	HB with Glu192
6	Thalassoilin C	91	HBs with Thr19, Gln192.
7	Thalassoilin A	95	Glu192, His41
8	Topsentin	106	Hydrophobic-hydrophobic interactions
9	Dragmacidin F	114	Val186, His41, Leu141
10	Macrolactin A	115	Hydrophobic-hydrophobic interactions. Two HBs with Ser144
11	22-O-(N-Me-L-valyl)-21-epiaflaquinolone B	116	Two HBs with Thr25, Val42
12	Neocitreoviridin	121	Two HBs with Gln192, Arg188
13	4,5-dihydro-6-deoxybromotopsentin	124	HB with Arg188
14	8,10,18-trihydroxy-2,6-dolabelladiene	128	HBs with THR: 190A, ALA:191A
15	Acarnidine C	128	Hydrophobic-hydrophobic interactions. Gunanidine moiety HB with SER:144 A.
16	Aplidiasphingosine	133	Ethanolamine moiety forms four HBs with CYS:145 A, SER: 144 A, GLY: 143 A, and SER: 144 A. another terminal hydroxyl forms HB with HIS:41 A.
17	3β-hydroxyergosta-8,14,24(28)-trien-7-one	135	Hydrophobic-hydrophobic interactions
18	Sceptrin	138	Multiple HBs with different pose from debromo sceptrin
19	Manzamine A	153	Hydrophobic-hydrophobic interactions
20	α-ketoamide (ligand)	158	HBs with Thr25–His 41, Asn142-Ser144, Glu166, Gln189, Ala191.

trast, an overlay was identified for all acyl indole alkaloids (topsentin derivatives), as shown in Fig. 5d.

2.1.2. Docking with the spike glycoporotein of SARS-CoV-2 (PDB ID: 6vsb)

Although more than 50 compounds showed stronger interaction than the standard ligand, most of them exhibited different binding modes or docked outside the created receptor. Compounds containing long-chain hydrocarbon parts behaved stickier in association with standard ligands. Fig. 6a displayed polyacetylenetriol, adopting the same interaction domain to the standard. However, as a result from M^pro docking studies,

Table 4 Consensus scores for marine bioactive compounds with spike glycoprotein 6vsb.

No.	Name	Consensus score	Key interactions
1	Thalassoilin B	30	HBs with Lys417 (sulphate part), Ser375, Phe377
2	Polyacetylenetriol	47	HB Tyr495. Hydrophobic hydrophobic interactions
3	Alterporriol Q	54	HBs with Tyr421, Phe374
4	Topsentin	63	HB Phe374, hydrophobic-hydrophobic interaction
5	Calyceramide C	84	HBs with Phe377 (through carbonyl), Gln409 (through hydroxyl functionality)
6	4,5-dihydro-6-deoxybromotopsentin	96	HBs with Phe377, Tyr369.
7	Calyceramide B	104	HBs with Phe377 (through carbonyl), Gln409 (through hydroxyl functionality)
8	Bromotopsentin	104	HBs with Phe377, Tyr369.
9	Thalassoilin A	114	HBs with Tyr375, Lys417
10	Varibiline	117	Two HBs with Asn422 through lactone group
11	TAN-931	126	HB Gln493.
12	Aspernolide A	131	HB with Thr415
13	Antimycin A10	132	hydrophobic-hydrophobic interaction
14	Dragmacidin F	134	HB with Asn422
15	Thalassoilin C	135	HBs with Phe377, Ser 375.
16	Crossasteroside B	149	HB Lys378 and hydrophobic-hydrophobic interactions
17	3β-hydroxyergosta-8,14,24(28)-trien-7-one	154	hydrophobic-hydrophobic interactions
18	Dehydrofurodendin	157	hydrophobic-hydrophobic interactions
19	debromo sceptrin	160	Outside the created receptor
20	S ligand	422	Two HBs with Asn422, HB with Pro491

Table 5 Consensus scores for marine bioactive compounds with RNA polymerase (PDB: ID 6 m71).

	Name	Consensus score	Key interactions
1	Lornemides A	7	Docked deeply with formation of two HBs with Tyr217 and Asp218.
2	bromotopsentin	44	hydrophobic-hydrophobic interaction
3	4,5-dihydro-6-deoxybromotopsentin	47	HB with Arg166
4	topsentin	49	HB with Arg166
5	Polyandrocarpidine A	66	HB with Phe219
6	Thalassoilin C	73	HBs with Phe219 and Asn79 through sugar moiety
7	Aplidiasphingosine	91	HB with Arg166 via OH near prenyl group
8	Acarnidine C	97	HB with Thr76 by guanidine moiety
9	Aspernolide A	104	Outside the receptor
10	Thalassoilin A	121	HB with Phe219 Carbonyl at C4.
11	thyrsiferyl 23- acetate	121	Docked with outer surface of receptor
12	Polyandrocarpidine B	138	HB with Asn79
13	Antimycin A1	139	Docked with outer surface of receptor but formed HB with Arg116.
14	Thalassoilin B	143	Out of the receptor domain but in side the inner grid
15	Dragmacidin F	144	Out of the receptor domain but in side the inner grid
16	sceptrin	155	Out of the receptor domain but in side the inner grid
17	2-(4-hydroxybenzoyl) quinazolin4(3H)-one	156	HB with Asp218 via carbonyl of phenol part.
18	polysaccharides kappa/beta-carrageenan	158	Outer surface with formation of two HBs with Arg116
19	Pulvic acid	159	HB with Phe219
20	3β-hydroxyergosta-8,14,24(28)-trien-7-one	163	Out of the receptor domain but inside the inner grid
21	Remdesivir	165	Two HBs with Gln166 and with HB with Asn79

some compounds adopted position outside the created receptor on the opposite side of the standard ligand. Among compounds incorporating the indole architecture, dragmacidin F was found to interact with receptor domains in the same clefts as the standard ligand, as shown in Fig. 6b. Compounds, including topsentin, calyceramide C, bromotopsentin, and the 4,5-dihydro-6-deoxybromotopsentin cluster inside the inner grid in the same domain away from standard sets were shown in Fig. 6c.

2.1.3. Docking with RNA polymerase of SARS-CoV-2 (PDB ID: 6m71).

Remdesivir is an anti-polymerase drug prescribed for the alleviation of the COVID-19 disease. Its docking mode showed its ability to form two HBs with Gln166 through its oxygen of the furanose ring (strong) and via the oxygen of the carbonyl functionality. Additionally, it participated in another HB with

Fig. 4 visual representation by Vida: **a)** Polyacetylenetriol occupied the receptor of 6lu7 domains in a hydrophobic-hydrophobic interaction; **b)** deoxybromotopsentin and bromotopsentin; **c)** sceptrin overlay with co crystalized ligand; **d)** Thalassoilin A and thalassoilin B illustrated same binding mode and pose.

Asn79 through the phosphor atom, as shown in Fig. 7a. Bromotopsentin, 4,5-dihydro-6-deoxybromotopsentin, and topsentin adopted the same position in the receptor and both 4,5-dihydro-6-deoxybromotopsentin and topsentin participated in the formation of HB with the essential amino acid Arg166, as depicted in Fig. 7b. The carbonyl was found to be essential for the activity as well as act in the molecule as a rigid amide. Many compounds with strong interactions, such as antimycin A1, thyrsiferyl 23- acetate, aspernolide A, thalassoilin B, dragmacidin, and the sceptrin cluster on the outer surface of the receptor and the side the inner grid closed the essential amino acid clefts, as shown in Fig. 7c.

2.2. Ligand Based Approach

2.2.1. Shape alignment and rapid overlay chemical structure (ROCS)

Rapid Overlay Chemical Structure (ROCS) is a chemoinformatic screening technique utilized to perceive a similarity between chemical entities based on their three-dimensional shape (Hawkins et al., 2007). **The** 3D shape structure displays good neighborhood behavior in which high a similarity in the shape reflects the high similarity in their biology in cases where a high similarity in biology is not reflected in a similarity in the structure. (Elbastawesy et al., 2020; Masek et al., 1993)

Shape similarity using the ROCS tool has different applications, including virtual screening, lead-hopping, molecular alignment, pose generation, as well as structural predictions (Hawkins et al., 2007). The approach of this study was to identify molecules that can adopt shapes which are extraordinarily similar to reported drug candidates (the query), as shown in Fig. 8–10. Based on this concept, matches are created according to only the volume of overlapping of optimally aligned molecules, which are virtually independent from the atom types and the bonding patterns of the query.

The ROCS study requires two files which must be in 3D format in case of most stable conformers generated by the Omega application in the Openeye software: a) a database file, which includes the collection of all compounds for this study; b) a query file, which includes the standard (reference) molecules or the lead compound used for the inhibition of SARS-CoV-2 replications.

The outputs from the ROCS analysis were 1) a shape counter, shape atoms, and color atom labels for the database set as well as the query compounds, as shown in Table 6 and Fig. 8–10. The shape was represented with a series of dotted lines around the molecule, while the color feature was shown as a filled colored circle representing the different kinds of chemical features; 2) the overlay, which was considered to be the alignment between the database as well as the molecules and the query, as visualized by the vROCS and VIDA applications; and 3) the set of scores were expressed in Tanimoto scores.

The most important score is the Tanimoto Combo (TC) that includes both shape fit and color. This score had a value between 0 and 2 and was the score used for ranking the hit list.

Based on previous docking results, the compounds were found to express multi target actions as most of them exhibited high scores with the different targets (main protease, RNA polymerase, or spike glycoprotein). In this regard, these marine candidates were selected to be the subjects of further cheminformatics studies. Compounds formed flavonoids nucleolus (thalassoilin B, thalassoilin C and subsequently thalassoilin A) showed higher TC scores (0.6800–0.4980) than PF-00835231 and α-ketomide (0.4960, 0.4950 respectively) in

Fig. 5   visual representation by Vida: **a)** Debromo sceptrin reprsents unique and strong interaction similar to co crystalized ligand N3; **b)** debromo sceptrin (pink) and Sceptrin (blue) not co overlay; **c)** Thalassoilin B and thalassoilin A overlay completely each other while thalassoilin C exhibited different binding mode; **d)** topsentin derivatives) overlay each other.

Fig. 6   **a)** Polyacetylenetriol adopted same interaction domain to the standard; **b)** dragmacidin F interacts with receptor domains in same clefts as standard ligand; **c)** topsentin, calyceramide C, bromotopsentin and 4,5-dihydro-6-deoxybromotopsentin cluster inside the inner grid.
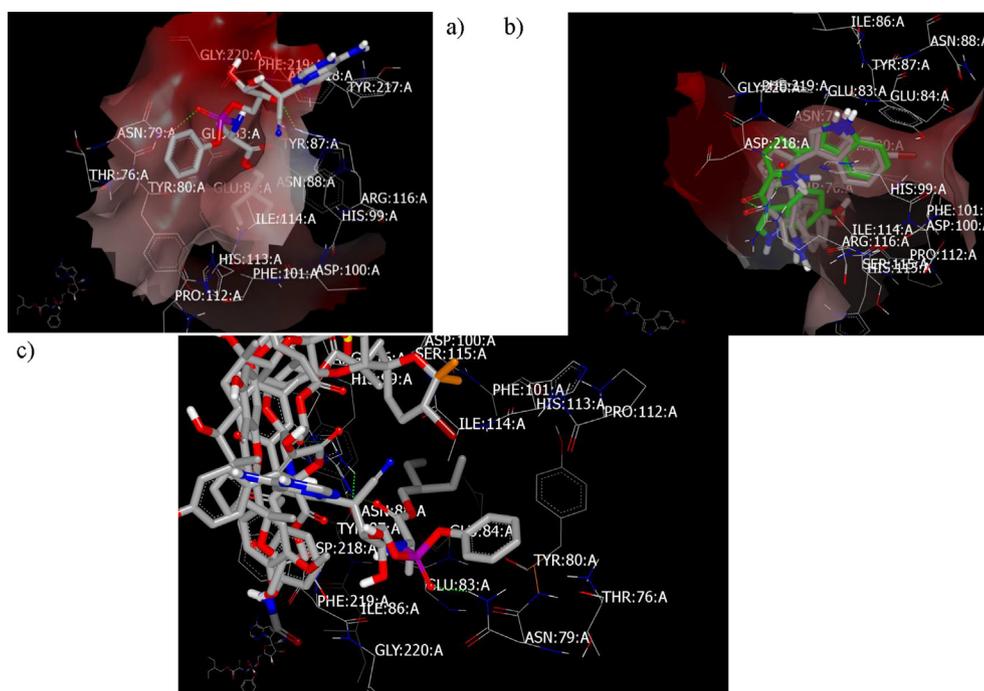
Fig. 7   a) Remidisver docked with formation two HB with GLN:166 A and HB with ASN:79 A; **b)** Bromotopsentin, 4,5-dihydro-6-deoxybromotopsentin and topsentin overlay with 4,5-dihydro-6-deoxybromotopsentin and topsentin; **c)** antimycin A1, thyrsiferyl 23-acetate, aspernolide A, thalassoilin B, dragmacidin, and fsceptrin cluster on the outer surface of the receptor and side the inner grid.
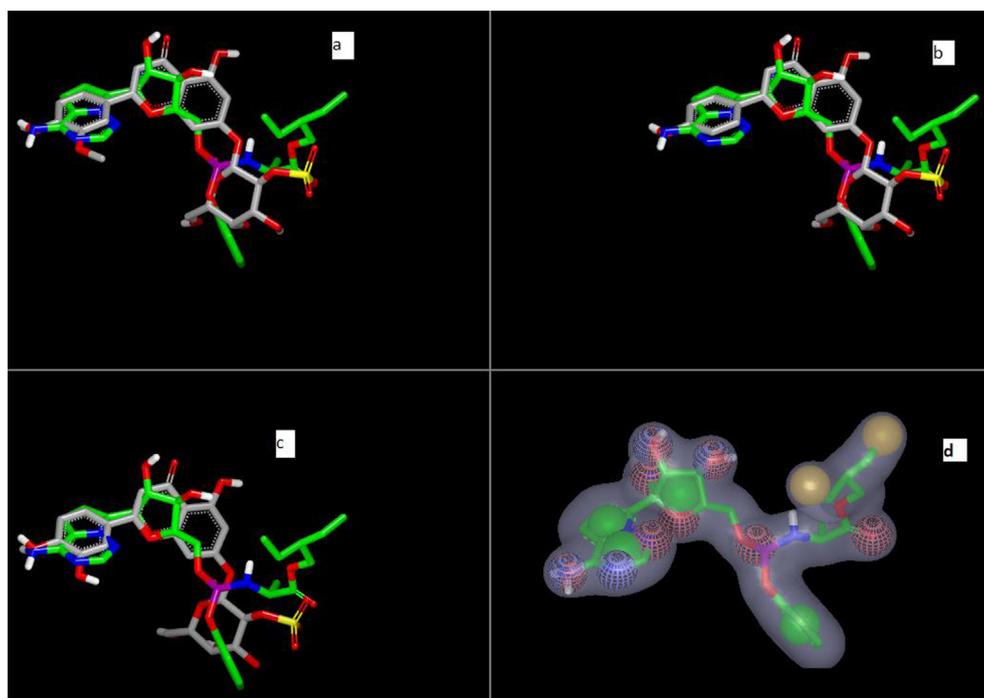
Fig. 8   ROCS view by Vida: **a)** thalassoilin B (grey) overlay with remdesivir (green); **b)** thalassoilin C (grey) overlay with remdesivir (green); **c)** thalassoilin A (grey) overlay with remdesivir (green); **d)** colour shape and volume of remdesivir.

Table 6 Tanimoto Combo scores and shape color and volume for marine compounds to different drugs recommended in COVID19 disease generated by *v*ROCS application.

Compound	Shape and color atoms	Inhibiting the RNA polymerase	Inhibiting the Main protease (PF-07304814)	Blocking Virus-Cell Membrane Fusion
		Remdesivir	PF-07304814	Umifenovir
1 Lornemides A	Ring: 1, donor: 2, acceptor: 2, hydrophobe: 1	0.4020	0.5570	0.7900
2 Atomaric acid	Ring: 3, donor: 1, acceptor: 4, hydrophobe: 1, anion: 1	0.4130	0.5660	0.7170
3 4,5-Dihydro-6-deoxybromotopsentin	Ring: 5, donor: 4, acceptor: 2, hydrophobe: 1	0.4760	0.6720	0.8910
4 Bromotopsentin	Ring: 5, donor: 3, acceptor: 1, hydrophobe: 1	0.4890	0.7070	0.9060
5 Debromo sceptrin	Ring: 5, donor: 8, acceptor: 6, hydrophobe: 1	0.4390	0.4820	0.6470
6 Sceptrin	Ring: 5, Donor: 8, acceptor: 6, hydrophobe: 2	0.4870	0.6130	0.6330
7 Polyacetylenetriol	Donor: 3, acceptor: 3, hydrophobe: 2	0.2980	0.3890	0.5420
8 Thalassoilin B	Ring: 4, Donor: 5, acceptor: 13, anion:1	0.6800	0.6270	0.6670
9 Thalassoilin C	Ring: 4, Donor: 5, acceptor: 12, anion:1	0.6710	0.7220	0.7080
10 Thalassoilin A	Ring: 4, Donor: 6, acceptor: 12, anion:1	0.6580	0.6910	0.6850
11 Topsentin	Ring: 5, donor: 4, acceptor: 2.	0.4960	0.6920	0.9460
12 Dragmacidin F	Ring: 5, donor: 7, acceptor: 3, cation: 3 hydrophobe: 2	0.4510	0.5790	0.7620
13 Macrolactin A	Donor: 3, acceptor: 4	0.4110	0.4520	0.6950
14 Remdesivir	Ring: 4, donor: 4, acceptor: 8, hydrophobe: 2	2.00	0.6260	0.5720
15 α-Ketomide	Ring: 3, donor: 4, acceptor: 5, hydrophobe: 1	0.4980	0.621	0.527
16 PF-07304814	Ring: 3, donor: 3, acceptor: 8, hydrophobe: 1	0.4950	2.00	0.5770
17 PF-00835231	Ring: 3, donor: 4, acceptor: 8, hydrophobe: 1	0.6300	1.5270	0.6890
18 Umifenovir	Ring: 3, acceptor: 2, donor: 1, hydrophobe:1.	0.450	0.512	2

comparison with remdesivir, as shown in Table 6. Furthermore, acyl indole alkaloids, such as topsentin (TC = 0.4960), bromotopsentin (TC = 0.4890), and 4,5-dihydro-6-deoxybromotopsentin (TC = 0.476) were analyzed, among which sceptrin (pyrrole carboxamide system alkaloids) exhibited TC = 0.4870. In the comparison of the 3D shape of marine compounds with that of PF-07304814 as an **M^{pro}** drug candidate, the flavonoid thalassoilin C exhibited higher TC scores than acyl indole alkaloids (bromotopsentin, and topsentin). Moreover, the flavonoids thalassoilin A, 4,5-dihydro-6-deoxybromotopsentin, and thalassoilin B were analyzed, as shown in Table 6. Concerning the comparative analysis of these compounds in association with drugs inhibiting the viral entry, such as umifenovir, the following marine compounds were analyzed: acylindole marine alkaloids (topsentin, bromotopsentin, 4,5-dihydro-6-deoxybromotopsentin respectively), lornemides A, dragmacidin F, atomeric acid, thalassoilin C, macrolactin A, PF-00835231, and subsequently thalassoilin A, exhibiting a high TC score (0.946–0.685, respectively), as shown in Table 6. All marine compounds exhibited TC score more than remdesivir in case of umifenovir as a query.

As a result, flavonoid compounds and acyl indoles exhibited high 3D shape similarity in association with most of the known reported drug candidates. Shape color and volume shape of all compounds were illustrated in Fig. 6 including supplementary data for their figures.

2.2.2. Predicted pharmacokinetics and pharmacodynamics parameters

Physicochemical properties determine the drug-likeness score in case of orally administered drugs. Drug candidates with high drug-likeness scores were previously reported to exhibit higher absorption and bioavailability in lower doses and show fewer drug-drug interaction warnings (Ritchie and Macdonald, 2014). Absorption, distribution, metabolism, excretion, and toxicity (ADMET) calculations contribute to the determination of the failure of approximately 60% of all drugs in the different clinical trial phases. In this regard, ADMET is determined at the beginning of the drug discovery phases to eliminate molecules coming with poor ADMET properties from an earlier drug discovery pipeline to save research costs. As demonstrated in Table 7 and among the

Table 7 Physicochemical Parameters and Predicted pharmacokinetics and pharmacodynamics parameters.

No.	Compound	Lipinskĭs rule (Rule of five)				PreADMET prediction						
		MW	LogP	HBD	HBA	BBB	PPB	PSA	HIA	Skin permeability log Kp (cm/s)	pKa basic/acidic	Drug-Likeness score
1	Lornemides A	273.37	3.01	2	3	Yes		63.32	High	−5.99	−1.05/11.86	−0.29
2	atomaric acid	442.63	5.86	2	4	No		66.76	Low	−3.9	< 0./4.67	0.58
3	bromotopsentin	420.26	3.97	4	2	No		84.67	High	−5.57	−3.41/9.68	−0.69
4	debromo sceptrin	541.4	0.75	8	4	No		199.18	Low	−9.21	9.01/12.63	−0.031
5	Polyacetylenetriol	426.55	5.34	3	3	Yes		60.69	High	−5.49	< 0./13.51	−1.05
6	Thalassoilin B	541.46	0.1	5	14	No		233.86	Low	−8.61	< 0./9.77	0.38
7	Thalassoilin C	511.43	0.11	5	13	No		224.63	Low	−8.4	< 0. /9.77	0.36
8	Thalassoilin A	527.43	−0.27	6	14	No		244.86	Low	−8.75	< 0. /9.12	0.39
9	Topsentin	341.36	3.36	4	2	No		84.67	High	−5.58	−3.41/9.68	−0.53
10	Dragmacidin F	532.39	2.75	7	4	No		147.64	Low	−7.37	9.58/11.31	−0.05
11	Macrolactin A	402.52	3.00	3	5	No		86.99	High	−5.58	< 0./15.03	−1.16
12	4,5-dihydro-6-deoxybromotopsentin	404.26	4.37	3	1	No		64.44	High	−5.22	−3.41/14.47	−0.98

HBD, hydrogen bond donor; HBA, hydrogen bond acceptor; BBB, blood brain barrier; PPB, plasma protein binding; HIA, percentage human intestinal absorption; MolPSA (Molecular Polar Surface Area (PSA); Caco-2 value, permeability to Caco-2 (human colorectal carcinoma) cells in vitro.

selected compounds, atomeric acid displayed the highest drug-likeness score (0.56), followed by flavonoids marine compounds (thalassoilin B, thalassoilin C, thalassoilin A) with a drug-likeness score in the range of 0.39–0.36. Atomeric acid had the highest lipophilicity value with a logP value of 5.86 and, therefore, low hydrophilicity, causing poor absorption or permeation. This result suggested us to semi-synthesize this compound in an ester form to optimize the lipophilicity to be a potential orally active compound.

3. Conclusion

Marine compounds with a flavonoid core, acyl indole, and pyrrole carboxamide alkaloids exhibited high scores in case of different virtual screening aspects in this study. Sceptrin and debromo sceptrin displayed complete overlay with the co-crystallized ligands of M^{pro}. Thalassoilin (A-B) exhibited the highest binding and similarity results among all the target proteins. These data highlight the importance of marine natural metabolites in regard of the conduction of future studies to discover new drugs to combat COVID-19.

4. Experimental

4.1. Data base collection

Marine bioactive compounds were retrieved from literatures, and reported drugs were identified from known data base collection.

4.2. Molecular modeling

4.2.1. Docking

The molecular docking studies were operated using the Open-Eye Modeling software [Fast Rigid Exhaustive Docking (FRED) Receptor, version 2.2.5; OpenEye Scientific Software, SantaFe, NM (USA); http://www.eyesopen.com], academic license (The Laboratory of Yaseen A. M. Mohamed Elshaier). A virtual library of target compounds was used, and their energies were minimized using the MMFF94 force field, followed by the generation of multi-conformers using the OMEGA application. The library was compiled in one file by Omega. The target proteins were retrieved from PDB and the created receptor was operated by OeDocking application. Both the ligand input file and the receptor input file were subjected to FRED to implement the molecular docking study. Multiple scoring functions were engaged to predict energy profile of the ligand-receptor complex. The vida application was used as a visualization method to represent the ligands pose and the potential binding interactions of the ligands to the receptor of interest.

4.2.2. Shape similarity and ROCS analysis

Basic method to represent shape and color features in ROCS is using ROCS application Open Eye scientific software. The query molecules were selected based on high similarity. [https://www.eyesopen.com/] Compounds library was adopted as the database file. Both query and database files were energy minimized by Omega applications. Personal PC in very fast

using vROCS interface employed ROCS runs. vROCS was employed to run and analyze/visualize the results. ROCS application searched the database with the query to find molecules with similar shape and colors. The result was visualized by Vida application. Compounds conformers were scored based upon the Gaussian overlap to the query and the best scoring parameters is Tanimoto Combo scores (shape + color), the highest score is the best matched with query compound.

4.3. ADME prediction

Lipinskǐs rule (Rule of five) and molecular property prediction was calculated at the following free access website https://www.molsoft.com/servers.html.

Regarding the PreADMET estimation, it was determined through utilizing the free access of the website https://preadmet.bmdrc.kr/.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.arabjc.2021.103092.

References

Allam A.E., Assaf, H.K., Hassan, H.A., Shimizuc, K., Elshaier, Y.A. M.M., 2020. An in silico perception for newly isolated flavonoids from peach fruit as privileged avenue for a countermeasure outbreak of COVID-19. RSC Adv., 10, 29983.

Battisti, V., Wieder, O., Garon, A., Seidel, T., Urban, E., Langer, T., 2020. A computational approach to identify potential novel inhibitors against the coronavirus SARS-CoV-2. Molecular Informatics, 39(10), e2000090.

Boopathy, N.S., Kathiresan, K., 2010. Anticancer drugs from marine flora: an overview. J. Oncol., 2010.

Cavasotto, C.N., Di Filippo, 2021. In silico drug repurposing for COVID-19: targeting SARS-CoV-2 proteins through docking and consensus ranking. Mol. Inform., 40(1), e2000115.

Cavasotto, C.N., Lamas, M.S., Maggini, J., 2020. Functional and druggability analysis of the SARS-CoV-2 proteome. Eur. J. Pharmacol.

Cavasottoa, C.N., Lamas, M.S., Magginice, J., Functional and druggability analysis of the SARS-CoV-2 proteome. Eur. J. Pharmacol., 890, 173705.

Che, C., 1991. Marine products as a source of antiviral drug leads. Drug Dev. Res. 23, 201–218.

Chen, B., Tian, E., He, B., 2020. Overview of lethal human coronaviruses. Sig Transduct. Target Ther. 5 (89).

Ciliberto, G., Mancini, R., Paggi, M.G., 2020. Drug repurposing against COVID-19: focus on anticancer agents. J. Exp. Clin. Cancer Res., 39(86).

Culletta, G., Gulotta, M.R., Perricone, Ugo, Zappalà, Maria, Almerico, Anna Maria, Tutone, Marco, 2020. Exploring the SARS-CoV-2 proteome in the search of potential inhibitors via structure-based pharmacophore modeling/docking approach. Computation, 8(3), 77.

Elbastawesy, M.A., El-Shaier, Y.A., Ramadan, M., Brown, A.B., Aly, A.A., Abuo-Rahma, G.E.D.A. 2020. Identification and molecular modeling of new quinolin-2-one thiosemicarbazide scaffold with antimicrobial urease inhibitory activity. Mol. Divers.

Ferraz, W.R., , R.A., Novaes, S., A.L., Goulart Trossini, G.H., 2020. Ligand and structure- based virtual screening applied to the SARS-CoV-2 main protease: an in silico repurposing study. Future Med. Chem., 12(20), 1815-1828.

Gao, Y., Yan, L., Huang, Y., Liu, F., Cao, L., Wang, T., Wang, Q., Lou, Z., Rao, Z., 2020. SARS-Cov-2 RNA-dependent RNA polymerase in complex with cofactors. RCSB Protien Data Bank.

Glowacka, I.B., Bertram, S., Muller, M.A., Allen, P., Soilleux, E., Pfefferle, S., Steffen, I., Tsegaye, T.S., He, Y., Gnirss, K., 2011. Evidence that TMPRSS2 activates the severe acute respiratory syndrome coronavirus spike protein for membrane fusion and reduces viral control by the humoral immune response. J. Virol., 2011. 85, 4122-4134.

Hawas, U.W., Abou El-Kassem, L.T., 2017. Thalassiolin D: a new flavone O-glucoside Sulphate from the seagrass Thalassia hemprichii. Nat. Prod. Res. 31 (20), 2369–2374.

Hawkins, P.C.D., Skillman, A.G., Nicholls, A., 2007. Comparison of shape-matching and docking as virtual screening tools. J. Med. Chem., 50(1), 74-82.

Heurich, A., Hofmann-Winkler, H., Gierer, S., Liepold, T., Jahn, O., Pohlmann, S., 2014. TMPRSS2 and ADAM17 Cleave ACE2 differentially and only proteolysis by TMPRSS2 augments entry driven by the severe acute respiratory syndrome coronavirus spike protein. J. Virol. 88, 1293-1307.

Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., Cheng, Z., Yu, T., Xia, J., Wei, Y., Wu, W., Xie, X., Yin, W., Li, H., Liu, M., Xiao, Y., Gao, H., Guo, L., Xie, J., Wang, G., Jiang, R., Gao, Z., Jin, Q., Wang, J., Cao, B., 2019. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China Lancet, 395, 497-506.

Ibrahim, M.A.A. et al, 2020a. Natural-like products as potential SARS-CoV-2 Mpro inhibitors: in-silico drug discovery. J. Biomol. Struct. Dyn., 1–13.

Ibrahim, M.A., Abdelrahman, A.H., Hussien, T.A., Badr, E.A.A., Mohamed, T.A., El-Seedi, H.R., Pare, P.W., Efferth, T., Hegazy, M.F., 2020. In silico drug discovery of major metabolites from spices as SARS-CoV-2 main protease inhibitors. . Comput Biol Med. 126, 104046.

Ibrahim, M.A., Abdelrahman, A.H., Hegazy, M.E.F., 2020. In-silico drug repurposing and molecular dynamics puzzled out potential SARS-CoV-2 main protease inhibitors J. Biomol. Struct. Dyn.

Ibrahim, M.A., Abdelrahman, A.H., Allemailem, K.S., Almatroudi, A., Moustafa, M.F., Hegazy, M.F., 2021. In silico evaluation of prospective anti-COVID-19 drug candidates as potential SARS-CoV-2 main protease inhibitors. Protein J. 1-14.

Jimeno, J., Faircloth, G., Fernández Sousa-Faro, J.M., Scheuer, P., Rinehart, K., 2004. New marine derived anticancer therapeutics—a journey from the sea to clinical trials. Marine Drugs, 2, 14-29.

Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y., Zhang, B., Li, X., Zhang, L., Peng, C., et al, 2020. Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. Nature 582, 289–293.

Jin, Z., Du, X., Yang, H., 2020. Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. Nature 582, 289-293.

Jin, Z., Zhao, Y., Sun, Y., Zhang, B., Wang, H., Wu, Y., Zhu, Y., Zhu, C., Hu, T., Du, X., Duan, Y., Yu, J., Yang, X., Yang, X., Yang, K., Liu, X., Guddat, L.W., Xiao, G., Zhang, L., Yang, H., Rao, Z.,

2020. Structural basis for the inhibition of SARS-CoV-2 main protease by antineoplastic drug carmofur. Nat. Struct. Mol. Biol. 27, 529–532.

Johnson, M.A., Chatterjee, A., Neuman, B.W., Wuthrich, K., 2010. SARS coronavirus unique domain: three-domain molecular architecture in solution and RNA binding. J. Mol. Biol. 400, 724–742.

Lai, C.C., Shih, T.P., Ko, W.C., Tang, H.J., Hsueh, P.R., 2020. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): the epidemic and the challenges. Int. J. Antimicrob Agents, 55(3), 105924.

Li, W., Moore, M.J., Vasilieva, N., Sui, J., Wong, S.K., Berne, M.A., Somasundaran, M., Sullivan, J.L., Luzuriaga, K., Greenough, T.C., Choe, H., Farzan, M., 2003. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature, 426, 450-454.

Li, X., Geng, M., Peng, Y., Meng, L., Lu, S., 2020. Molecular immune pathogenesis and diagnosis of COVID-19. J. Pharm. Anal. 19, 1-7.

Masek, B.B., Merchant, A., Matthew, J.B., 1993. Molecular shape comparison of Angiotensin II receptor antagonists, J. Med. Chem. 36(9), 1230-1238.

Meyer-Almes, F.J., 2020. Repurposing approved drugs as potential inhibitors of 3CL- protease of SARS-CoV-2: Virtual screening 60 and structure based drug design. Comput. Biol. Chem. 88, 107351.

Ritchie, T.J., Macdonald, S.J.F., 2014. How drug-like are 'ugly' drugs: do drug-likeness metrics predict ADME behaviour in humans?. Drug Discovery Today 19 (4), 489–495.

Senanayake, S.L., 2020. Drug repurposing strategies for COVID-19. Future Drug Discov., 2(2).

Shyr, Z.A., Gorshkov, K., Chen, C.Z., Zheng, W., 2020. Drug discovery strategies for SARS- CoV-2. J. Pharmacol. Exp. Ther. 375(1), 127-128.

Touret, F., Gilles, M., Barral, K., Nougairède, A., van Helden, J., Decroly, E., de Lamballerie, X., Coutard, B. 2020. In vitro screening of a FDA approved chemical library reveals potential inhibitors of SARS-CoV-2 replication. Sci. Rep., 10(1), 13093.

Worldometer, W., 2020. Coronavirus Update: Cases And Deaths From COVID-19 Virus Pandemic–Worldometer. info (2020).[online] Available at:[Accessed 27 May 2020].

Wrapp, D., Wang, N., Corbett, K.S., Goldsmith, J.A., Hsieh, C.L., Abiona, O., Graham, B.S., McLellan, J.S., 2020. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science 367(6483), 1260-1263.

Zhang, L., Lin, D., Sun, X., Hilgenfeld, R. Crystal structure (monoclinic form) of the complex resulting from the reaction between SARS-CoV-2 (2019-nCoV) main protease and tert-butyl (1-((S)-1-(((S)-4-(benzy-lamino)-3,4-dioxo-1-((S)-2-oxopyrrolidin-3-yl)butan-2-yl)amino)-3-cyclopropyl-1-oxopropan-2-yl)-2-oxo-1,2-dihydropyridin-3-yl)carba-mate (alpha-ketoamide 13b). Protien Data Bank.