Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Computers in Biology and Medicine 141 (2022) 104714

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Abrogation of SARS-CoV-2 interaction with host (NRP1) neuropilin-1 receptor through high-affinity marine natural compounds to curtail the infectivity: A structural-dynamics data

Fahad Humayun [a,1], Abbas Khan [a,1], Sajjad Ahmad [b], Wang Yuchen [c], Guoshen Wei [d], N. Nizam-Uddin [e], Zahid Hussain [f], Wajid Khan [f], Nasib Zaman [f], Muhammad Rizwan [f], Muhammad Waseem [g], Dong-Qing Wei [a,h,i,*]

[a] Department of Bioinformatics and Biological Statistics, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, 200240, PR China
[b] Department of Health and Biological Sciences, Abasyn University, Khyber Pakhtunkhwa, Pakistan
[c] Beijing 161 High School, No. 94, Nanheng West Street, Xicheng District, Beijing, PR China
[d] Yangpu High School, Yangpu, Shanghai, PR China
[e] Biomedical Engineering Department, HITEC University, Taxila, Pakistan
[f] Center for Biotechnology and Microbiology, University of Swat, Swat, KP, Pakistan
[g] Faculty of Rehabilitation and Allied Health Science, Riphah International University, Islamabad, Pakistan
[h] Peng Cheng Laboratory, Vanke Cloud City Phase I Building 8, Xili Street, Nashan District, Shenzhen, Guangdong, 518055, PR China
[i] State Key Laboratory of Microbial Metabolism, Shanghai-Islamabad-Belgrade Joint Innovation Center on Antibacterial Resistances, Joint Laboratory of International Cooperation in Metabolic and Developmental Sciences, Ministry of Education and School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200030, PR China

ARTICLE INFO

Keywords:
SARS-CoV-2
Marine natural products
NRP1
Virtual drugs screening
KD. (dissociation constant)
Biophysical simulation

ABSTRACT

The evolution of new severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) variants around the globe has made the coronavirus disease 2019 (COVID-19) pandemic more worrisome, pressuring the health care system and resulting in an increased mortality rate. Recent studies recognized neuropilin-1 (NRP1) as a key facilitator in the invasion of the new SARS-CoV-2 into the host cell. Therefore, it is considered an imperative drug target for the treatment of COVID-19. Hence, a thorough analysis was needed to understand the impact and to guide new therapeutics development. In this study, we used structural and biomolecular simulation techniques to identify novel marine natural products which could block this receptor and stop the virus entry. We discovered that the binding affinity of CMNPD10175, CMNPD10017, CMNPD10114, CMNPD10115, CMNPD10020. CMNPD10018, CMNPD10153, CMNPD10149 CMNPD10464 and CMNPD10019 were substantial during the virtual screening (VS). We further explored these compounds by analyzing their absorption, distribution, metabolism, and excretion and toxicity (ADMET) properties and structural-dynamics features. Free energy calculations further established that all the compounds exhibit stronger binding energy for NRP1. Consequently, we hypothesized that these compounds might be the best lead candidates for therapeutic interventions hindering virus binding to the host cell. This study provides a strong impetus to develop novel drugs against the SARS-CoV-2 by targeting NRP1.

1. Introduction

The global coronavirus disease 2019 (COVID-19) pandemic was caused by acute respiratory syndrome coronavirus 2 (SARS-CoV-2), a highly contagious virus first identified in Wuhan, China [1]. In 2013, a similar coronavirus named SARS-CoV caused an epidemic. SARS-CoV-2 infects the lower respiratory system and spreads quickly, mainly via pharyngeal viral shedding [2,3]. Angiotensin-converting enzyme 2

* Corresponding author. Department of Bioinformatics and Biological Statistics, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, 200240, PR China.
E-mail addresses: fahadhamayun@sjtu.edu.cn (F. Humayun), abbaskhan@sjtu.edu.cn (A. Khan), sahmad@bs.qau.edupk (S. Ahmad), 772961234@qq.com (W. Yuchen), 2178970883@qq.com (G. Wei), nizam.uddin@hitecuni.edu.pk (N. Nizam-Uddin), dqwei@sjtu.edu.cn (D.-Q. Wei).
[1] Contributed Equally.

https://doi.org/10.1016/j.compbiomed.2021.104714
Received 4 May 2021; Received in revised form 5 July 2021; Accepted 27 July 2021
Available online 31 July 2021
0010-4825/© 2021 Elsevier Ltd. All rights reserved.

(ACE2), a host receptor, is responsible for the uptake of SARS-CoV and SARS-CoV-2 viruses and is thus considered a primary target for viral entry [4,5].

Recently, neuropilin-1 (NRP1) was identified as an additional entry channel that facilitates the entry of SARS-CoV-2 into the host cell. NRP1 is a member of the catalytic and signalling proteins family that is reported to be involved in the cellular invasion of SARS-CoV-2 and to potentially cause *in vitro* infectivity of SARS-CoV-2 [6]. However, it is uncertain whether NRP1 can enable the receptor-mediated endocytosis of the virus. NRP1 is also recognized as a cellular signalling and cell surface receptor [7,8]. The host receptor binding protein of SARS-CoV-2, known as the spike protein, is cleaved into S1 and S2 polypeptides by protease and furin within the host cell, thus exposing the CendR motif in the cleaved S1 protein. The CendR motif is named after the C-end terminal rule requiring the presence of a cationic amino acid, generally arginine, at the carboxyl terminus, which causes an RXXR configuration in the ligand. The binding pocket of CendR is present in NRP1's b1 domain [9]. Daly et al. found that the virus's infectivity is caused by the binding of the CendR motif to NRP1 in the S1 protein. Neutralization of the virus is initiated by the attachment of a monoclonal antibody to the extracellular binding pocket (ECP) of the CendR region of the b1 domain in NRP1 [10]. The CendR motif b1 domain is exposed as soon as the viral spike protein is cleaved by the proteases and thus mediates the binding and entry through NRP1 [11]. This motif's C-end terminal rule name is due to the interspersed cationic amino acids, particularly arginine, that result in an RXXR configuration [12–14]. According to Daly et al. [7], the CendR motif in the SARS-CoV-2 S1 protein interacts with NRP1 and increases viral infectivity. NRP1 RNA expression analysis of bronchial and alveolar cells isolated from patients with severe COVID-19 revealed elevated expression of NRP1 in SARS-CoV-2-positive cells but not in uninfected cells [8]. Blocking the association between NRP1 and SARS-CoV-2 may be a useful therapy in the battle against COVID-19, but further research is required. The results of these studies show that in addition to the function of ACE2 in causing the cellular entry of SARS-CoV-2, NRP1 can also perform as a mediator to increase the virus's infectivity.

The conclusions of these well-designed studies indicate that NRP1 may act as a host cell mediator to enhance the virus's attachment and entry into host cells; thus, it may maximize tissue tropism of SARS-CoV-2 [15]. Therefore, in the present study, molecular modelling-based approaches were employed to target NRP1 in order to identify novel hits against SARS-CoV-2 within a comprehensive marine natural products (CMNP) database (https://www.cmnpd.org/). Computational structural biology and biophysical approaches are the most widely used approaches in drug discovery for the treatment of viral pathogens [16–25]. We used virtual drug screening approaches and binding free energy calculations to decipher the binding, possible mechanism, and long-term association strength between the identified drugs and NRP1. These efforts to discover effective drug candidates for the treatment of COVID-19 will help combat the infection caused by SARS-CoV-2.

2. Materials and methods

2.1. Marine library filtration

Before structure-based VS, the CMNP database was filtered based on drug-like soft, toxicophores, removal of Pan Assay Interference (PAINS) and Eli Lilly MedChem rules in FAFDrugs4 server (https://fafdrugs4.rpbs.univ-paris-diderot.fr/) to obtain drug-like, non-toxic, and non-PAINS compounds [26]. The different parameters used in discarding non-drug-like molecules are presented in Table 1.

Table 1
Different drug-like parameters used in CMNP filtration.

Parameters	Value
Molecular weight	100–600
logP	−3 to 6
Hydrogen bond acceptor	≤12
Hydrogen bond donors	≤7
tPSA	≤180
Rotatable Bonds	≤11
Rigid Bonds	≤3
Rings	≤6
Max size system ring	≤18
Carbons	3–35
HeteroAtoms	1–15
H/C Ratio	0.1 to 1.1
Charges	≤4
TotalCharge	−4 to 4
RO5 Violations	No

2.2. Protein and ligand structure preparation

The crystal structure of NRP1 (PDB ID: 1KEX) was retrieved from Protein Databank (http://www.rcsb.org/) [27]. The structure was prepared by removing water molecules and co-crystallized ligands using the protein preparation wizard incorporated in Schrödinger software [28]. The Amber force field was used for protein's energy minimization using Chimera [28,29]. For ligands preparation, the filtered compounds were imported to PyRx and subjected to energy minimization using MM2 force field [30].

2.3. High-throughput VS

AutoDock Vina was used to virtually screen compounds at the binding site of protein collectively formed by residues Tyr297, Trp301, Thr316, Asp320, Ser346, Thr349 and Tyr353 [31]. The grid box size was defined as 6.04 x −70.97 × 24.028, while the grid dimensions of 56 Å × 58 Å × 50 Å were generated. The aforementioned experimentally validated residues were selected to specify the final active site grid and the maps generated by the AutoDock tool. Three steps of VS were performed to screen the CMNP database [32]. Initially, the whole database was screened using AutoDock Vina with exhaustiveness set as 16, and then the top hits were screened again using exhaustiveness as 32. Finally, for the best scoring compounds, induced-fit docking(IFD) was carried out using 64 exhaustiveness to confirm the final hits. For IFD, we used AutoDockFR–AutoDock for Flexible Receptors (ADFR) software which handles the receptor sidechain conformational optimization up to 14 different sidechains to enhance the success rate of docking [33]. AutoDockFR achieved higher accuracy than AutoDock Vina in cross-validation, and also the speed of the docking is much higher.

2.4. Absorption, distribution, metabolism, excretion and toxicity (ADMET) properties and bioactivity prediction

For the top hits, parameters like drug-like, lead likeness, toxicity, pharmacokinetics and solubility were also explored using SwissADME [34]. Molinspiration, a cheminformatic tool (https://www.molinspiration.com/cgi-bin/properties), was used to foresee the IC_{50} value of each compound. For instance, Molinspiration is widely employed by thousands of studies to forecast bioactivity scores.

2.5. Determination of dissociation constant (K_D)

Furthermore, to give a persuasible understanding of the dissociation constant (K_D), PROtein binDIng enerGY prediction (PRODIGY) (https://wenmr.science.uu.nl/prodigy/) server was used to computationally predict K_D values for different biological complexes [35].

2.6. Molecular dynamics simulation of protein-ligand complexes

All-atoms MD simulation of the top hits from the CMNP database was performed using the AMBER18 package [36]. For drug topologies,

antechamber module was used while the Amber general force field (GAFF) and ff14SB forcefields were employed for parameterization of the ligands and NRP1, respectively. To solvate each protein-ligand complex, a TIP3P water box was used, while Na + counter ions were used to neutralize each system subsequently. Energy minimization of systems was carried out in two stages (steepest descent and conjugate gradient), followed by heating and equilibration [23]. The Particle Mesh Ewald (PME) algorithm was used to treat long-range electrostatic interactions [37]. A 1.4 nm cutoff value was set for Van Der Waal interactions, as well as for Columbic interactions of short-range. Langevin thermostat was employed to temperature constant at 300 K, whereas for pressure control, Berendsen barostat was considered. A time step of 2fs and a total simulation time of 200 ns for each complex was performed. The dynamics, stability and other features of the ligand-protein complexes were evaluated by using CPPTRAJ and PTRAJ [38].

2.7. The binding free energy calculations

For all protein-ligand complexes, the free binding energy was calculated using the script MMPBSA.PY by considering 500 snapshots using equations described below [39–42]. Different studies have used this free energy calculation method to estimate the binding of ligands for a given biological macromolecule [43,44].

$$\Delta G_{bind} = \Delta G_{complex} - \left[\Delta G_{receptor} + \Delta G_{ligand}\right]$$

Here, ΔG_{bind} denotes total free binding energy, while others denote the free energy of the protein, the ligand and complex. The following equation was used to calculate specific energy term contribution to the total free energy:

$$G = G_{bond} + G_{ele} + G_{vdW} + G_{pol} + G_{npol}$$

Bonded, electrostatic, polar, non-polar and van der Waal energy terms are represented by the above equation.

3. Results and discussion

3.1. Computer VS and binding modes of top hits

The outbreak of SARS-CoV-2 has distressed the world by spreading exponentially to every nook and cranny of the world. The pathophysiology of this virus shows that it primarily affects the respiratory tract by entering the host cell through Spike protein. The attachment of the Spike receptor-binding domain to the host ACE2 protein enables the entry to the host cell, and it is considered the main drug target for novel drug discovery [45,46]. However, recent findings deemed NRP1 as a supplementary cellular intermediary that enables SARS-CoV-2 attachment and entrance into host cells [7,8,11]. Therefore, it could serve as an alternative therapeutic target. Keeping in view the importance of NRP1 in the entry of SARS-CoV-2 and infection herein, we applied a computational VS approach to identify potential inhibitors from the marine CMNP database using AutoDock Vina. The database has 47,451 molecules collected from different marine sources. Our analysis identified that among the 47,451 compounds, only 21,000 compounds passed the ADMET criteria. Using AutoDock Vina for virtual screening reported that for these 21,000 compounds, the scores ranged from −8.0 to −2.3 kcal/mol. Compounds with a score lower than −6.0 kcal/mol were selected for further analysis. This was done to select highly efficient binders as a very small number of compounds from the entire library were found to have stable binding energy with the protein because of weak interaction profile. Therefore, to increase selection chances of efficient binding molecules to the protein, a stringent criteria of −6.0 kcal/mol to −8.0 kcal/mol was used. With this criterion, only those compounds were selected which interact with the key hot spot residues of the receptor protein. Based on the high docking score and interactions with the key residues, ten compounds were shortlisted for further analysis. The identified top ten compounds are given in Table 2. These compounds were further analyzed for drug-like rules, which include Lipinski, Ghose, Veber, Egan and Muegge parameters for drug feature calculation [47]. The shortlisted compounds include CMNPD10175, CMNPD10017, CMNPD10114, CMNPD10115, CMNPD10020. CMNPD10018, CMNPD10153, CMNPD10149 CMNPD10464 and CMNPD10019. The shortlisted compounds were reported to act drug-like with no toxicity and good pharmacokinetic properties. From the top ten compounds, only five compounds, presented in Fig. 1, were selected for interactions, molecular dynamics simulation and free energy calculations. The following sections describe the binding mode of each of these compounds in more detail.

3.2. Binding mode of CMNPD10175

Madangolide (CMNPD10175) is a marine natural compound originally isolated from cyanobacteria known *Lyngbya bouillonii,* which entails one large macrocyclic lactone ring [48]. It belongs to a group of antibiotics known as macrolides. Herein, Madangolide has been identified as the most potent drug with a docking score of −8.0 kcal/mol. This compound formed two hydrogens and two pie-alkyl bonds with the key active site residues. Among the two hydrogen bonds, one was formed with Phe297 while the other one was reported with Lys351. The residue Phe297 was also involved in pie-alkyl along with Tyr353. The interaction pattern of Madangolide (CMNPD10175) is given in Fig. 2a.

3.3. Binding mode of CMNPD10017

Isogranulatimide (CMNPD10017), marine pyrrolocarbazole, is a cell-permeable alkaloid primarily used for phosphorylation and dephosphorylation and has been reported to act as competitive inhibitors in cancer [49]. Initially isolated from a marine source, *Didemnum conchyliatum*, this compound act as a major G2 cell cycle checkpoint inhibitor. This compound identified herein through virtual drug screening possesses a docking score of −6.9 kcal/mol against the NRP1. Similar to the Madangolide, this compound formed 4 key interactions with the NRP1 receptor. Among the total interactions two hydrogen bonds with Asp320 and Thr413 while the two pie-alkyl interactions were formed with Tyr297, which is considered as an important residue for the recognition. The interaction pattern of Isogranulatimide (CMNPD10017) is given in Fig. 2b.

3.4. Binding mode of CMNPD10115

Likewise, the CHMPDB10115, Luisol B (CMNPD10115) also belongs to Naphthofurans class of marine natural products produced by a marine bacterium from the *Streptomyces (Actinomycetales)* genus [50,51]. This compound identified herein through virtual drug screening possesses a docking score of −6.9 kcal/mol against the NRP1. Unlike the Madangolide and Isogranulatimide, this molecule formed three hydrogen bonds and two pie-alkyl interactions with the key residues. Among the hydrogen bonds Tyr297, Tyr353 and Thr413 are involved, while the two pie-alkyl interactions are formed with Tyr297 and Trp301. This compound shows robust interactions with the key active site residues. The interaction pattern of Luisol B (CMNPD10114) is given in Fig. 3a.

3.5. Binding mode of CMNPD10114

Luisol A (CMNPD10114) belongs to Naphthofurans class of marine natural products produced by a marine bacterium from the *Streptomyces (Actinomycetales)* genus [50,51]. This compound secured a docking score of −6.8 kcal/mol against the NRP1. This molecule formed two hydrogen bonds and two pie-alkyl interactions with the key residues. Residues Trp301 and Glu348 are involved in hydrogen bonding while the two pie-alkyl interactions are formed with Tyr297 and Trp301. The interaction pattern of Luisol A (CMNPD10115) is given in Fig. 3b.

Table 2
Top ten marine drug compounds shortlisted through VS and applied drug-like filters. The docking score of each compound is given in kcal/mol.

Compound ID	Compound Names	drug like	Lead likeness	Toxicity	Pharmacokinetics	Solubility	Score
CMNPD10175	Madangolide	Yes	Yes	No	Good	Good	−8.0
CMNPD10017	Isogranulatimide	Yes	Yes	No	Good	Good	−6.9
CMNPD10114	Luisol B	Yes	Yes	No	Good	Good	−6.9
CMNPD10115	Luisol A	Yes	Yes	No	Good	Good	−6.8
CMNPD10020	Meridianin A	Yes	Yes	No	Good	Good	−6.6

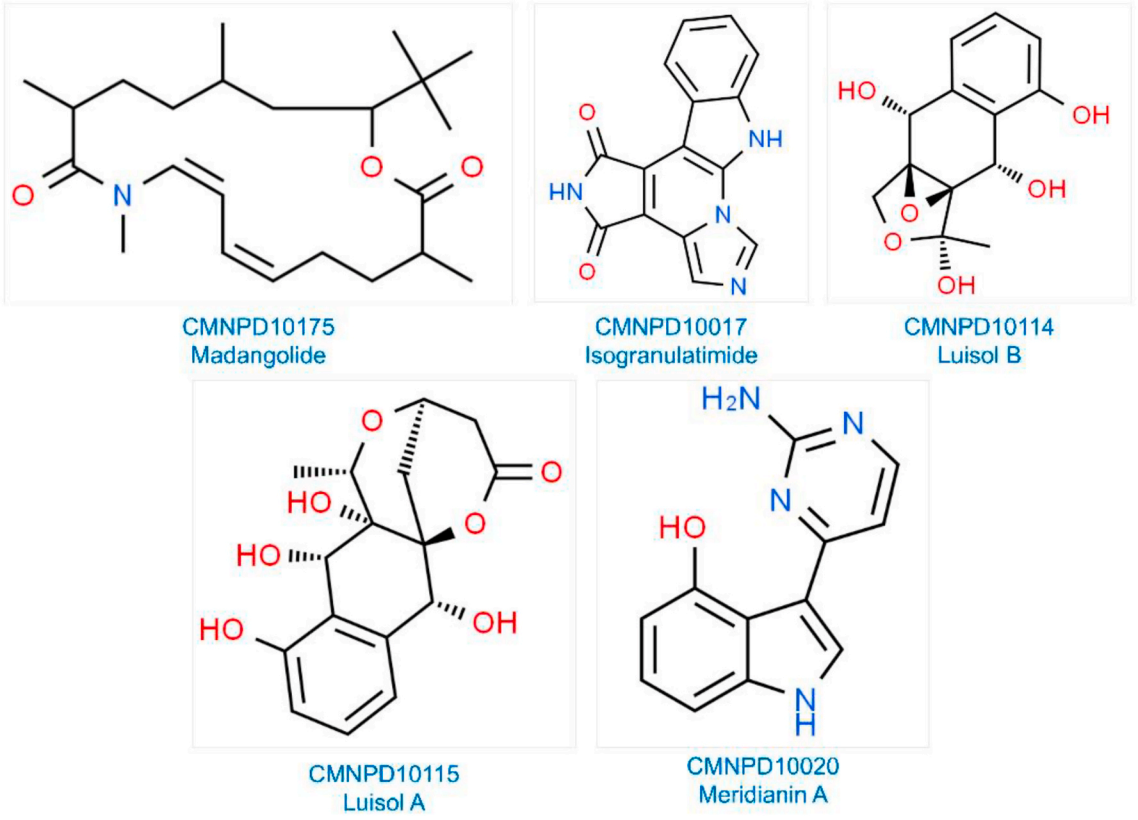

Fig. 1. 2D structure representation of the top five hits along with their CMNPD. database accession ID and original scientific names are given. The structural scaffold of all the compounds seems very similar, which reflects the same activity of these compounds.

3.6. Binding mode of CMNPD10020

Meridianin A (CMNPD10020) is a marine-derived hydroxyindoles primarily identified as kinase inhibitor [52]. Meridianin A identified as a potential drug with a docking score −6.6 kcal/mol. This compound formed two hydrogens and five pie-alkyl bonds with the key active site residues. Among the two hydrogen bonds were formed with Trp413 while the five pie-alkyl interactions involved Tyr297, Trp301 and Tyr353 residues. The interaction pattern of Meridianin A (CMNPD10020) is given in Fig. 4.

For instance, a comparative analysis of these top five hits with the already experimentally reported compounds revealed that the shortlisted compounds possess better docking scores than the experimentally reported. The experimentally reported best compound EG00229 has docking score of −6.4 kcal/mol, while the others have docking energy as follow: EG01377 (docking score = −5.8 kcal/mol) [53], CB7739526 (docking score = −4.1 kcal/mol) [54] and NRPa308 (docking score = −3.5 kcal/mol) [55]. In contrast, the docking scores of our shortlisted compounds are far better than the experimentally reported, which confirms the stronger inhibitory effects of our compounds. Similarly, a recent study [56] also reported some compounds including Amantadine, Carvacrol, Thymoquinone, and Thymol identified through VS against the NRP1 also possess lower docking scores than our reported top hits. This shows the robustness of our top hits, which bind more strongly than the previously reported.

3.7. Prediction of bioactivity and K_D Determination

On the other hand, molinspiration predicted the bioactivity of each compound against each class of receptors. The server's results demonstrate that all of the compounds on the shortlist are active against the nuclear receptor target (**Table 3**). The reported scores for CMNPD10175 (0.34), CMNPD10017 (0.64), CMNPD10114 (0.50), CMNPD10115 (0.46) while CMNPD10020 possess (0.04) bioactivity score against the nuclear receptor target. The scores for other classes such as GPCR, Ion Channels, Kinases, protease and enzymes are also given in Table 3. As a result, these findings conclude that the compounds shortlisted could effectively inhibit NRP1 in the experimental design and could be evaluated in clinical trials.

We also determined the K_D of all the complexes to provide a clear insight into the binding variations. K_D is a method for assessing and ranking the strength of biomolecular interactions [57]. This method is widely used for the binding strength determination of biological molecules [58,59]. We employed PRODIGY, to calculate the binding affinity

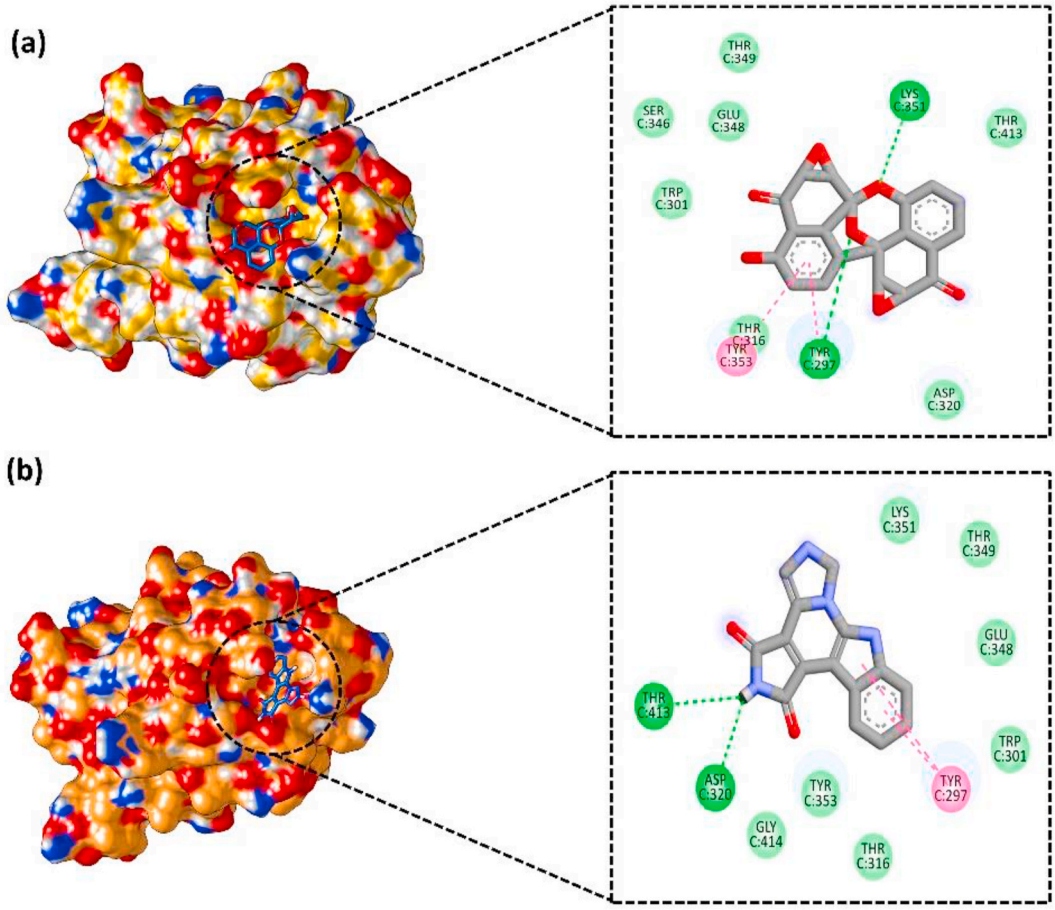

Fig. 2. Binding modes of Madangolide (CMNPD10175) **(a)** and Isogranulatimide (CMNPD10017) **(b)**. The hydrogen-bonding interactions are shown in green colour, while the pie-alkyl interactions are shown as pink.

K_D for these protein-drugs complexes. Prodigy reported the binding score for each drug as −7.9 for CMNPD10175, -7.3 for CMNPD1017, CMNPD10114 and CMNPD10115, respectively. While for CMNPD10020, the prodigy score was predicted to be −7.1 kcal/mol. This shows the tighter binding of these molecules with the host receptor NRP1 and its inhibition. The predicted scores by prodigy are given in Table 3.

3.8. Deciphering structural stability via RMSD

Predictions made by molecular docking about enzyme-ligand docked conformation stability were tested in a dynamic environment over 200 ns of MD simulation time. To understand the generated MD simulation trajectories, CPPTRAJ. module of AMBER was employed to carry out several statistical parameters that helped to understand overall complexes structural stability. In this regard, the first root mean square deviation (RMSD) was performed that plotted all Carbon alpha atoms deviation from original docked conformation versus time. Higher RMSD. implies higher structural fluctuations with respect to a reference, whereas lower RMSD. corresponds to good structure stability, i.e. higher intermolecular affinity and better-docked mode. The mean RMSD of CMNPD10017, CMNPD10020, CMNPD10114, CMNPD10115, and CMNPD10149 was reported to be 1.2 Å, 1.4 Å, 1.3 Å, 1.3 Å and 1.2 Å, respectively. Overall, all the complexes indicated good structural stability as no significant spike in the RMSD was highlighted. The CMNPD10149-enzyme complex remained quite stable till 50 ns with RMSD. ~1 Å. The followed RMSD suffers from an initial minor deviation touching RMSD of 2 Å. Then, the RMSD plot is more uniform, with no global or local structural changes noticed until 170 ns. Afterwards, a sudden rise and fall of RMSD till 175 ns was depicted. Towards the end, the RMSD of this system was seen as the highly stable indicating equilibrium of the intermolecular interactions and overall system. The CMNPD10017-enzyme complex is subject to a continuous pattern of RMSD with a minor steady increase till 165 ns. The highest RMSD seen for this system is 2.5 Å. Towards the end time, system RMSD is drop-down touched 1.5 Å and suffered from a sudden upsurge. This concluded the system RMSD not converged and need more MD simulation time. The CMNPD10114 system is more consistent with no structural alterations and a mean RMSD of 1.3 Å to 1.4 Å. CMNPD10115 and CMNPD10020 complexes like CMNPD10114 behave very stably except for one sharp spike at 125 ns in the case of CMNPD10115. The last three systems were reported to show the greater intermolecular affinity of the compounds for the enzyme that further suggested that proposed docked complexes by docking studies are valid in terms of lowest energy minima and conformational stability. The RMSDs of all the complexes are given in Fig. 5.

3.9. Investigating residues fluctuations

Investigating the residues fluctuation, we employed root-mean-square fluctuation (RMSF) to calculate the residual flexibility of carbon alpha residues of the targeted systems. Generally speaking, all the systems mean RMSF is very low, indicating the residues of enzyme enjoying excellent energy minima in the presence of compounds. This can be indeed affirmed by the very low RMSF of enzyme binding site residues which is well below under 2 Å. However, the C- and N-terminal of the enzyme were found highly unstable, which might be flexible and is a mechanism to attain a more stable conformation of the compounds

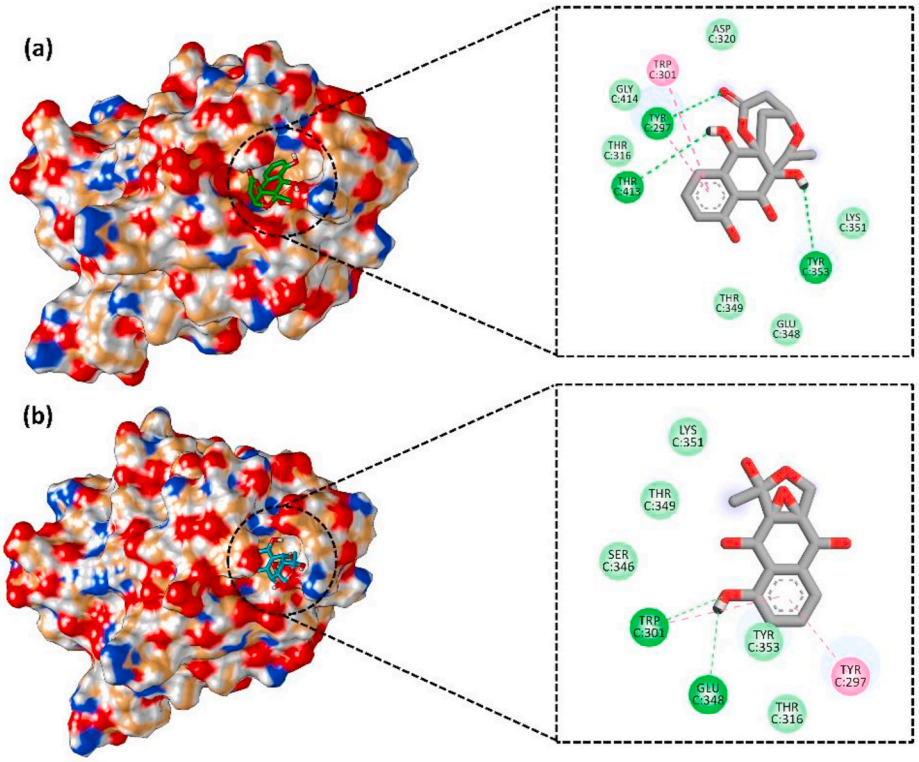

Fig. 3. Binding modes of Luisol B (CMNPD10114) **(a)** and Luisol A (CMNPD10115) **(b)**. The hydrogen bonding interactions are shown in green colour while the pie-alkyl interactions are shown as pink.

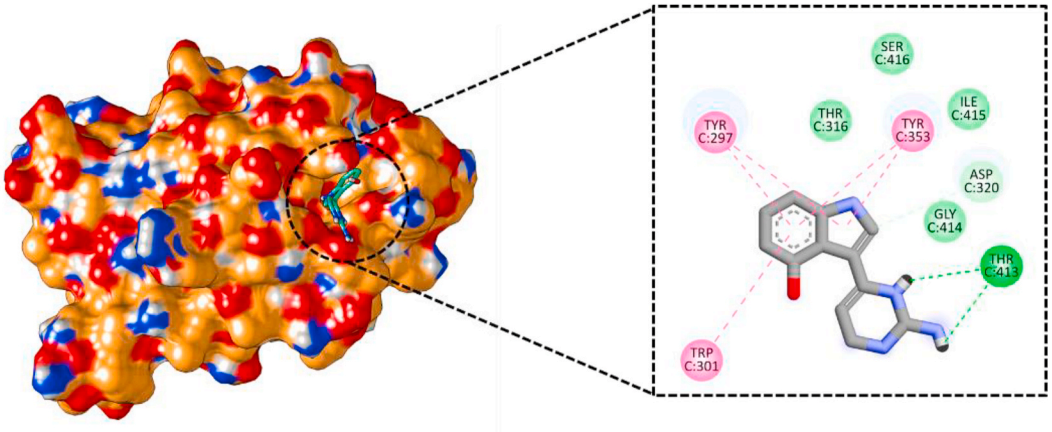

Fig. 4. Binding modes of Meridianin A (CMNPD10020). The hydrogen-bonding interactions are shown in green colour, while the pie-alkyl interactions are shown as pink.

Table 3
Top five marine drug compounds shortlisted through VS and drug-like rules. The docking score of each compound is given in kcal/mol.

Compound ID	GPCR ligand	Ion Channels	Kinase Inhibitors	Nuclear Receptor ligand	Protease inhibitor	Enzymes inhibitor	Prodigy score
CMNPD10175	0.18	−0.06	−0.42	0.34	0.28	0.21	−7.9
CMNPD10017	−0.07	−0.10	1.29	0.64	−0.20	0.41	−7.3
CMNPD10114	0.25	0.17	−0.10	0.50	0.21	0.51	−7.3
CMNPD10115	−0.04	0.01	−0.20	0.46	−0.11	0.34	−7.3
CMNPD10020	0.20	0.34	1.0	0.04	−0.46	0.60	−7.1

at the active pocket. RMSF data complement RMSD as both agree on the highly stable nature of complexes. The RMSFs of all the complexes are given in Fig. 6.

3.10. Determining strength of intermolecular interactions

In biological systems, hydrogen bonding is key to strong intermolecular binding and therefore easing molecular recognition and

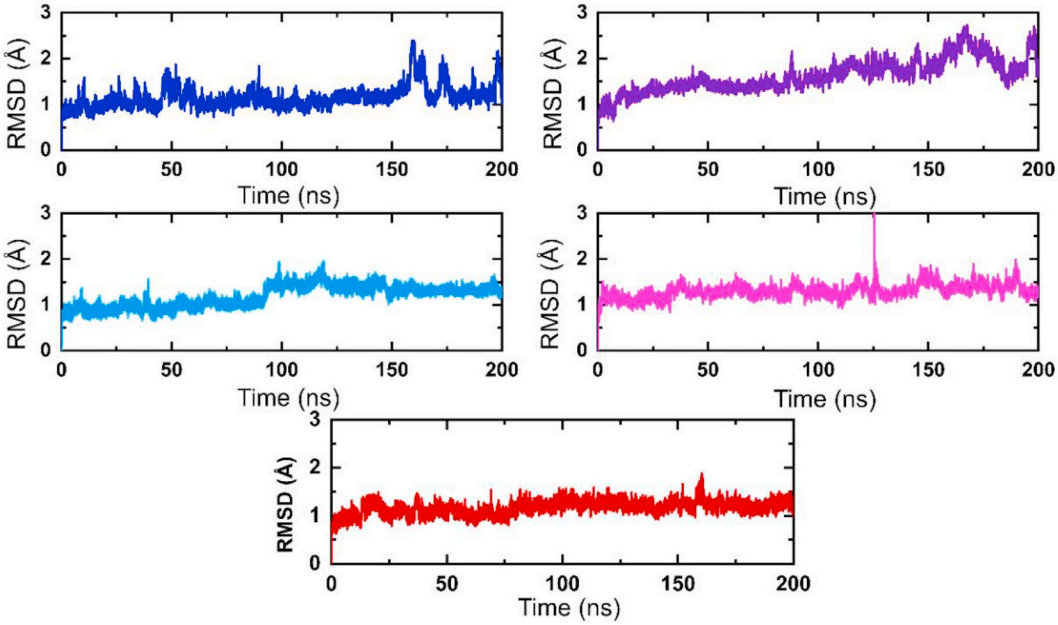

Fig. 5. Thermodynamics stability calculated as RMSD of all the complexes. **(a)** CMNPD10175, **(b)** CMNPD1017, **(c)** CMNPD10114, **(d)** CMNPD10115 and **(e)** CMNPD10020 represent each complex.

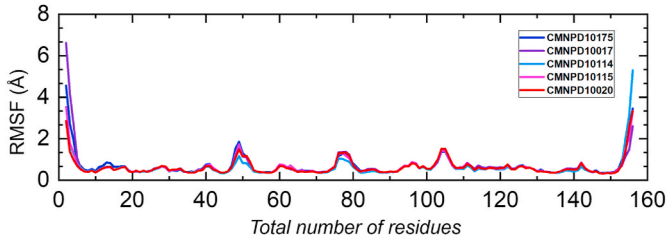

Fig. 6. Residual flexibility calculated as RMSF of all the complexes. All the complexes CMNPD10175, CMNPD1017, CMNPD10114, CMNPD10115 and CMNPD10020 represented with different colour.

ultimately performing biological functions. All complexes were subjected to hydrogen bond analysis to calculate the number of hydrogen bonds formed in each frame of MD simulation. As can be visualized in the figure, all the targeted complexes are dominated by a rich network of hydrogen bonds reflecting on the strong stability of the docked compounds with the enzyme. The average number of hydrogen bonds in each frame of the complexes is 85, 83, 83, 82 and 82, respectively. As observed in the above RMSD and RMSF, the systems are showing significant stability, which is the outcome of a large number of hydrogen bonding. The *H-bonds* count of all the complexes is given in Fig. 7.

3.11. Unveiling atomic level interaction energies

Determining binding free energies at the residue level of protein-ligand complex is now in the routine of computer-aided drug designing (CAAD) to cross-validate docking and MD simulation results. This was achieved by performing MM-GBSA analysis on the simulation trajectories by picking all the frames. Binding free energies estimation by MM-GBSA is a common practice as they are more reliable than traditional docking scores and less computationally expensive compare

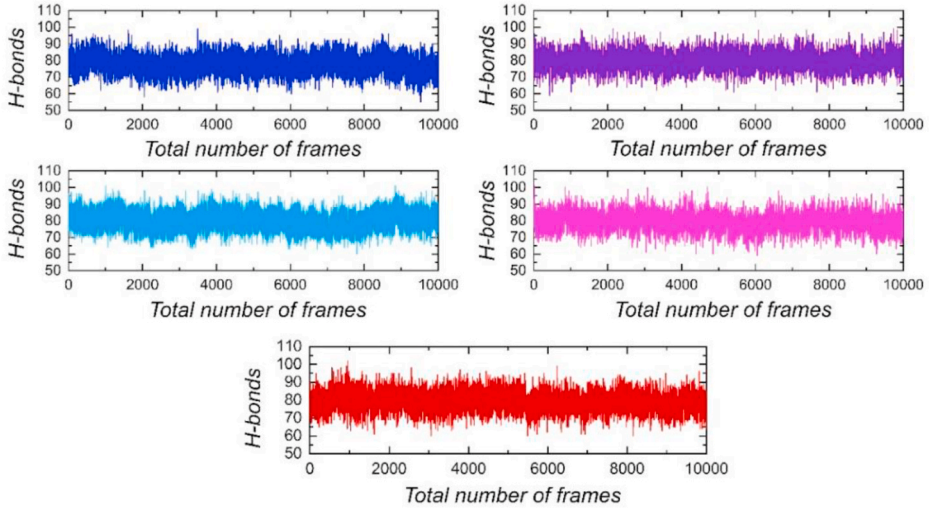

Fig. 7. Thermodynamics stability calculated as RMSD of all the complexes. **(a)** CMNPD10175, **(b)** CMNPD1017, **(c)** CMNPD10114, **(d)** CMNPD10115 and **(e)** CMNPD10020 represent each complex.

to Alchemical free energy methods. As can be understood from the data tabulated in Table 4, the CMNPD10017 is relatively more stable by securing the lowest binding energy of −43.08 kcal/mol, followed by CMNPD10020 (−40.86 kcal/mol). This can be reasoned by the structural deviations the complexes acquired during simulation time; these deviations might help these two complexes to get more stable conformation by establishing new chemical interactions. The last three complexes, CMNPD10114, CMNPD10115 and CMNPD10149, are classified as less stable than the mentioned before that might be because of the relaxed complex structure. The complexes preferred to deviate from the original conformation as this might lose their significant bonding and will push the systems to instability. CMNPD10017, CMNPD10020 and CMNPD10149 interactions with the enzyme are dominated by van der Waals energy as well as a favourable contribution from electrostatic energy was also unraveled. In the case of CMNPD10114 and CMNPD10115, the systems favoured electrostatic energy to produced stable complex while less significant help was considered from van der Waals energy. In all complexes, the non-polar energy of the solvation phase seems also to contribute positively, making the systems stable. The polar energy is a non-contributor to system stabilization as can be interpreted by positive values. Together, docking, MD simulations and binding free energies all are in support of good docked conformation of the compounds to the enzyme and resulting in highly stable complexes.

4. Conclusions

In conclusion, this study identified novel hits from the marine sources against SARS-CoV-2 NRP1 receptor. Though the findings of the current work are promising, the study suffers from several limitations of the methods applied herein. Docking limitations are overcome by the long run of MD simulations and downward sophisticated computational analysis. We further determined the bioactivity and K_D of the top hits. CMNPD10175, CMNPD10017, CMNPD10114, CMNPD10115, CMNPD10020. CMNPD10018, CMNPD10153, CMNPD10149 CMNPD10464 and CMNPD10019 which were identified through VS and further explored by analyzing its ADMET properties and structural-dynamics features. Free energy calculations further established that all the compounds exhibit strong binding energy for NRP1. Though the findings of the current work are promising, the study suffers from several limitations of the methods applied herein. For example, the docking predictions are often found non-reliable. The force fields used in MD simulations need further refinement whereas the binding free energy methods lacked entropy estimation as well as missed the role of water molecules that bridge ligands with receptor hotspot residues. Consequently, we hypothesized that these compounds might be the best lead candidates for therapeutic interventions. This study provides a strong impetus to develop novel drugs against the SARS-CoV-2 by targeting NRP1.

Author contributions

"Conceptualization, AK and A.M.; methodology, A. K; software, DW; validation, DW, FA and JA; formal analysis, AK DW SA; investigation, KK, SS.A. JA, NU, AS; resources, DW; data curation, AK, AS; writing—original draft preparation, AK, KK, SS.A. JA, NU, AS; writing—review and editing, A.M, DW; visualization, AKFA, JA; supervision, DW; project administration, DW; funding acquisition, DW All authors have read and agreed to the published version of the manuscript."

Funding

Dong-Qing Wei received funding from Key Research Area Grant 2016YFA0501703 of the "Ministry of Science and Technology of China", "the National Science Foundation of China". The grants Numbers are 32070662, 61832019, 32030063. This work is also supported by "Science and Technology Commission of Shanghai Municipality" under the grant number 19430750600 and Joint Research Fund and Joint Research Funds for Medical and Engineering and Scientific Research from Shanghai Jiao Tong University grant number YG2021ZD02.

Table 4
Binding free energy results of all the complexes calculated using the MM-GBSA approach. The free score of each compound is given against it in kcal/mol.

Complex	vdW	Electrostatic	GB	ESURF	Total
CMNPD10017	−30.97	−25.60	16.78	−3.29	**−43.08**
CMNPD10020	−27.00	−18.17	7.08	−2.59	**−40.86**
CMNPD10114	−24.65	−93.53	89.10	−3.30	**−32.38**
CMNPD10115	−20.76	−140.92	131.88	−2.99	**−32.80**
CMNPD10149	−27.93	−11.65	4.83	−3.00	**−34.45**

Data availability statement

All the data is available on RCSB, UniProt and any simulation data would be provided on demand. The accession numbers to access this data are given in the manuscript.

Declaration of competing interest

Declared None.

Acknowledgments

The computational resources were provided by the Pengcheng Lab. HPC center (High-Performance Computing) of Shanghai Jiao Tong University. We acknowledge their help.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2021.104714.

References

[1] D. Benvenuto, M. Giovanetti, A. Ciccozzi, S. Spoto, S. Angeletti, M. Ciccozzi, The 2019-new coronavirus epidemic: evidence for virus evolution, J. Med. Virol. 92 (2020) 455–459.
[2] W.-j. Guan, Z.-y. Ni, Y. Hu, W.-h. Liang, C.-q. Ou, J.-x. He, L. Liu, H. Shan, C.-l. Lei, D.S. Hui, Clinical Characteristics of 2019 Novel Coronavirus Infection in China, MedRxiv, 2020.
[3] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, The lancet 395 (2020) 497–506.
[4] D. Wrapp, N. Wang, K.S. Corbett, J.A. Goldsmith, C.-L. Hsieh, O. Abiona, B. S. Graham, J.S. McLellan, Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation, Science 367 (2020) 1260–1263.
[5] T.N. Starr, A.J. Greaney, S.K. Hilton, D. Ellis, K.H. Crawford, A.S. Dingens, M. J. Navarro, J.E. Bowen, M.A. Tortorici, A.C. Walls, Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ACE2 binding, Cell 182 (2020) 1295–1310, e1220.
[6] I. Kyrou, H.S. Randeva, D.A. Spandidos, E. Karteris, Not only ACE2—the quest for additional host cell mediators of SARS-CoV-2 infection: neuropilin-1 (NRP1) as a novel SARS-CoV-2 host cell entry mediator implicated in COVID-19, Signal transduction and targeted therapy 6 (2021) 1–3.
[7] J.L. Daly, B. Simonetti, K. Klein, K.-E. Chen, M.K. Williamson, C. Antón-Plágaro, D. K. Shoemark, L. Simón-Gracia, M. Bauer, R. Hollandi, Neuropilin-1 is a host factor for SARS-CoV-2 infection, Science 370 (2020) 861–865.
[8] L. Cantuti-Castelvetri, R. Ojha, L.D. Pedro, M. Djannatian, J. Franz, S. Kuivanen, F. van der Meer, K. Kallio, T. Kaya, M. Anastasina, Neuropilin-1 facilitates SARS-CoV-2 cell entry and infectivity, Science 370 (2020) 856–860.
[9] B.S. Mayi, J.A. Leibowitz, A.T. Woods, K.A. Ammon, A.E. Liu, A. Raja, The role of Neuropilin-1 in COVID-19, PLOS, Pathogens 17 (2021), e1009153.
[10] A. Moutal, L.F. Martin, L. Boinon, K. Gomez, D. Ran, Y. Zhou, H.J. Stratton, S. Cai, S. Luo, K.B. Gonzalez, SARS-CoV-2 Spike protein co-opts VEGF-A/Neuropilin-1 receptor signaling to induce analgesia, Pain 162 (2021) 243.
[11] L. Cantuti-Castelvetri, R. Ojha, L.D. Pedro, M. Djannatian, J. Franz, S. Kuivanen, K. Kallio, T. Kaya, M. Anastasina, T. Smura, Neuropilin-1 Facilitates SARS-CoV-2 Cell Entry and Provides a Possible Pathway into the Central Nervous System, BioRxiv, 2020.

[12] M. Seyran, K. Takayama, V.N. Uversky, K. Lundstrom, G. Palù, S.P. Sherchan, D. Attrish, N. Rezaei, A.A. Aljabali, S. Ghosh, The structural basis of accelerated host cell entry by SARS-CoV-2, FEBS J. (2020).

[13] T. Teesalu, K.N. Sugahara, V.R. Kotamraju, E. Ruoslahti, C-end rule peptides mediate neuropilin-1-dependent cell, vascular, and tissue penetration, Proc. Natl. Acad. Sci. Unit. States Am. 106 (2009) 16157–16162.

[14] H.-F. Guo, C.W. Vander Kooi, Neuropilin functions as an essential cell surface receptor, J. Biol. Chem. 290 (2015) 29120–29126.

[15] C.A. Devaux, J.-C. Lagier, D. Raoult, New insights into the physiopathology of COVID-19: SARS-CoV-2-associated gastrointestinal illness, Front. Med. 8 (2021) 99.

[16] D. Wang, J. Mai, W. Zhou, W. Yu, Y. Zhan, N. Wang, N.D. Epstein, Y. Yang, Immunoinformatic analysis of T- and B-cell epitopes for SARS-CoV-2 vaccine design, Vaccines 8 (2020) 355.

[17] S. Ismail, S. Ahmad, S.S. Azam, Immunoinformatics characterization of SARS-CoV-2 spike glycoprotein for prioritization of epitope based multivalent peptide vaccine, J. Mol. Liq. 314 (2020) 113612.

[18] M. Bhattacharya, A.R. Sharma, P. Patra, P. Ghosh, G. Sharma, B.C. Patra, S.-S. Lee, C. Chakraborty, Development of epitope-based peptide vaccine against novel coronavirus 2019 (SARS-COV-2): immunoinformatics approach, J. Med. Virol. 92 (2020) 618–631.

[19] Y. Dai, H. Chen, S. Zhuang, X. Feng, Y. Fang, H. Tang, R. Dai, L. Tang, J. Liu, T. Ma, G. Zhong, Immunodominant regions prediction of nucleocapsid protein for SARS-CoV-2 early diagnosis: a bioinformatics and immunoinformatics study, Pathog. Glob. Health 114 (2020) 463–470.

[20] R. Saha, P. Ghosh, VLSP Burra, Designing a next generation multi-epitope based peptide vaccine candidate against SARS-CoV-2 using computational approaches, 3, Biotech 11 (2021) 47.

[21] A. Khan, S.S. Ali, M.T. Khan, S. Saleem, A. Ali, M. Suleman, Z. Babar, A. Shafiq, M. Khan, D.-Q. Wei, Combined drug repurposing and virtual screening strategies with molecular dynamics simulation identified potent inhibitors for SARS-CoV-2 main protease (3CLpro), J. Biomol. Struct. Dyn. (2020) 1–12.

[22] A. Khan, T. Zia, M. Suleman, T. Khan, S.S. Ali, A.A. Abbasi, A. Mohammad, D.-Q. Wei, Higher infectivity of the SARS-CoV-2 new variants is associated with K417N/T, E484K, and N501Y mutants: an insight from structural data, J. Cell. Physiol. (2021) n/a.

[23] A. Khan, W. Heng, Y. Wang, J. Qiu, X. Wei, S. Peng, S. Saleem, M. Khan, S.S. Ali, D.-Q. Wei, In Silico and in Vitro Evaluation of Kaempferol as a Potential Inhibitor of the SARS-CoV-2 Main Protease (3CLpro), Phytotherapy research: PTR.

[24] A. Khan, M. Khan, S. Saleem, Z. Babar, A. Ali, A.A. Khan, Z. Sardar, F. Hamayun, S. S. Ali, D.-Q. Wei, Phylogenetic analysis and structural perspectives of RNA-dependent RNA-polymerase inhibition from SARs-CoV-2 with natural products, Interdisciplinary Sciences (2020) 1–14. Computational Life Sciences.

[25] A. Khan, D.-Q. Wei, K. Kousar, J. Abubaker, S. Ahmad, J. Ali, F. Al-Mulla, S.S. Ali, N. Nizam-Uddin, A.M. Sayaf, Preliminary Structural Data Revealed that the SARS-CoV-2 B. 1.617 Variant's RBD Binds to ACE2 Receptor Stronger than the Wild Type to Enhance the Infectivity, ChemBioChem.

[26] D. Lagorce, L. Bouslama, J. Becot, M.A. Miteva, B.O. Villoutreix, FAF-Drugs4: free ADME-tox filtering computations for chemical biology and early stages drug discovery, Bioinformatics 33 (2017) 3658–3660.

[27] P.W. Rose, A. Prlić, A. Altunkaya, C. Bi, A.R. Bradley, C.H. Christie, L.D. Costanzo, J.M. Duarte, S. Dutta, Z. Feng, The RCSB Protein Data Bank: Integrative View of Protein, Gene and 3D Structural Information, Nucleic acids research, 2016, p. gkw1000.

[28] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera—a visualization system for exploratory research and analysis, J. Comput. Chem. 25 (2004) 1605–1612.

[29] T.D. Goddard, C.C. Huang, T.E. Ferrin, Software extensions to UCSF chimera for interactive visualization of large molecular assemblies, Structure 13 (2005) 473–482.

[30] S. Dallakyan, A.J. Olson, Small-molecule Library Screening by Docking with PyRx, Chemical Biology, Springer, 2015, pp. 243–250.

[31] O. Trott, A.J. Olson, AutoDock Vina, Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, J. Comput. Chem. 31 (2010) 455–461.

[32] C. Lyu, T. Chen, B. Qiang, N. Liu, H. Wang, L. Zhang, Z. Liu, CMNPD: a comprehensive marine natural products database towards facilitating drug discovery from the ocean, Nucleic Acids Res. 49 (2021) D509–D515.

[33] P.A. Ravindranath, S. Forli, D.S. Goodsell, A.J. Olson, M.F. Sanner, AutoDockFR: advances in protein-ligand docking with explicitly specified binding site flexibility, PLoS Comput. Biol. 11 (2015), e1004586.

[34] A. Daina, O. Michielin, V. Zoete, SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules, Sci. Rep. 7 (2017) 1–13.

[35] L.C. Xue, J.P. Rodrigues, P.L. Kastritis, A.M. Bonvin, A. Vangone, PRODIGY: a web server for predicting the binding affinity of protein–protein complexes, Bioinformatics 32 (2016) 3676–3678.

[36] D.A. Case, T.E. Cheatham III, T. Darden, H. Gohlke, R. Luo, K.M. Merz Jr., A. Onufriev, C. Simmerling, B. Wang, R.J. Woods, The Amber biomolecular simulation programs, J. Comput. Chem. 26 (2005) 1668–1688.

[37] D.J. Price, C.L. Brooks III, A modified TIP3P water potential for simulation with Ewald summation, J. Chem. Phys. 121 (2004) 10096–10103.

[38] D.R. Roe, T.E. Cheatham III, PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data, J. Chem. Theor. Comput. 9 (2013) 3084–3095.

[39] H. Sun, Y. Li, S. Tian, L. Xu, T. Hou, Assessing the performance of MM/PBSA and MM/GBSA methods. 4. Accuracies of MM/PBSA and MM/GBSA methodologies evaluated by various simulation protocols using PDBbind data set, Phys. Chem. Chem. Phys. 16 (2014) 16719–16729.

[40] T. Hou, N. Li, Y. Li, W. Wang, Characterization of domain–peptide interaction interface: prediction of SH3 domain-mediated protein–protein interaction network in yeast by generic structure-based models, J. Proteome Res. 11 (2012) 2982–2995.

[41] F. Chen, H. Liu, H. Sun, P. Pan, Y. Li, D. Li, T. Hou, Assessing the performance of the MM/PBSA and MM/GBSA methods. 6. Capability to predict protein–protein binding free energies and re-rank binding poses generated by protein–protein docking, Phys. Chem. Chem. Phys. 18 (2016) 22129–22139.

[42] B.R. Miller III, T.D. McGee Jr., J.M. Swails, N. Homeyer, H. Gohlke, A.E. Roitberg, MMPBSA py: an efficient program for end-state free energy calculations, J. Chem. Theor. Comput. 8 (2012) 3314–3321.

[43] Y. Wang, A. Khan, A. Chandra Kaushik, M. Junaid, X. Zhang, D.-Q. Wei, The systematic modeling studies and free energy calculations of the phenazine compounds as anti-tuberculosis agents, J. Biomol. Struct. Dyn. 37 (2019) 4051–4069.

[44] A. Khan, A.C. Kaushik, S.S. Ali, N. Ahmad, D.-Q. Wei, Deep-learning-based target screening and similarity search for the predicted inhibitors of the pathways in Parkinson's disease, RSC Adv. 9 (2019) 10326–10339.

[45] M. Letko, A. Marzi, V. Munster, Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses, Nature microbiology 5 (2020) 562–569.

[46] J. Lan, J. Ge, J. Yu, S. Shan, H. Zhou, S. Fan, Q. Zhang, X. Shi, Q. Wang, L. Zhang, Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor, Nature 581 (2020) 215–220.

[47] B. Kirchweger, J.M. Rollinger, Virtual Screening for the Discovery of Active Principles from Natural Products, Natural Products as Source of Molecules with Therapeutic Potential, Springer, 2018, pp. 333–364.

[48] D. Klein, J.C. Braekman, D. Daloze, L. Hoffmann, G. Castillo, V. Demoulin, Madangolide, A. Laingolide, Two novel macrolides from Lyngbya bouillonii (cyanobacteria), J. Nat. Prod. 62 (1999) 934–936.

[49] B. Hugon, F. Anizon, C. Bailly, R.M. Golsteyn, A. Pierré, S. Léonce, J. Hickman, B. Pfeiffer, M. Prudhomme, Synthesis and biological activities of isogranulatimide analogues, Bioorg. Med. Chem. 15 (2007) 5965–5980.

[50] X.C. Cheng, P.R. Jensen, W. Fenical, Luisols A and B, new aromatic tetraols produced by an estuarine marine bacterium of the genus Streptomyces (Actinomycetales), J. Nat. Prod. 62 (1999) 608–610.

[51] W.N. Hozzein, M. Mohany, S.M. Alhawsawi, M.Y. Zaky, S.S. Al-Rejaie, D.H. M. Alkhalifah, Flavonoids from Marine-Derived Actinobacteria as Anticancer Drugs, Current Pharmaceutical Design, 2020.

[52] S. B Bharate, R. R Yadav, S. Battula, R. a Vishwakarma, Meridianins: marine-derived potent kinase inhibitors, Mini Rev. Med. Chem. 12 (2012) 618–631.

[53] A. Jarvis, C.K. Allerston, H. Jia, B. Herzog, A. Garza-Garcia, N. Winfield, K. Ellard, R. Aqil, R. Lynch, C. Chapman, Small molecule inhibitors of the neuropilin-1 vascular endothelial growth factor A (VEGF-A) interaction, J. Med. Chem. 53 (2010) 2215–2226.

[54] J. Powell, F. Mota, D. Steadman, C. Soudy, J.T. Miyauchi, S. Crosby, A. Jarvis, T. Reisinger, N. Winfield, G. Evans, Small molecule neuropilin-1 antagonists combine antiangiogenic and antitumor activity with immune modulation through reduction of transforming growth factor beta (TGFβ) production in regulatory T-cells, J. Med. Chem. 61 (2018) 4135–4154.

[55] A. Starzec, M.A. Miteva, P. Ladam, B.O. Villoutreix, G.Y. Perret, Discovery of novel inhibitors of vascular endothelial growth factor-A–Neuropilin-1 interaction by structure-based virtual screening, Bioorg. Med. Chem. 22 (2014) 4042–4048.

[56] s. mohamed, S. Mohamed, A. Aya, N.Z. Abdel Rahman, Research Square, 2021.

[57] J.P. Landry, Y. Fei, X. Zhu, Simultaneous measurement of 10,000 protein-ligand affinity constants using microarray-based kinetic constant assays, Assay Drug Dev. Technol. 10 (2012) 250–259.

[58] J.P. Landry, Y. Fei, X. Zhu, High throughput, label-free screening small molecule compound libraries for protein-ligands using combination of small molecule microarrays and a special ellipsometry-based optical scanner, International drug discovery (2011) 8.

[59] J. Landry, Y. Sun, X. Guo, X. Zhu, Protein reactions with surface-bound molecular targets detected by oblique-incidence reflectivity difference microscopes, Appl. Opt. 47 (2008) 3275–3288.