Application of docking-based inverse high throughput virtual screening to found phytochemical covalent inhibitors of SARS-CoV-2 main protease, NSP12 and NSP16

Yaroslav V. Faletrov
Research Institute for Physical Chemical Problems, Belarusian State University, Leningradskaya str. 14, Minsk, Belarus

Viktoryia A. Staravoitava
Research Institute for Physical Chemical Problems, Belarusian State University, Leningradskaya str. 14, Minsk, Belarus

Anastasiya R. Dudko
Research Institute for Physical Chemical Problems, Belarusian State University, Leningradskaya str. 14, Minsk, Belarus

Vladimir M. Shkumatov (vlad.shkumatov@tut.by)
Belarusian State University

Research Article

Keywords: SARS-CoV-2, virtual screening, docking, phytochemicals, unsaturated carbonyl, covalent inhibitor

Posted Date: March 21st, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1456627/v1

Abstract

Computer-aided docking-based inverse high-throughput virtual screening (inv HTVS) was applied for a quick evaluation of compounds as new affine ligands for SARS-CoV-2 proteins. We performed Autodock Vina-based inv HTVS using more than 450 structures of the virus proteins from Protein Data Bank (PDB) and totally ~ 2000 structures of compound with electrophilic motifs (mainly, α,β-unsaturated carbonyls) from Pubchem (CID codes) or Biogem databases. The inv HTVS calculations were semi-automatically organized, run and analyzed using our helper program tool composed of Python scripts and Excel files with macroses. Criteria for binding energy (Ebind, kcal/mol) values (no more than 6.9), distance between thiol groups of cysteine residues and electrophilic groups of the compounds within 0.4 nm (distance criterium) and total number of structures meeting the criteria were used. New in silico interactions were found for Mpro in the cases of such phytochemicals as peperomine E, 15-oxosteviol, zambesiacolactone B, isodehydroeupatundin, 2deoxycucurbitacin and jatrophone. Analogously, potential covalent ligands for NSP12 were found to include ixerin D, xindongnin C, rabdosin B, epinodosin as well as such ligands for NSP16 were found to include anhydroverlotolin-like compound CID325147, mexicanin E, insulicolide B and geoditin A. These and others hits are discussed with respect to their availability and biological properties. Thus, a set of natural compounds were pointed out as potential new covalent inhibitors of SARS-CoV-2 proteins Mpro, Nsp12 and Nsp16. This information could be used for further evaluations as prophylaxis tools or drugs against COVID-19.

Intoduction

Coronavirus disease 2019 (COVID-19) is a worldwide pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). SARS-CoV-2 is an enveloped, positive-sense single-stranded RNA (+ ssRNA) virus belonging to the family of betacoronaviruses. [1]. Its genome encodes the structural Spike (S), Envelope (E), Membrane (M), Nucleocapsid (N), and nonstructural proteins (nsp1-nsp16) which include papain-like proteases (PLpro, part of NSP3) and 3C-like protease (3CLpro, NSP5) [2][3].

On 30 January 2020, the WHO declared the COVID-19 outbreak as the sixth public health emergency of international concern, following H1N1 (2009), polio (2014), Ebola in West Africa (2014), Zika (2016) and Ebola in the Democratic Republic of Congo (2019). Presently, SARS-CoV-2 has infected over 300 million people, and over 5 million deaths have been reported [4]. Mutations in S protein on the surface of SARS-CoV-2 can increase either transmission or mortality rate like it was observed for the cases of omicron and delta strains and also decrease vaccines effectivity. These circumstances highlight necessity to develop a small molecule drug suitable for coronavirus infection management. Quite effective therapeutics which was repurposed for the treatment of COVID-19 have been identified by screening known antivirals, other approved drugs and clinical-trial drug candidates. Among of them RdRp inhibitors remdesivir, ribavirin, favipiravir, galidesivir, Mpro covalent inhibitors nirmatrevir (a component of the Paxlovid), ebselen, boceprevir, carmofur, Mpro noncovalent inhibitors cinanserin, atazanavir and novel ensitrelvir are currently available for use.

Virtual screening methods can significantly reduce the number of candidates for *in vitro* biological activity assay. [5][6][7] [8] [9]. Developing new effective anti-CoV-2 drugs may require several years of drug development efforts. Nevertheless, given the urgency of the ongoing COVID-19 pandemic, screening for substances found in plants and other living organisms, that are presumably to be safe for humans, can provide a quick and effective approach to combating this deadly infection. Examples of successful in silico screening of phytochemicals able to bind with SARS-CoV-2 proteins [10][11][12] has encouraged us to do efforts to find out more natural compounds with a potential as covalent ligands for the targets. Using all our experience in docking calculations [13][14], we performed such a search using AutoDock Vina, ~ 450 SARS-CoV-2 protein 3D structures and over 2000 structures, mainly α,β-unsaturated carbonyl-bearing lactones and ketones.

Methods

Hardware

Computations were done using mainly Hewlett Packard Z230 workstation with intel Xeon CPU-1225 3.2 GHz, RAM 8 Gb, Windows 7 Professional 64-bit (compA) and FK BY workstation with Intel Core i7, 3 GHz, RAM 8 Gb, Windows 10 64-bit (CompF) as well as various laptops with worse parameters. Inverse high-throughput virtual screening (inv HTVS) data processing was done using Hewlett Packard Probook 450 notebook with Intel Core i5-3230M 2.6 GHz, RAM 4 GB working under 64-bit Windows 7. Information from web resources www.stackoverflow.com, www.geeksforgeeks.org, www.support.microsoft.com, https://github.com, www.planetaexcel.ru and others was used to realize coding for the helper software creation.

Dataset preparation

The 461 SARS-CoV-2 protein 3D structures files with code range from 5R7Y to 7LKV, available to the beginning of 2021 (now at the Jan 2022 ~1700 are available), were taken from Protein Data Bank database (https://www.rcsb.org). Supplementary Table S1 lists PDB accession numbers and names for proteins under consideration. Only records about AAs atom coordinates belonging to chain A were used for further processing, but the rest of information was removed. Further formatting was done using standard MGL Tools [15] Python script prepare_receptor4.py resulting in a set of protein files in ready-to-use pdbqt format. Ligand 3D or, if unavailable, 2D structures files were taken from Pubchem (https://pubchem.ncbi.nlm.nih.gov) and BioGem (https://pdt.biogem.org) [16] databases. Smiles codes and substructure search options of the databases were used to find structures with α,β-unsaturated carbonyl and few others electrophilic motifs. Structures for compounds with trivial names were selected preferably. Finally, a library of ~2000 compounds was collected. Open Babel 2.4.1 software [17] was used for conversion of the selected files from both sdf and smi formats to .pdb format using 3D-structure generation and minimization commands (–gen3d and -ff uff -n 800 -cg -c 1e-8, respectively). Then the files were formatted using standard MGL Tools Python script prepare_ligand4.py resulting in ligand files in pdbqt format.

Description of selected ligands set

The 1788 structures of compounds were used. Majority of them were α,β-unsaturated carbonyl-bearing linear esters and lactones, ketones and amides which can be found in plants or other natural sources. Such structures were selected because they were reported to make covalent bonds with nucleophilic atoms of AAs in various proteins. In particular, α,β-unsaturated amide motif has become a classic for design of approved drug covalent inhibitors targeting Cys residues in kinases [18]. α,β-Unsaturated lactones are reported to attack Cys [22][23], His [22] and even Lys [22], whereas similar properties are monitored for linear esters too [24]. α,β-unsaturated ketones were reported to bind with Cys [25] and His [24]. Previous reports demonstrated examples of covalent inhibition of coronaviral enzymes by such electrophilic natural compounds, e.g. MPro by unsaturated lactone andrographolide [25], artecanin [12] and oridonin [21].

Docking and docking-based data processing

Docking-based virtual screening was performed using Autodock Vina [19]. Parameters for grid box (4x4x4 cubic nm, centered on the center of each individual protein pdbqt file), exhaustiveness (12) and number of models (5) were identical for all calculations.

The docking results were processed using python script process_VinaResults.py and the best poses were selected. For further analysis of predicted protein-ligand complexes Python script binana.py [20] was used. The script was modified to save protein-ligand complex data in one file and to operate with more than 1000 AAs numbers.

Parameters description

Various parameters for each protein-ligand interaction were tabulated and analyzed, in particular, binding energy (E_{bind}, kcal/mol), AAs and their atoms within 0,4 nm from atoms of a ligand (close contacts), ligand codes, PDB codes and trivial names for proteins etc. Only interactions with E_{bind} of -6.9 or less were taken for detailed analysis. The colocalization of ligands electrophilic atoms (selected manually) with nucleophilic atoms of Cys (-SH, SG in pdbqt files) and His (any of imidazole N-atoms, ND1 or NE2) residues within 0.4 nm was analyzed to guess whether the interaction have a potential to covalent bond formation (“distance criterium”).

The number of different protein-ligand complexes meeting the distance criterium for a certain amino acid residue (Cys was used as default) and E_{bind} criterium for the same protein was also used as a main criterium to highlight repeatability of such type of interaction in spite of variability between different PDB structures (“repeatability criterium”). Trivial names for hit ligands were retrieved from aforesaid databases. In the cases of absence of the names, trivial names of similar compounds with suffix -like were mentioned; the similarity was determined based on “Substructure” or, if not, “Similarity” functions of Pubchem database.

Also, two more parameters were calculated for more effective interaction assessment: $E_{relative}$ and $E_{average}$. The relative binding energy ($E_{relative}$, kcal/mol×atom) of protein-ligand interactions was calculated using the

following equation:

$$E_{relative}, \frac{kcal}{mol * atom} = \frac{E_{bind}, \frac{kcal}{mol}}{Number\ of\ ligand\ atoms} * 1000$$

where the number of ligand atoms was taken from the corresponding ligand.pdbqt file for a calculated interaction, E_{bind} is the minimal binding energy for this interaction. The criterium depict effectivity of a ligands' atoms "usage" to bind with a certain protein and should be less than -250 ("$E_{relative}$ criterium").

The average binding energy ($E_{average}$, kcal/mol×complex) was evaluated as follows:

$$E_{average}, \frac{kcal}{mol * complex} = \frac{\sum E_{bind}}{\text{Repeatability criterium}}$$

where denominator is the number of complexes meeting all criteria for the same protein and numerator is the sum of this complexes minimal binding energies. The parameter allows for a more accurate assessment of the formation a bond between a given ligand and protein possibility. We assume that the closer $E_{average}$ to E_{bind}, the better the ligand binds to various protein structures of the same protein, and therefore the greater probability of the ligand effective binding to the protein in practice.

All the procedures in the workflow described above were semi-automatically done using an original helper program tool named FYTdock, which is based on aforesaid known Python scripts and few original scripts as well as Microsoft Excel book files with macroses (see Supplementary Video_1 and Video_2).

Results And Discussion

We tested 1788 structures of low molecular weight compounds and 461 structures of SARS-CoV-2 proteins (chains A were only used), assuming *in silico* modeling of ~ 824 000 protein-ligand complexes. Overall for all ligands and SARS-CoV-2 protein's structures under consideration, about 190 000 poses were calculated to have E_{bind} values − 6.9 or less; among them a subset of ~ 21 800 complexes demonstrated proximity of thiol sulfur (SG atom) to any atom of the corresponding ligand no more than 0,4 nm. To the date the information about the complexes in the set was additionally analyzed according to the aforesaid E_{bind} and $E_{relative}$ criteria. $E_{average}$ and repeatability criteria are used for an additional analysis to characterize the interactions. Usage of the aforesaid helper program based on widely-used Autodock Vina and minorly-modified published binana.py script allowed to generate valid data and process them in a convenient way. A possibility to arrange such inv HTVS on own computers strongly impacts on confidentiality of such projects and allow to avoid queues on public web-resources. For the analysis we used semi-automatic proximity assay of manually-assigned electrophilic atoms for 703 ligands matching the above criteria

(Supplementary Table S2) yielding 875 suitable protein-ligand complexes (Supplementary Table S3). Among the complexes we have highlighted hits for MPro (Table 1), NSP16 (Table 2) and NSP12 (Table 3).

Table 1

The PubChem IDs, ligand names, pdb codes, AutoDock Vina minimal binding energies (E_{bind}, kcal/mol) and other criteria and information for the list of structures in complexes with given protein stuctures of MPro sorted with respect to the repeatability criterium (RC).

PubChem ID & ligand name	Pdb code	E_{bind}	$E_{relative}$	Distance, Å	Targeted AA	$E_{average}$	RC (CYS)	RC (HIS)
CID101821135 Ovatodiolide	5R80	-8.6	-358	3.61	CYS145	-7.8	26	3
CID16116606	5RGN	-7.7	-321	3.99	CYS156	-7.3	18	-
CID13890811 15-Oxosteviol	5RH1	-7.2	-277	3.89	CYS156	-7.0	15	-
CID453153	5RH1	-7.2	-300	3.89	CYS156	-7.0	14	-
CID392454	5RGN	-7.2	-300	3.89	CYS156	-7.0	12	-
CID10476763 Peperomine E	5RH9	-8.4	-250	3.66	CYS145	-7.7	12	4
CID57380389 Methylgeopyxin E	5RH3	-7.2	-240	3.97	CYS156	-7.0	11	-
CID16114788	5RH9	-7.2	-313	3.83	CYS156	-7.1	10	-
CID5281373 Jatrophone	5RH8	-8	-285	3.44	CYS145	-7.5	9	-
CID5384663 Isodehydroeupatundin	6XMK	-8.1	-218	3.72	CYS145	-7.4	9	-
CID11624798 Zambesiacolactone B	7JYC	-7.9	-272	3.79	CYS145	-7.5	8	5
CID102239748 Agriantholide	5REA	-7.6	-262	3.65	CYS145	-7.4	7	3
CID44398293 Thieleanin	7C8U	-7.1	-394	3.83	CYS145	-7.0	6	-
CID10410482 Insulicolide A	7C8B	-7.8	-236	3.81	CYS145	-7.4	4	-
CID11056679 Xeniolide F	7D1O	-8	-240	3.88	CYS145	-7.6	3	-

PubChem ID & ligand name	Pdb code	E_{bind}	$E_{relative}$	Distance, Å	Targeted AA	$E_{average}$	RC (CYS)	RC (HIS)
CID21597452 2-deoxycucurbitacin D	6Y2G	-8.1	-291	3.87	CYS145	-7.7	2	-
CID21603251 Chaetoglobosin C	6ZRT	-9.4	-229	3.98 3.68	CYS145	-9.4	1	-
BXGC0053659 Sculponeatin A	5RFD	-7.5	-278	3.92	CYS145	-7.5	1	-
CID9975896 Epinodosin	7CX9	-7.6	-271	3.52	CYS145	-7.6	1	1

Table 2

The PubChem IDs, ligand names, pdb codes, AutoDock Vina minimal binding energies (E_{bind}, kcal/mol) and other criteria and information for the list of structures in complexes with given protein stuctures of NSP16 sorted with respect to the repeatability criterium (RC).

PubChem ID & ligand name	Pdb code	E_{bind}	$E_{relative}$	Distance, Å	Targeted AA	$E_{average}$	RC (CYS)
CID325147 Anhydroverlotolin-like	7JPE	-8.2	-456	3.80	CYS115(6913)	-8.0	11
CID15866741 Costunolide-like	6YZ1	-8.1	-426	3.67	CYS115(6913)	-7.7	11
CID14605949 Ivalbatin-like	7JHE	-7.5	-417	3.94	CYS115(6913)	-7.2	9
CID22226 Mexicanin E	7JHE	-7.8	-459	3.89	CYS115(6913)	-7.5	8
CID282784	7JHE	-7.2	-400	3.76	CYS115(6913)	-7.0	8
CID139589877 Insulicolide B	7JHE	-9.5	-306	3.92	CYS115(6913)	-9.1	5
CID5932623	6W75	-9.4	-269	3.78	CYS115(6913)	-9.1	5
CID44575708 Geoditin A	7C2J	-9.7	-294	3.92	CYS115(6913)	-9.5	4
CID14109732	6W61	-8	-333	3.61	CYS115(6913)	-7.6	3
CID10476763 Peperomine E	6W4H	-8.4	-280	3.90	CYS115(6913)	-8.3	2
CID10957236	7BQ7	-7.6	-362	3.90	CYS115(6913)	-7.6	2
CID133556460	6WKQ	-7.5	-417	3.84	CYS115(6913)	-7.5	2
CID14355826 Ludartin	7JPE	-7.1	-394	3.85	CYS115(6913)	-7.1	2

Table 3

The PubChem IDs, ligand names, pdb codes, AutoDock Vina minimal binding energies (E_{bind}, kcal/mol) and other criteria and information for the list of structures in complexes with given protein stuctures of NSP12 sorted with respect to the repeatability criterium (RC).

PubChem ID & ligand name	Pdb code	E_{bind}	$E_{relative}$	Distance, Å	Targeted AA	$E_{average}$	RC (CYS)
CID101553163 Ixerin D	7BZF	-8.8	-251	CYS395	3.91	-8.5	4
CID132494155 Xindongnin C	7CTT	-7.8	-236	CYS395	3.81	-7.6	4
CID145993575 Isotelekin ester	6XQB	-9.1	-314	CYS395	3.98	-9.0	2
CID45268596	7BV2	-9	-281	CYS395	3.58	-9.0	2
CID101883322	6XEZ	-8.9	-387	CYS395	3.57	-8.2	2
CID14109027 Rabdosin B	6XQB	-8.4	-255	CYS395	3.89	-7.9	2
CID6439017 Kericembrenolide A	7BW4	-7.4	-285	CYS395	3.60	-7.2	2
CID71718234 Withametelin F	7BTF	-8.9	-270	CYS395	3.989	-8.9	1
CID101891076 Insulicolide B	6M71	-8.2	-248	CYS395	3.91	-8.2	1
CID6443871 Peniophorin A	6XQB	-7.7	-335	CYS697	3.24	-7.7	1
CID102034411 Rabdosin A	7BV1	-7.7	-275	CYS395	3.62	-7.7	1
BXGC0053409 Epinodosin	7BW4	-7.7	-275	CYS395	3.92	-7.7	1
CID20056132 Maoecrystal P	7BW4	-7.7	-296	CYS395	3.79	-7.7	1
CID44605338 Crotonkinensin A	7BZF	-7.6	-304	CYS395	3.75	-7.6	1

The Mpro hits include 15-oxosteviol, a semi-synthetic derivative of stevioside from stevia plant, used as a common sweetening 933 [27][28], methylgeopyxin E from endolichenic fungi *Geopyxis sp.* [29], peperomine E (Fig. 1) from *Peperomia dindygulensis* [30], agriantholide from plant *Agrianthus pungens* [31], three 15-oxo-kaur-16-en acids, thieleanin (Fig. 1) from *Viguiera sylvatica* [32], jatrophone from *Jatropha gossypifolia* [33], ovatodiolide from *Anisomeles indica* (catmint) [26], isodehydroeupatundin (Fig. 1), zambesiacolactone B from Aframomum zambesiacum [34] and *Helianthus annuus* (sunflower) [35], agriantholide from Agriantus pungens, thieleanin from Viguiera sylvatica [32][36], insulicolide A from the marine fungus *Aspergillus insulicola* [37], xeniolide F from coral of the genus Xenia [38], 2-deoxycucurbitacin D from *Luffa amara* [63], chaetoglobosin C from the mold *Chaetomium globosum*, sculponeatin A and epinodosin from Isodon species (Table 1). To the best of our knowledge no one of the compound have not been reported yet as a potential covalent inhibitor of SARS-CoV-2 Mpro. Peperomine E from a non-toxic plant *Peperomia dindygulensis* (a type of peperomia is commonly known as a “radiator plant”) with a reported ability to act as Cys-targeting covalent inhibitor of DNA-methyl transferase DNMT1 [30][41][42] seems to provide good options to be recommended for further evaluation as a new covalent inhibitor of the viral protein. Anti-inflammatory properties of jatrophone related with NF-kB pathway suppression [33][43] [45] could be an advantage in reducing the exposure of the lungs to Mpro SARS-CoV-2. [44]. An ovatodiolide suppressing effect on functions of YAP1 protein [39], a substrate of MPro [40], might cause interactions at COVID-19 organism state. Other compounds meeting aforesaid E_{bind}, $E_{relative}$ and distance criteria can be found (Supplementary Table S3). It is an important point to note on the example of peperomine E in silico interactions with MPro. Notably that only 11 structures of Mpro from 197 used for the job have met the all criteria to mark Peperomine E as a potential covalent inhibitor of Mpro. Thus, it is highly likely to miss a potential hit in the case of a single chosen structure for direct HTVS. From the other hand, the numbers indicate that peperomine E has not met all criteria for major part of Mpro structures and, thus, has a chance to be not a suitable inhibitor of the SARS-CoV-2 enzyme.

The NSP16 hits (Table 2) include anhydroverlotolin-like compound (CID325147, Fig. 2) with both unsaturated lactone and unsaturated ketone moieties from *Artemisia verlotiorum* [47], ivalbatin-like compound (CID14605949, Fig. 2) from *Xanthium sibiricum* and *Xanthium strumarium*, mexicanin E from *Helenium* [46], insulicolide B from the marine-derived fungus *Aspergillus ochraceus* [48], geoditin A (Fig. 2) from marine sponge *Geodia japonica* [49], peperomine E and some others (Supplementary Table S3).

The NSP12 hits include ixerin D from *Taraxacum* (common dandelion) species, xindongnin C, rabdosins A and B, epinodosin from different *Isodon* species, cytotoxic kericembrenolide A from soft coral *Clavularia koellikeri* [50], pipernonaline from *Piper longum* [51], withametelin F from *Datura metel* or other *Solanaceous sp.* [52], insulicolide B, isotelekin ester (isotelekin itself is found in Inula racemosa [53], Fig. 3), peniophorin A from fungi *Peniophora affinis* [54], maoecrystal P [55], crotonkinensin A from *Croton tonkinensis* [56] (Table 3). Notably that crotonkinensin A, rabdosin B and withametelin F (Fig. 3) are reported to have anti-inflammatory properties. Anti-inflammatory properties xindongnin A and B were reported to be related with suppression of NF-kB pathway [57]. The same was found for steroid hormone-like guggulsterone E [58] (Supplementary Table S3).

Conclusion

Thus, about 30 new potential covalent inhibitors of SARS-CoV-2 proteins were predicted based on *in silico* inverse high-throughput virtual screening (inv HTVS) of only about 2000 initial structures (Tables 1–3). Such relatively high number of hits (more than 1%) with respect to relatively small initial set of structures was achieved due to direct docking evaluation of each protein-ligand interaction as well as using multiple different structures for the same SARS-CoV-2 protein. Usage of our small helper program to arrange, run, tabulate and process results of inv HTVS has allowed us to do all principal steps of the job within about 8 months only using for calculations mainly two workstations with moderate parameters. Among the hits found for Mpro we would like to highlight one more time α-methylene ketone 15-oxosteviol, which could be received semi-synthetically and related with commonly-used food sweetening 933, α-methylene lactones peperomine E from a non-toxic peperomia species and zambesiacolactone B from easily available non-toxic of sunflower leaves. The tree hit compounds are from available sources. Similarly, for NSP12 hits we would like to highlight ixerin D from common non-toxic dandelion, which is used orally in different folk medicines. The compounds have no easy hydrolysable ester bonds in spite of the lactone moiety and this circumstance is also in favor of their potential further evaluations as novel tools to cure or reducing severity of SARS-CoV-2 infection. This report offers some compounds for further evaluation as potential covalent inhibitors of SARS-CoV-2 proteins aiming a global task to develop new natural compounds-based tools for prophylaxis, reduction of severity and might be even treatment of COVID-19.

Declarations

Acknowledgement

The work was supported by grant BRFFI X21COVID-030; helper software creation was supported by GPSR № 20210560.

Impacts

Y.F. designed the research, developed FYTdock helper program (coded on Python and Visual Basic, organized data workflow and processing), participated in compounds library creation and annotation, discussed results.

V.St. consulted about Python coding, tested FYTdock helper program, participated in compounds library creation and annotation, wrote and formatted the paper's text, discussed results.

N.D. participated in compounds library creation and annotation,

V.Sh. wrote and formatted the paper's text, discussed results, participated in compounds library creation and annotation.

References

1. Machhi J et al (2020) The natural history, pathobiology, and clinical manifestations of SARS-CoV-2 infections. J Neuroimmune Pharmacol 15:359–386
2. Wong NA, Saier MH (2021) The SARS-coronavirus infection cycle: a survey of viral membrane proteins, their functional interactions and pathogenesis. Int J Mol Sci 22:1308
3. Astuti I, Ysrafil Y (2020) Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2): An overview of viral structure and host response. Diabets Metab Syndr 14:407–412
4. Cucinotta D, Vanelli M (2020) WHO declares COVID-19 a pandemic. Acta Biomed 91:157–160
5. Xu J (2021) Drug repurposing approach to combating coronavirus: potential drugs and drug targets. Med Res Rev 41:1375–1426
6. Fu L et al (2020) Both Boceprevir and GC376 efficaciously inhibit SARS-CoV-2 by targeting its main protease. Nat Commun 11:4417
7. Ribone SR et al (2021) Target identification for repurposed drugs active against SARS–CoV–2
8. via high–throughput inverse docking.Journal of Computer-Aided Molecular Designhttps://doi.org/10.1007/s10822-021-00432-3
9. Sarkar C et al (2021) Ebselen suitably interacts with the potential SARS-CoV-2 targets: an in-silico approach. J Biomol Struct Dyn. https://doi.org/10.1080/07391102.2021.1971562
10. Jin Z et al (2020) Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. Nature 582:289–293
11. Patel CN et al (2021) Identification of antiviral phytochemicals as a potential SARS-CoV-2 main protease (Mpro) inhibitor using docking and molecular dynamics simulations. Sci Rep 11:20295
12. Chen Z et al (2021) Ginkgolic acid and anacardic acid are specific covalent inhibitors of SARS-CoV-2 cysteine proteases. Cell Biosci 11:45
13. Al-Shuhaib MBS et al (2022) Artecanin of Laurus nobilis is a novel inhibitor of SARS-CoV-2 main protease with highly desirable drug-likeness. J Biomol Struct Dyn. https://doi.org/10.1080/07391102.2022.2030801
14. Mineeva IV et al Synthesis of new thiazolo[3,2-a]pyrimidine derivatives and in silico analysis of their bioactivity. *Proceedings of the National Academy of Sciences of Belarus, Chemical Series* **57**:456462
15. Faletrov YV et al (2021) In silico modeling of isoniazid-steroid conjugates interactions with mycobacterial cytochromes P450 and their bioconversion in vitro by the cells. Biochem (Moscow) Supplement Ser B: Biomedical Chem 15:111–118
16. Sanner MF (1999) Python: a programming language for software integration and development. J Mol Graph Model 17:57–61
17. Ashok Kumar T, Rajagopal B (2017) PDTDB - an integrative structural database and prediction server for plant metabolites and therapeutic drug targets. Int J Curr Res 9:46537–46541
18. O'Boyle NM (2011) Open Babel: an open chemical toolbox. J Cheminform. https://doi.org/10.1186/1758-2946-3-33

19. Estupiñán HY et al (2021) BTK gatekeeper residue variation combined with cysteine 481 substitution causes super-resistance to irreversible inhibitors acalabrutinib, ibrutinib and zanubrutinib. Leukemia 35:1317–1329
20. Trott O, Olson AJ (2009) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 31:455–461
21. Jacob D, Durrant and J. Andrew McCammon. BINANA: a novel algorithm for ligand-binding characterization.J. Mol. Graph. Model.**29**:888–893
22. Zhong B et al (2022) Oridonin Inhibits SARS-CoV-2 by Targeting Its 3C-Like Protease. Small sci https://doi.org/10.1002/smsc.202100124
23. Shin M et al (2017) Hsp72 is an intracellular target of the α,β-unsaturated sesquiterpene lactone, Parthenolide. ACS Omega 2:7267–7274
24. Balaguer FA et al (2019) Crystal structure of the cyclostreptin-tubulin adduct: implications for tubulin activation by taxane-site ligands. Int J Mol Sci 20:1392
25. Yoshizawa M et al (2018) Identification of the histidine residue in vitamin D receptor that covalently binds to electrophilic ligands. J Med Chem 61:6339–6349
26. Yu X et al Eriocalyxin B inhibits STAT3 signaling by covalently targeting STAT3 and blocking phosphorylation and activation of STAT3.PLoS One.https://doi.org/10.1371/journal.pone.0128406
27. Yu CY et al (2018) Ovatodiolide isolated from Anisomeles indica induces cell cycle G2/M arrest and apoptosis via a ROS-dependent ATM/ATR signaling pathways. Eur J Pharmacol 819:16–29
28. Procinska E, Bridges BA, Hanson JR (1991) Interpretation of results with the 8-azaguanine resistance system in Salmonella typhimurium: no evidence for direct acting mutagenesis by 15-oxosteviol, a possible metabolite of steviol. Mutagenesis 6:165–167
29. Khazayel S et al Derivative of Stevioside, CPUK02, Restores ESR1 gene methylation in MDA-MB 231.Asian Pac. J. Cancer Prev.**19**:2117–2123
30. Wijeratne EM et al (2012) Geopyxins A-E, ent-kaurane diterpenoids from endolichenic fungal strains Geopyxis aff. majalis and Geopyxis sp. AZ0066: structure-activity relationships of geopyxins and their analogues. J Nat Prod 75:361–369
31. Wang XZ et al (2018) Peperomin E induces promoter hypomethylation of metastatic-suppressor genes and attenuates metastasis in poorly differentiated gastric cancer. Cell Physiol Biochem 50:2341–2364
32. Vujić B et al (2020) Composition, antioxidant potential, and antimicrobial activity of Helichrysum plicatum DC. various extracts. *Plants.* https://doi.org/10.3390/plants9030337 (2020)
33. Taylor PG et al (2008) Anticancer activities of two sesquiterpene lactones, millerenolide and thieleanin isolated from Viguiera sylvatica and Decachaeta thieleana. Fitoterapia 79:428–432
34. Li K, Xu Y, Yue W (2020) Anti-viral activity of jatrophone against RSV-induced respiratory infection via increase in interferon-γ generating dendritic cells. Environ Toxicol 35:888–894
35. Kenmogne M et al (2006) Five labdane diterpenoids from the seeds of Aframomum zambesiacum. Phytochemistry 67:433–438

36. El Marsni Z et al (2015) Isolation of bioactive compounds from sunflower leaves (Helianthus annuus L.) extracted with supercritical carbon dioxide. J Agric Food Chem 63:6410–6421
37. Dupuy OA et al (2008) [Sesquiterpene lactones of the plants Viguiera sylvatica and Decachaeta thieleana (Asteraceae) modulate nitric oxide production and phagocytosis of macrophages cell line RAW]. Rev Biol Trop 56:1063–1073
38. Rahbæk L et al (1997) Insulicolide A: a new nitrobenzoyloxy-substituted sesquiterpene from the marine fungus Aspergillus insulicola. J Nat Prod 60:811–813
39. Pollex A, Hiersemann M (2005) Catalytic asymmetric claisen rearrangement in natural product synthesis: synthetic studies toward (-)-xeniolide F. Org Lett 7:5705–5708
40. Huang YJ et al (2017) Ovatodiolide suppresses colon tumorigenesis and prevents polarization of M2 tumor-associated macrophages through YAP oncogenic pathways. J Hematol Oncol. https://doi.org/10.1186/s13045-017-0421-3
41. Pablos I et al (2021) Mechanistic insights into COVID-19 by global analysis of the SARS-CoV-2 3CL pro substrate degradome. Cell Rep. https://doi.org/10.1016/j.celrep.2021.109892
42. Lin M et al (2021) Peperomin E Induces Apoptosis and Cytoprotective Autophagy in Human Prostate Cancer DU145 Cells In Vitro and In Vivo. Planta Med https://doi.org/ 10.1055/a-1348-1634
43. Wang XZ et al (2016) The natural secolignan peperomin E induces apoptosis of human gastric carcinoma cells via the mitochondrial and PI3K/Akt signaling pathways in vitro and in vivo. Phytomedicine 23:818–827. https://doi.org/10.1016/j.phymed.2016.04.001
44. Fröhlich JK et al (2017) Antinociceptive and anti-inflammatory activities of the Jatropha isabellei dichloromethane fraction and isolation and quantitative determination of jatrophone by UFLC-DAD. Pharm Biol 55:1215–1222
45. Li W et al (2021) SARS-CoV-2 Nsp5 activates NF-κB pathway by upregulating SUMOylation of MAVS. Front Immunol 12:750969
46. Yuan L et al (2020) Inhibition of glioma cell viability by jatrophone via NF-κB down-regulation and apoptosis inhibitor up-regulation. Archives of Medical Science. https://doi.org/10.5114/aoms.2020.94439
47. Elsohly MA et al (1979) Constituents of Helenium amarum. II. Isolation and characterization of heleniamarin and other constituents. J Nat Prod 42:450–454
48. Geissman TA (1970) Sesquiterpene lactones of Artemisia A. verlotorum and A. vulgaris. Phytochemistry 9:2377–2381
49. Tan Y et al (2022) A nitrobenzoyl sesquiterpenoid insulicolide a prevents osteoclast formation via suppressing c-Fos-NFATc1 signaling pathway. Front Pharmacol. https://doi.org/10.3389/fphar.2021.753240
50. Cheung FW et al (2012) Anti-melanogenic property of geoditin A in murine B16 melanoma cells. Mar Drugs 10:465–476
51. Kobayashi M et al (1986) Kericembrenolides A, B, C, D, and E, five new cytotoxic cembrenolides from the Okinawan soft coral Clavularia koellikeri. Chem Pharm Bull 34:2306–2309

52. Lee W et al (2013) Pipernonaline from Piper longum Linn. induces ROS-mediated apoptosis in human prostate cancer PC-3 cells. Biochem Biophys Res Commun 430:406–412
53. Zhang H, Cao CM, Gallagher RJ (2014) Timmermann, B. N. Antiproliferative withanolides from several solanaceous species. Nat Prod Res 28:1941–1951
54. Ma Y, Zhao DG, Gao K (2013) Structural investigation and biological activity of sesquiterpene lactones from the traditional Chinese herb Inula racemosa. J Nat Prod 76:564–570
55. Gerber NN, Shaw SA, Lechevalier HA (1980) Structures and antimicrobial activity of peniophorin A and B, two polyacetylenic antibiotics from Peniophora affinis Burt. Antimicrob. Agents Chemother 17:636–641
56. Su F et al (2018) Total synthesis of Maoecrystal†…P: application of a strained bicyclic synthon. Angew Chem Int Ed Engl 57:760–764
57. Thuong PT et al (2009) Crotonkinensins A and B, diterpenoids from the Vietnamese medicinal plant Croton tonkinensis. J Nat Prod 72:2040–2042
58. Leung CH et al (2005) Novel mechanism of inhibition of nuclear factor-kappa B DNA-binding activity by diterpenoids isolated from Isodon rubescens. Mol Pharmacol 68:286–297
59. Lv N et al Guggulsterone, a plant sterol, inhibits NF-kappaB activation and protects pancreatic beta cells from cytokine toxicity.Mol. Cell Endocrinol.**289**:49–59

Supplementary Materials

Supplementary materials to the work can be downloaded from Google drive link https://drive.google.com/drive/folders/1Z31UbVvHL3VIXajl83bJ2AUfO04GSSDx?usp=sharing.
The supplementary materials include
1) Supplementary Table S1 in Supplementary_Tables.xlsx file containing a list of PDB accession numbers and names of proteins under consideration;

2) Supplementary Table S2 in Supplementary_Tables.xlsx file containing a list of accession numbers for ligands under consideration;

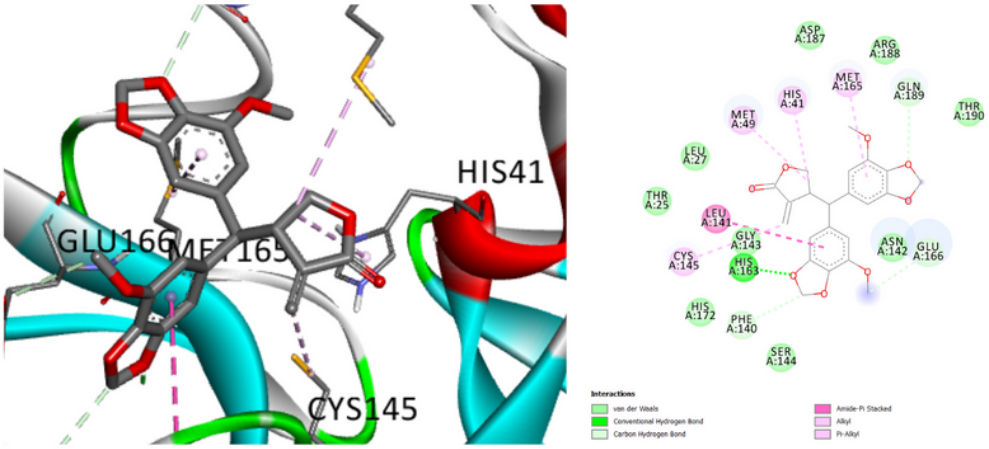
3) Supplementary Table S3 in Supplementary_Tables.xlsx file containing a table describing selected in silico modeled protein-ligand complexes in which ligands' electrophilic atoms are within 0.4 nm from sulfur of the proteins' cysteine residues (for detailed interpretation of the information given in Table S3, please, see sheet "description" in Supplementary_Tables.xlsx file;

4) FYTdock_video_1.avi and FYTdock_video_2.avi video files describing major details about workflow of inverse high-throughput virtual screening presented in the paper;

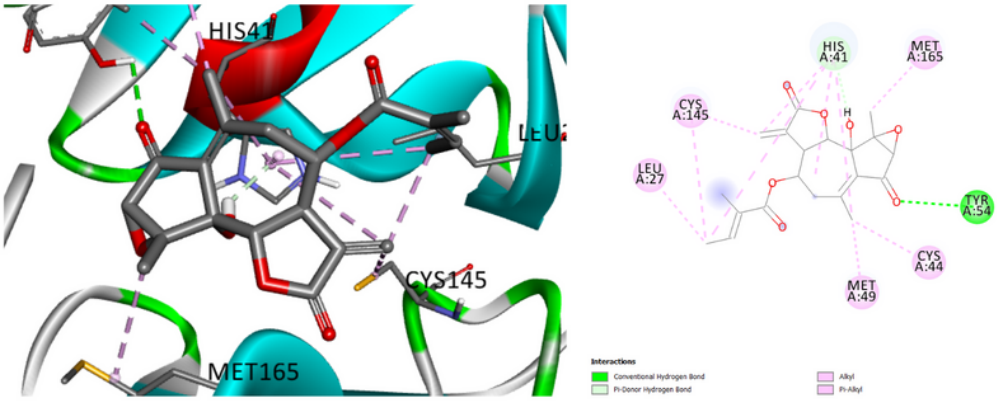
5) Supplementary_FYTdock_version_1_files.zip archive file containing files of FYTdock helper program tool used to inverse high-throughput virtual screening presented in the paper;

6) Supplementary_selected protein-ligand complexes described in TABLES 1,2 & 3.zip archive file containing information about structures (coordinates of atoms)of some selected in silico modeled protein-ligand complexes (the files can be open using MGL Tools) and the data about the in silico protein-ligand interactions analysis.

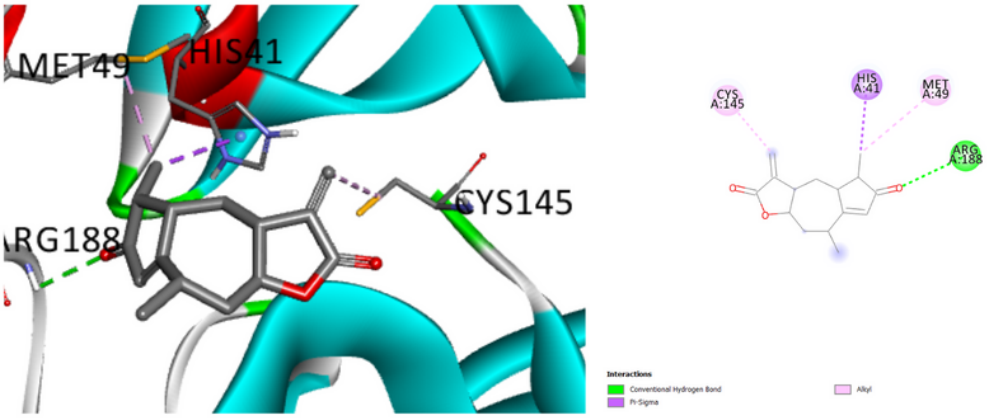
Figures



Peperomine E and 5RH9
Binding energy = -8.4 kcal/mol



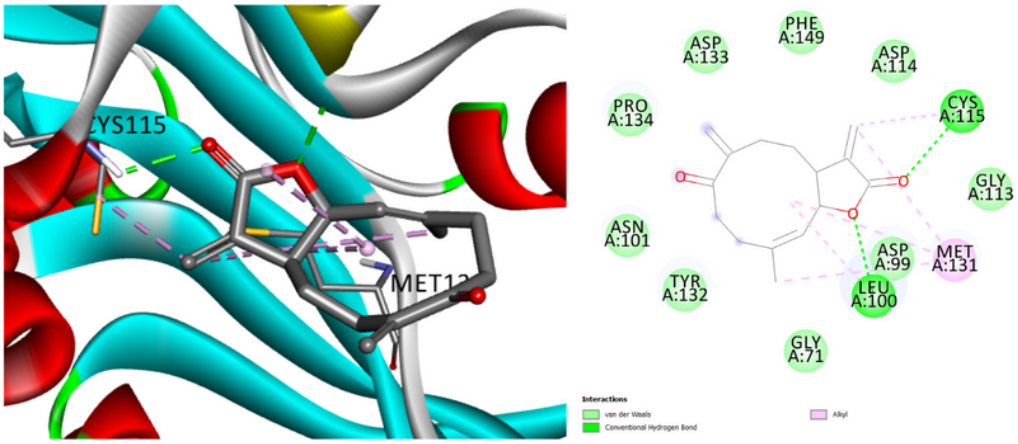
Isodehydroeupatundin and 6XMK
Binding energy = -8.1 kcal/mol



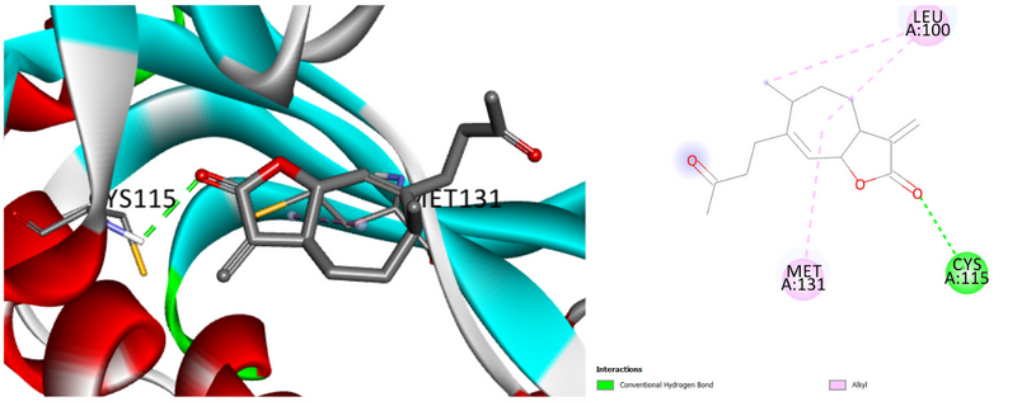
Thieleanin and 7C8U
Binding energy = -7.1 kcal/mol

Figure 1

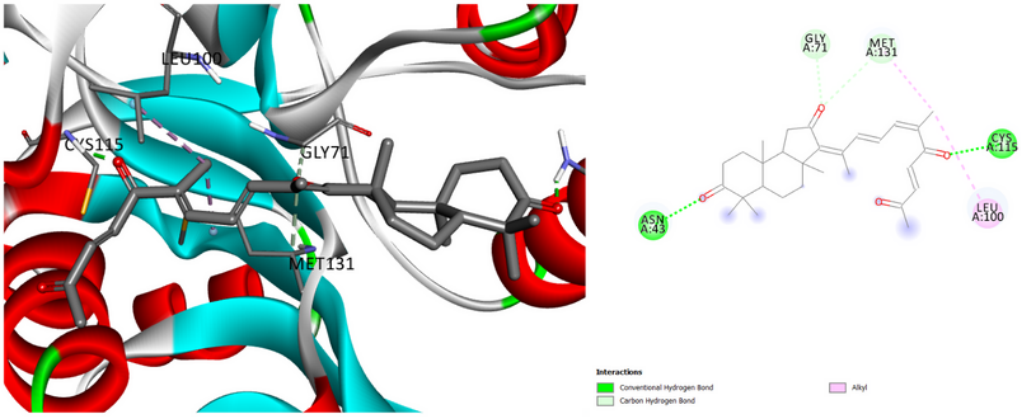
Complexes of ligands in the binding pocket of corresponding structures of SARS-CoV-2 MPro shown in 3D and 2D representations demonstrating ligand electrophilic atom-CYS145 interplay and other ligand-protein interactions (Discovery studio visualizer v21.1.0.20298; https://www.3ds.com/products-services/biovia/).



CID325147 and 7JPE
Binding energy = -8.2 kcal/mol



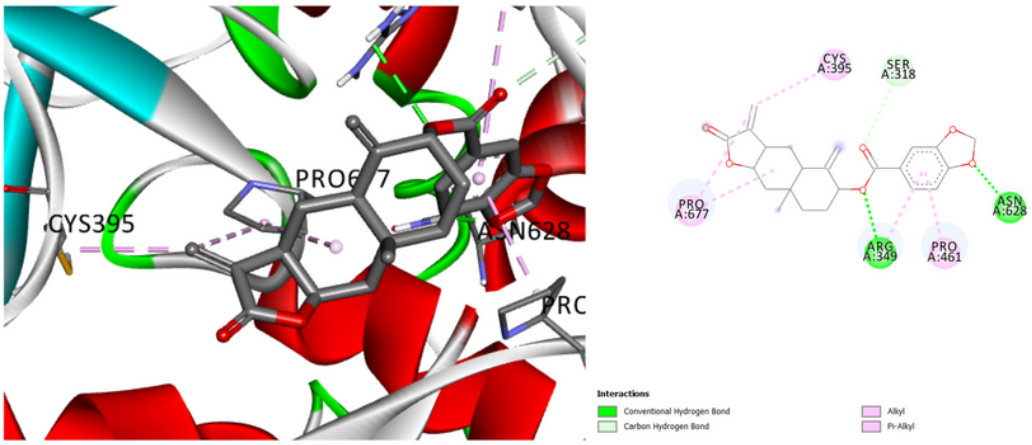
CID14605949 and 7JHE
Binding energy = -7.5 kcal/mol



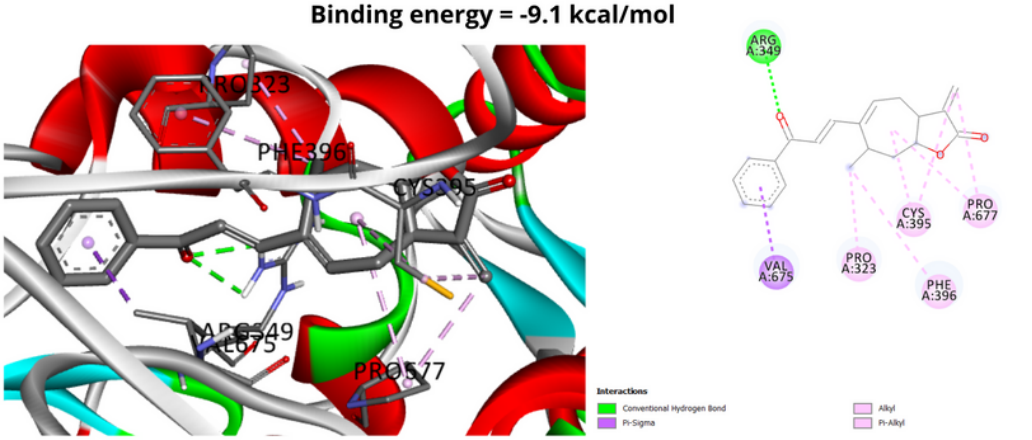
Geoditin A and 7C2J
Binding energy = -9.7 kcal/mol

Figure 2

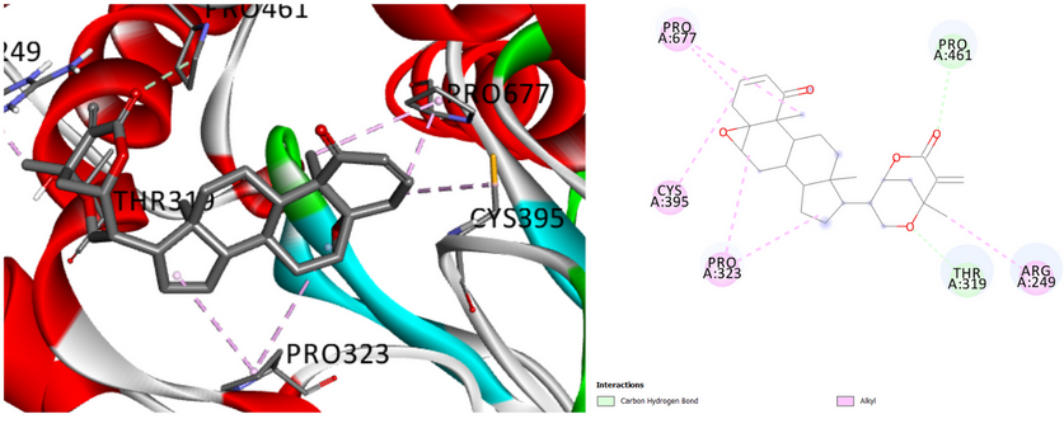
Interaction of three ligands in the binding pocket of corresponding structures of SARS-CoV-2 NSP16 shown in 3D and 2D representations demonstrating ligand electrophilic atom-CYS115 interplay and other ligand-protein interactions (Discovery studio visualizer v21.1.0.20298; https://www.3ds.com/products-services/biovia/).



Isotelekin ester and 6XQB
Binding energy = -9.1 kcal/mol



CID101883322 and 6XEZ
Binding energy = -8.9 kcal/mol



Withametelin F and 7BTF
Binding energy = -8.9 kcal/mol

Figure 3

Interaction of three ligands in the binding pocket of corresponding structures of SARS-CoV-2 NSP12 shown in 3D and 2D representations demonstrating ligand electrophilic atom-CYS395 interplay and other ligand-protein interactions (Discovery studio visualizer v21.1.0.20298; https://www.3ds.com/products-services/biovia/).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTables.xlsx
- SupplementaryselectedproteinligandcomplexesdescribedinTABLES123.zip
- Graphicalabstract.jpg