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.

Journal of Molecular Structure 1261 (2022) 132951

Contents lists available at ScienceDirect

Journal of Molecular Structure

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

Virtual screening of molecular databases for potential inhibitors of the NSP16/NSP10 methyltransferase from SARS-CoV-2

João Pedro Agra Gomes[a], Larissa de Oliveira Rocha[a], Cíntia Emi Yanaguibashi Leal[b], Edilson Beserra de Alencar Filho[a,b,c,*]

[a] College of Pharmacy, Federal University of San Francisco Valley, Petrolina, Pernambuco, Brazil
[b] Postgraduate Program in Biosciences, Federal University of San Francisco Valley, Petrolina, Pernambuco, Brazil
[c] Postgraduate Program in Health and Biological Sciences, Federal University of San Francisco Valley, Petrolina, Pernambuco, Brazil

ARTICLE INFO

Article history:
Received 23 November 2021
Revised 22 March 2022
Accepted 27 March 2022
Available online 28 March 2022

Keywords:
SARS-CoV-2
Virtual screening
Molecular docking
NSP10/NSP16
S-adenosylmethionine
Caatinga Brazilian biome

ABSTRACT

COVID-19 is a disease caused by the SARS-CoV-2 virus and represents one of the greatest health problems that humanity faces at the moment. Therefore, efforts have been made with the objective of seeking therapies that could be effective in combating this problematic. In the search for ligands, computational chemistry plays an essential role, since it allows the screening of thousands of molecules on a given target, in order to save time and money for the *in vitro* or *in vivo* pharmacological stage. In this paper, we perform a virtual screening by docking looking for potential inhibitors of the NSP16-NSP10 protein dimer (methyltransferase) from SARS-CoV-2, by evaluating a homemade databank of molecules found in plants of the Caatinga Brazilian biome, compounds from ZINC online molecular database, as well as structural analogues of the enzymatic cofactor s-adenosylmethionine (SAM) and a known inhibitor in the literature, sinefungin (SFG), provided at PubChem database. All the evaluated sets presented molecules that deserve attention, highlighting four compounds from ZINC as the most promising ligands. These results contribute to the discovery of new molecular hits, in the search of potential agents against SARS-CoV-2 virus, still unveiling a pathway that can be used in combined therapies.

1. Introduction

An outbreak of a new Severe Acute Respiratory Syndrome hit the city of Wuhan, China, in late 2019. Caused by SARS-CoV-2, COVID-19 has reached global proportions, differing from respiratory epidemics caused by other coronaviruses, such as SARS-CoV (Severe Acute Respiratory Syndrome) and Mers-CoV (Middle East Respiratory Syndrome), by its high transmission potential, resulting so far in the largest pandemic of the 21st century, with more than 4.4 million deaths [1–3]. SARS-CoV-2 is a single-stranded, positively polar RNA virus of the family Coronaviridae [4]. There are seven types of coronaviruses pathogenic to humans: 229E and NL63 (*Alphacoronaviruses*), OC43, HKU1, SARS-CoV, MERS-CoV and SARS-CoV-2 (*Betacoronaviruses*) [5–7]. Structurally, SARS-CoV-2 contains 4 major structural proteins (the nucleocapsid protein, the transmembrane protein, the envelope protein, and the spike protein), 16 nonstructural proteins (NSP), and 9 open reading frame (ORF) sequences [8].

Among the nonstructural proteins (NSP), the NSP16 is one of the most studied. NSP16 is a methyl transferase dependent on the S-adenosylmethionine (SAM) cofactor, which has enhanced activity and greater stability when associated with the nonstructural protein 10 (NSP10). NSP16 is made up of seven stranded β-sheet surrounded by α-helices and loops [9], whereas NSP10 can be described in a general way into three regions: a helical domain at the N terminus followed by an irregular β-sheet region, and a C-terminal loop region. It has also a positively charged and hydrophobic surface that interacts with a hydrophobic pocket and negatively charged surface from NSP16, aiding in the stabilization of the SAM binding site [10]. These two proteins interact with each other through salt bridges, hydrophobic interactions, and a large network of hydrogen bonds, resulting in a stable complex [9].

NSP16-NSP10 dimer is responsible for a step in the capping of messenger RNA (mRNA) [11], which is a mechanism of the virus to evade the innate immune response mechanisms. It involves the insertion of three phosphate groups, a guanosine and a methyl group (methylation, which is carried out by the NSP16-NSP10 dimer) to the 5′ end of the RNA. This modification is responsible for preventing the action of RNAses enzymes on mRNA strand [11,12]. Thus, interfering on the activity of this dimer using molecules that have

* Corresponding author.
E-mail address: edilson.beserra@univasf.edu.br (E.B.d.A. Filho).

https://doi.org/10.1016/j.molstruc.2022.132951
0022-2860/

affinity for the SAM bind site can inhibit the methylation process, thereby making the virus susceptible to the body's immune system. Although SAM is important for various processes in the human body, the dimer is still considered an interesting target by several research groups [7,13-15] due to its virus-fighting potential. It is still a strategy that needs refinement and testing, but the selection of molecules by theoretical methods represents a first and important step in drug development.

Several studies have been alredy conducted to understand the activity of 29 viral proteins [16]. In this sense, understanding the structures, functions and interactions of these viral proteins enables the development of targeted strategies [7]. Computational chemistry has a key role in this process, since it has tools capable of evaluating the pharmacological potential of thousands of molecules through in silico strategies, using the structures of several protein targets of SARS-CoV-2 [17]. Molecular modeling methods aim to build representative models of complex molecular systems, using computational mathematical models [18]. Thus, they save money and time in relation to experimental studies, since it is possible to direct future investigations only with ligands that have a greater chance of success in *in vivo* and *in vitro* tests.

Molecular docking is a computational strategy that involves determining the best conformational pose of a ligand in a region of a target macromolecule, through algorithms and graphical interfaces implemented in computer programs. Screening of molecules against a molecular target of interest is one of the major applications of this type of study [19]. Thus, obtaining the structure of the NSP10-NSP16 protein dimer and understanding that the functionality of the dimer is dependent on a cofactor (SAM), a virtual screening can be performed aiming to find structures of several potential inhibitors available in databases, allowing the graphical visualization of the various ligand-target complexes, the identification of interactions, as well as the understanding of the structure-affinity relationship [9,11,19,20].

One of the best-characterized drug targets among coronaviruses is the main protease (M^{pro}, also called $3CL^{pro}$), because this enzyme is essential for processing the polyproteins that are translated from the viral RNA [5]. However, treatments for viral infections with single drugs have not been successful, as exemplified by human immunodeficiency virus (HIV) and hepatitis C virus (HCV) infections. Combination therapy has led to improved clinical outcomes, since it can enhance therapeutic efficacy through additive, and ideally synergistic effects [21].

Natural products are a promising source for the discovery of new drugs. Besides the great variability of compounds, they generally have good efficacy with tolerable toxicity levels [22]. Caatinga is an exclusively Brazilian phytogeography domain, located in the Northeast region of the country. It has great diversity, with 744 endemic species, thus being an important source of new drug candidate molecules [23]. Ethnobotanical and ethnopharmacological studies reflect this diversity: reports in the literature range from use for gastrointestinal problems like diarrhea, to respiratory problems like asthma, and antiproliferative activity [24–26]. Moreover, in silico studies were recently conducted using a database containing only natural products from the semi-arid region of Bahia (NatProDB), and revealed three (secotrachylobanoic acid, hexahydronaphthalen-1-yl, indolinedione, and benzopyran-4-one) promising scaffolds for Mpro inhibition of SARS-CoV-2 [27].

So, this study aims to search for new ligands that can be potential inhibitors of the NSP16-NSP10 protein dimer, evaluating a homemade databank of molecules from plants of the Caatinga, aleatory database (ZINC) as well as from structural analogs based on the scaffold of its enzyme cofactor (SAM), using virtual screening by docking. We hope to contribute with the studies and the proposition of potential chemotherapeutic alternatives for future tests and treatment of COVID-19.

Table 1
Max-range of different types of intermolecular interactions visualized with Discovery Studio.

Intermolecular Interactions	Max-range (Å)
Hydrogen bonds	3.8
Salt bridges	4.0
Charge-charge	5.6
Pi-Interactions	6.0
Hydrophobic	6.0

2. Methods

2.1. Redockings

Initially, a search for the 3D structure of the NSP10-NSP16 protein dimer of SARS-CoV-2 was performed using the Protein Data Bank virtual platform, obtaining two crystallographic structures of the same dimer: 6W75 (resolution 1.95 Å), bounded with its natural ligand, S-adenosylmethionine (SAM); and the 6WKQ (resolution 1.98 Å), bounded with an inhibitor already known in the literature,

Table 2
Binding energies and respective RMSD from redockings calculations

	Subunit A		Subunit C	
	Energy (kcal/mol)	RMSD	Energy (kcal/mol)	RMSD
SAM (14×16×14)	-8.7	0.365	-8.7	0.276
SFG (16×14×16)	-8.7	0.554	-8.5	0.672

Table 3
Best results for Caatinga secondary metabolites database (kcal/mol), in the NSP16-NSP10 dimer from SARS-CoV-2.

Caatinga database compound	6W75		6WKQ	
	A	C	A	C
15 (5'-hidroxyamentoflavone)	-10.9	-10.6	-10.7	-10.8
98 (amentoflavone)	-8.9	-10.3	-11.0	-10.7

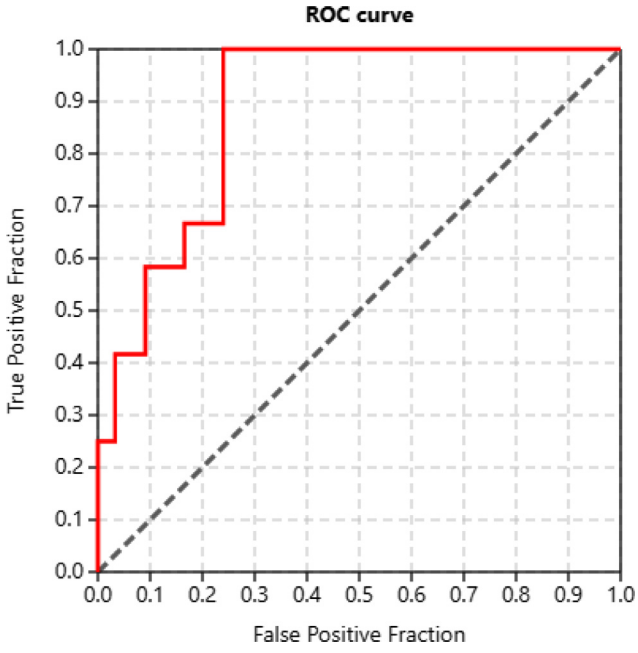
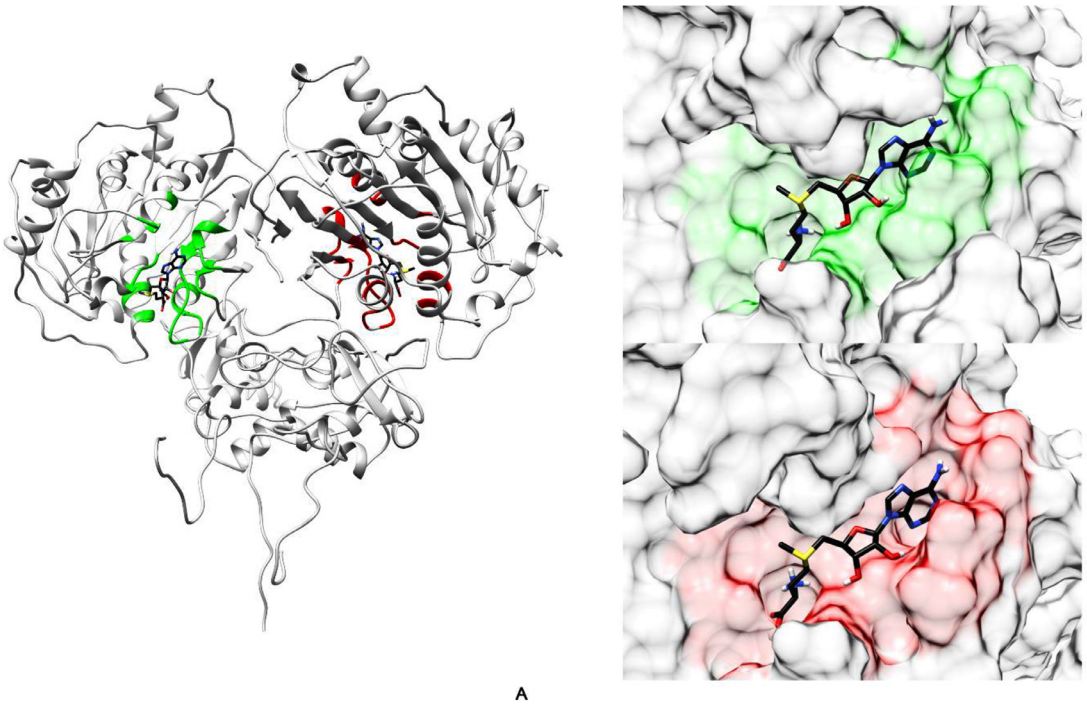
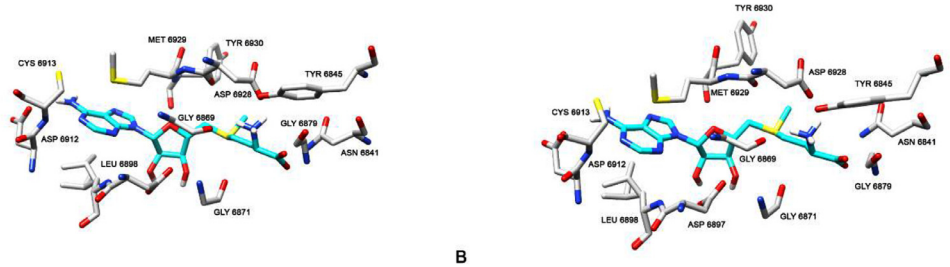


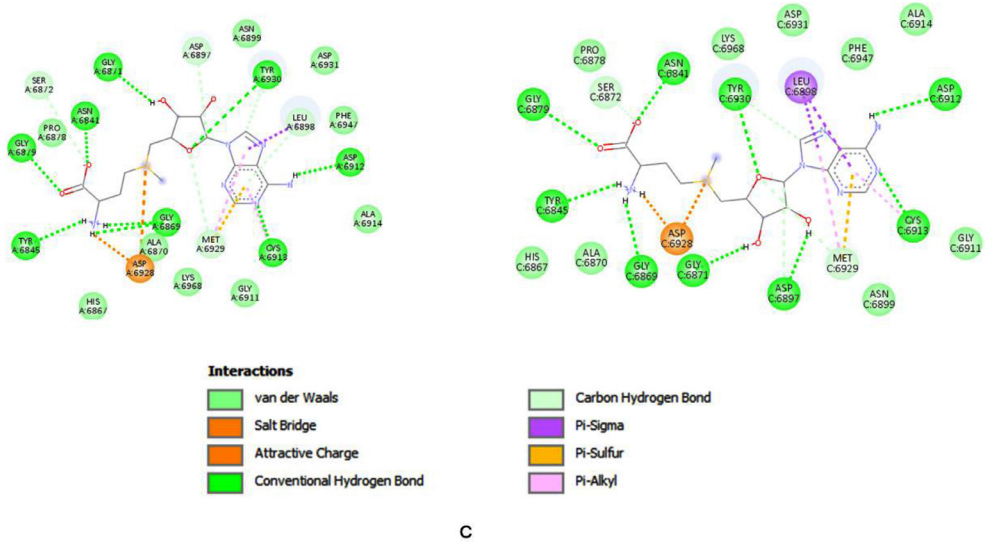
Fig. 1. ROC curve



A

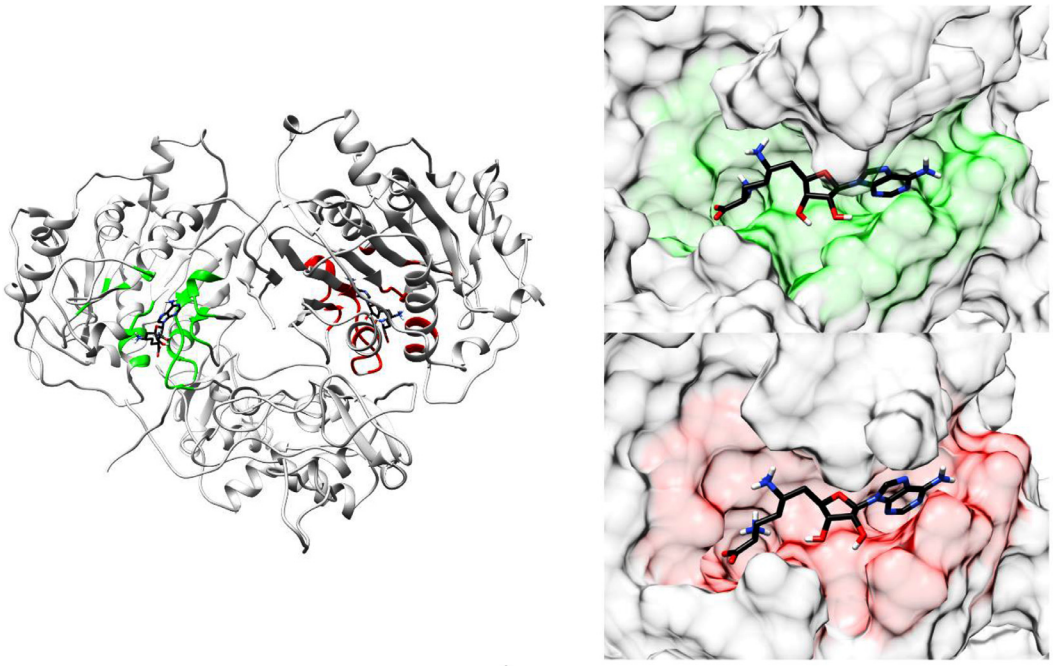


B

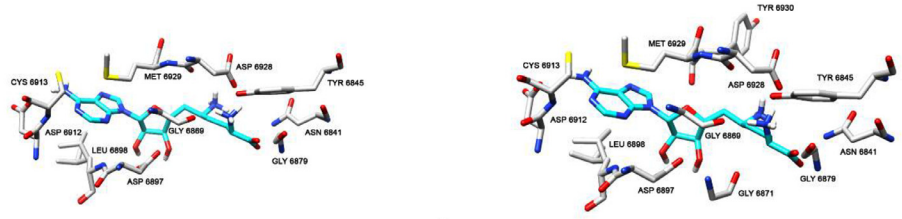


C

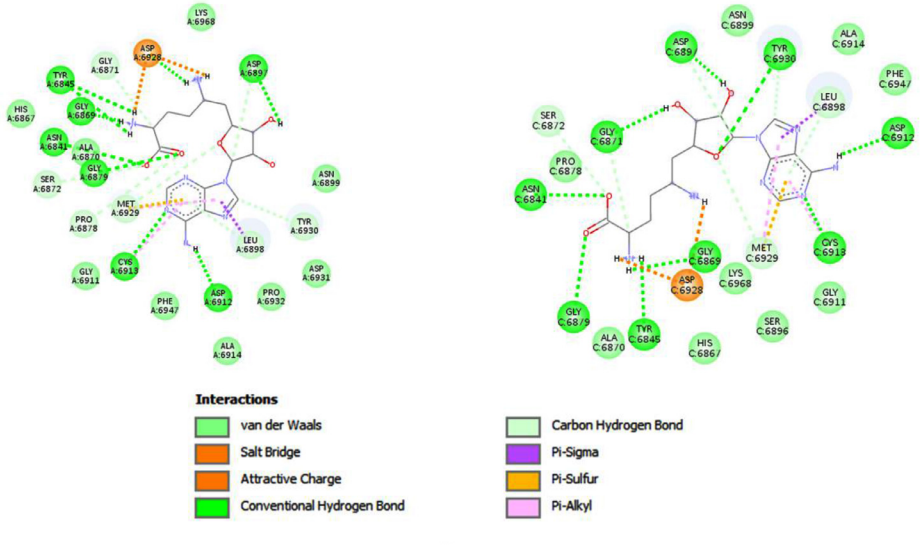
Fig. 2. **A**) Results obtained from redockings of SAM - 6W75 (*the green site refers to the A subunit and the red site refers to the C subunit*); **B)** A and C interaction sites of SAM (6W75), respectively; **C)** 2D diagram of intermolecular interactions between SAM and A and C subunits, respectively, of NSP16-NSP10 protein dimer (6W75).



A



B



C

Fig. 3. A) Results obtained in the redockings of SFG - 6WKQ (*green site is for subunit A and red site is for subunit C*); **B)** SFG (6WKQ) interaction sites A and C, respectively; **C)** 2D diagram of intermolecular interactions between SFG and subunits A and C, respectively, of the protein dimer NSP16-NSP10 (6WKQ).

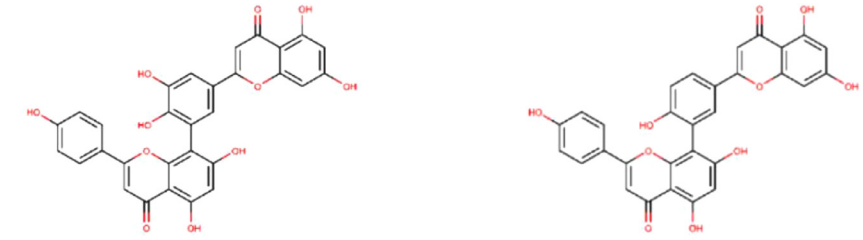


Fig. 4. 2D structures of the best molecules in the Caatinga secondary metabolites database.

Table 4
Best results obtained by dockings from ZINC database (kcal/mol) on the A and C subunits of the NSP16-NSP10 dimer from SARS-CoV-2.

ZINC Identification Code	6W75		6WKQ	
	A	C	A	C
1581896711	-10.8	-10.7	-10.8	-10.3
1581897071	-10.7	-10.4	-10.7	-9.9
1573660102	-10.5	-10.3	-10.4	-9.9
1573660106	-10.5	-10.2	-10.4	-9.8
1606009988	-10.5	-10.3	-10.4	-10
1573660101	-10.4	-10.2	-10.3	-9.4
1592942129	-10.4	-10.1	-10.2	-10
498694188	-10.3	-10.2	-10.3	-10
1573657043	-10.3	-10.2	-10.2	-10
1573660105	-10.3	-10	-10.2	-9.7
1579400860	-10.3	-10	-10.2	-9.7
1579403252	-10.3	-10	-10.2	-9.8
1593194102	-10.3	-10.3	-10.5	-10.2
1573657046	-10.2	-10.1	-10.1	-9.8
1579400626	-10.2	-9.9	-10	-9.6
1579403255	-10.2	-9.9	-10.1	-9.6
1579944224	-10.2	-9.8	-9.9	-9.8
1581896691	-10.2	-10.1	-10.2	-9.8
1582178908	-10.2	-10.1	-10.1	-9.7
1582524624	-10.2	-9.7	-10.1	-9.4
535551028	-10.1	-9.9	-10.0	-9.8
1579310996	-10.1	-9.4	-9.4	-8.8
1579320216	-10.1	-9.8	-9.7	-9.8
1581896585	-10.1	-10.0	-9.1	-9.7
1581896712	-10.1	-9.9	-10.3	-9.9
1581897687	-10.1	-9.5	-10.2	-9.8
1582205261	-10.1	-10.0	-10.0	-9.6
1582524573	-10.1	-9.7	-10.1	-9.5
1582777819	-10.1	-9.8	-10.0	-9.8
677698257	-10.0	-9.6	-10.0	-9.8
1571269484	-10.0	-9.7	-9.6	-8.8
1573658859	-10.0	-9.8	-9.7	-9.7
1579320217	-10.0	-10.0	-9.2	-9.8
1579406074	-10.0	-9.9	-9.8	-9.6
1580329703	-10.0	-9.6	-9.9	-8.1
1581757430	-10.0	-9.5	-9.4	-8.4
1582178909	-10.0	-9.6	-10.0	-9.6
1582730750	-10.0	-9.7	-9.6	-9.0
1582787156	-10.0	-9.7	-10.0	-9.6
1590296524	-10.0	-9.6	-9.9	-9.5
1605416137	-10.0	-9.7	-9.6	-9.3
1606009989	-10.0	-9.8	-10.0	-9.4
1607039176	-10.0	-10.0	-9.1	-9.7

Table 5
Results obtained on dockings in kcal/mol in subunit A of the two crystallographic structures of the NSP16-NSP10 dimer of SARS-CoV-2, selected by energy ranges better or equal to -8.7 kcal/mol.

SAM Analogs (PubChem Identification Code)	6W75		6WKQ	
	A	C	A	C
20592073	-8.8	-8.6	-8.6	-7.4
45268044	-8.6	-8.6	-8.8	-8.6
SFG Analogs (PubChem Identification Code)	**6W75**		**6WKQ**	
	A	**C**	**A**	**C**
446317*	-8.9	-8.7	-9.1	-9.0
24941256	-8.7	-7.8	-8.9	-8.7
46936669*	-8.9	-8.8	-9.0	-9.1
49787001*	-8.7	-8.8	-9.1	-9.1
52921643*	-9.1	-9.0	-9.3	-9.1
91240073	-9.1	-8.9	-8.8	-8.6
100965957*	-8.8	-8.8	-8.8	-8.8
101396128	-8.7	-7.8	-8.9	-8.8
102187413*	-8.9	-8.7	-9.0	-8.9
131854545	-9.1	-9.0	-9.2	-7.7
132561820*	-9.1	-8.9	-9.3	-9.0
132561822	-9.0	-8.9	-9.2	-7.8
132561823	-8.7	-7.9	-8.9	-8.7
152457420*	-9.3	-9.2	-9.5	-9.4
153484384	-8,6	-8.7	-8.8	-8.9

sinefungin (SFG) [15,20]. The co-crystallographic ligands were extracted from original structures and prepared for redocking using AutoDock Tools, in which the conformation and coordinates of the ligands were previously randomized to avoid biased results. After that, the proteins were prepared for docking using Chimera software 1.15rc [21], and the virtual platform ABPS, for the calculation of Poisson-Boltzmann electrostatic analysis, being adjusted to pH 7.4 [22]. The determination of the gridbox values was performed by AutoDock Tools software. Then, redocking calculations were performed for both structures using the AutoDock Vina module [23]. RMSD calculations were performed on Zhang Lab's DockRMSD virtual platform [24].

2.2. Docking validation

For the validation of the method used, a ROC (Receiver Operating Characteristic Curve) curve was constructed using an on-

Table 6
ADMET properties of the best molecules.

	Zinc database					
	Consensus Log P	GI absorption	BBB permeant	PAINS	Synthetic accessibility	Water solubility
1581896711	-0.57	Low	No	0 alert	3.69	Soluble
1581897071	-0.69	Low	No	0 alert	3.72	Soluble
1593194102	-0.06	Low	No	0 alert	3.72	Soluble
1606009988	-0.62	Low	No	0 alert	4.08	Soluble

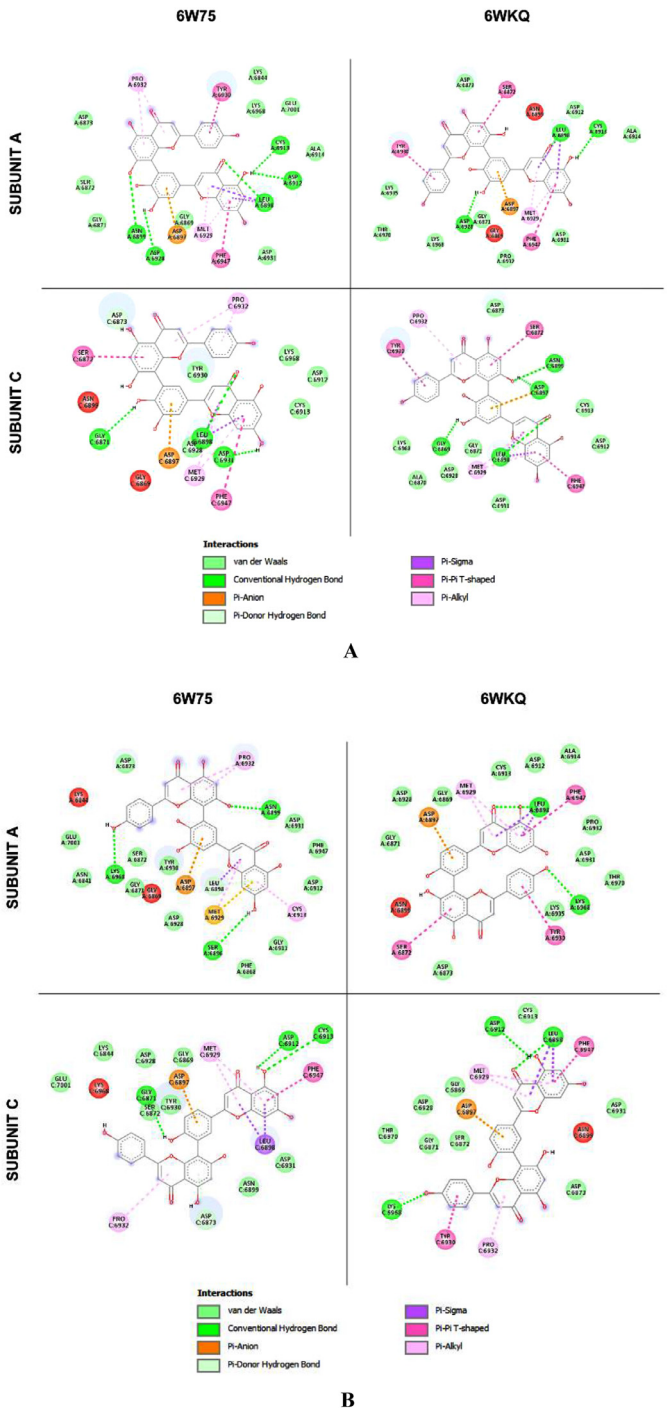


Fig. 5. A) 2D diagram of intermolecular interactions between molecule **15** in the NSP16-NSP10 protein dimer. **B)** Same for molecule **98**.

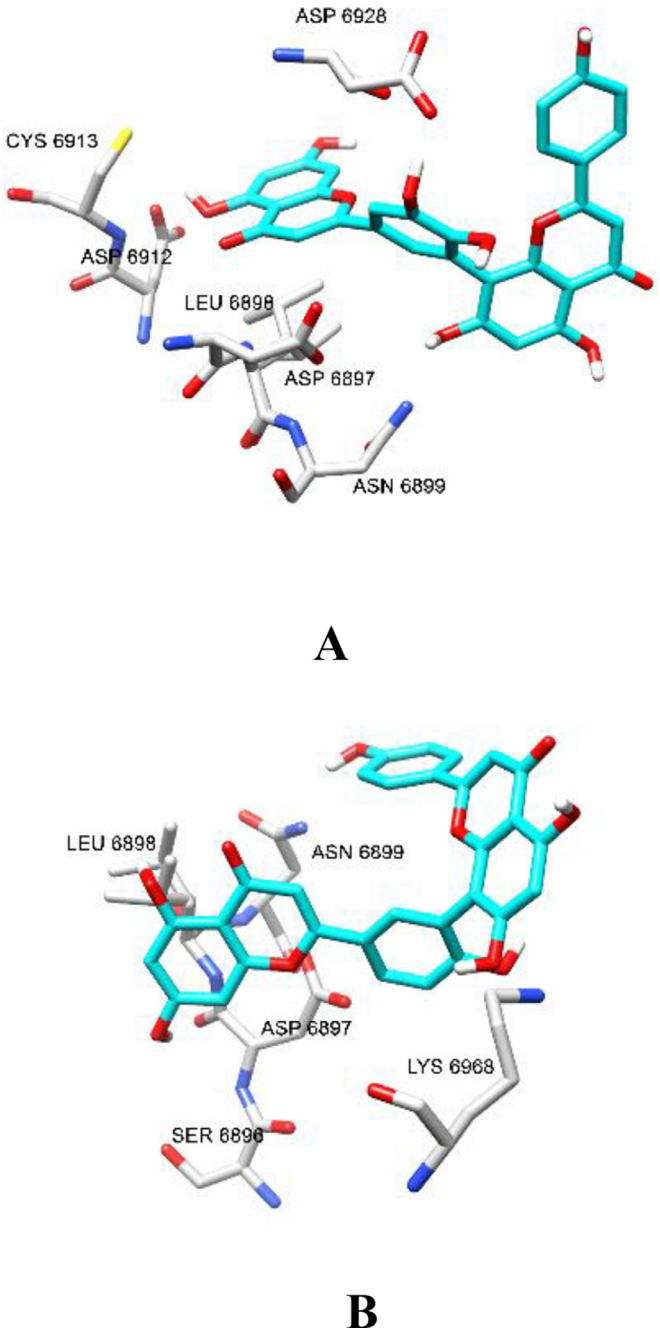


Fig. 6. A) 3D visualization of the binding site (on 6W75, subunit A) with the amino acids that can possibly interact with the molecule **15. B)** Same for molecule **98**.

line platform (http://stats.drugdesign.fr/) [28]. Since it is still a recent topic, not many ligands were found that bind in the site of interest on the NSP16/10 dimer. Therefore, molecules tested in vitro for the SAM binding site on NSP14 of SARS-CoV-2 were also considered. The library of decoys was generated using DUD – E [29], and a total of 12 true positive ligands were used [30–33]. Fifty decoys were generated for each active ligand, and 10 decoys corresponding to each real ligand were randomly chosen (total 120). Then, ROC curve calculations were performed using 132 molecules.

2.3. Database of secondary metabolites from Caatinga

A homemade databank of Caatinga biome was previously generated with 248 molecules, which were obtained through the search for secondary metabolites isolated from plants of this phytogeographic domain, performed with the online platform SciFinder®. The search was done based on the scientific name of 40 species cataloged in the book "Caatinga" [34], using filters for articles resulting from the keywords: "LC-MS", "GC-MS", "secondary metabolites" and "chemical composition", ensuring that only studies with identified and/or isolated molecules were selected. The compounds were drawn with ChemSketch and had their geometry optimized with Hyperchem, at the RM1 semiempirical level. After

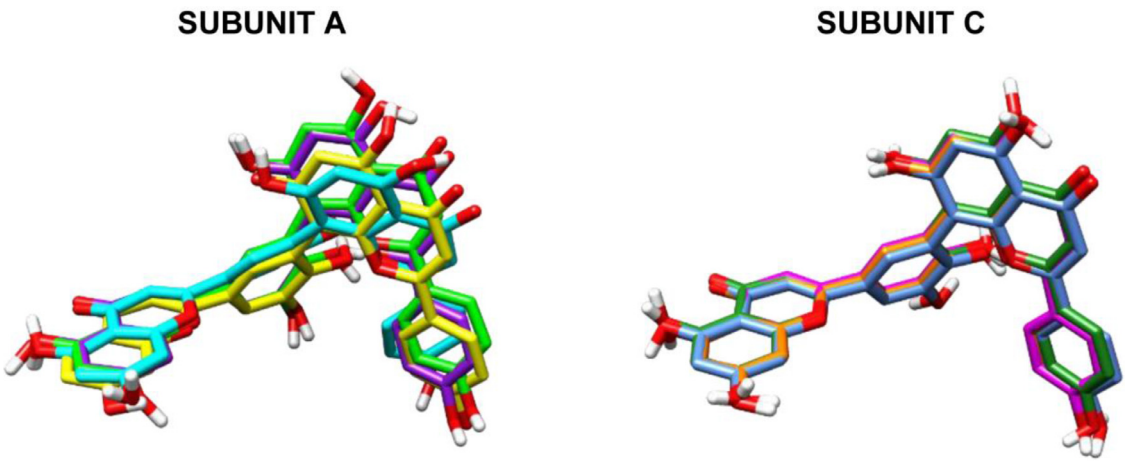


Fig. 7. Comparative overlay of the best results, 15 and 98.

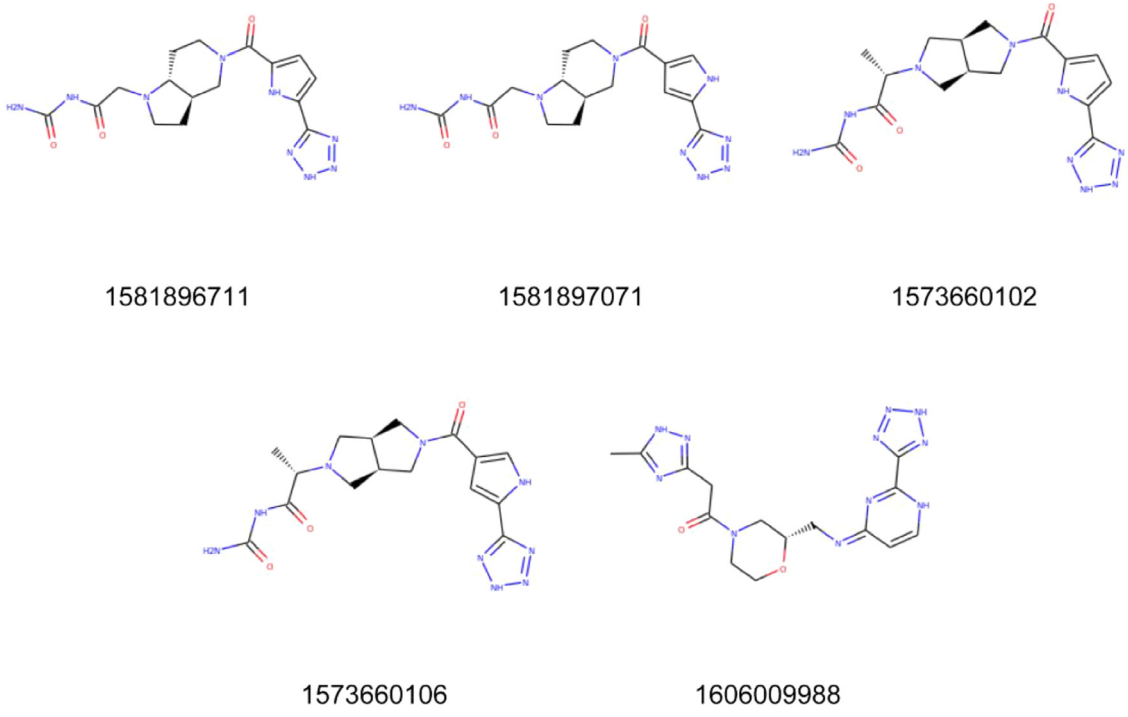


Fig. 8. 2D structures of the best selected molecules after the screening of ZINC database.

geometry optimization, the molecules were converted to '.pdbqt' with the aid of OpenBabel GUI, for docking by AutoDock Vina [35,36].

2.4. Zinc databank

Using the ZINC Database, a search was performed for molecules whose molecular weight and LogP were similar to those of the endogenous coenzyme of the NSP10-NSP16 protein dimer, S-adenosylmethionine (399.4 g/mol and -5.3, respectively), yielding 28,926 structures in this selection. The compounds were downloaded as '.sdf' and converted to '.pdbqt' with OpenBabel GUI [36]. Docking was performed with AutodDock Vina following the same parameters as before [35].

2.5. PubChem databank

A search for related compounds analogous to S-adenosylmethionine was performed in the PubChem molecule database, using Lipinski's rule of five as filter [37], obtaining 711 molecules. Similarly, a search for analogs of sinefungin was performed, obtaining 358 molecules. After this, the molecules of both groups were downloaded in a sdf file. In this case, the molecules were in 2D format, so they were adjusted to the three-dimensional form using ChemSketch. After this adjustment, the molecules were converted to pdbqt. with Open Babel GUI 2.3 for docking [36]. Docking calculations were performed by AutoDock Vina [35].

2.6. Intermolecular interactions visualization

For the visualization of intermolecular interactions between ligand and and macromolecule, the software Discovery Studio was used, which creates a 2D diagram of the possible interactions between the ligand and the amino acids in the biding site. The software considers interactions in different ranges from the ligand, for different types of interactions, as shown in Table 1.

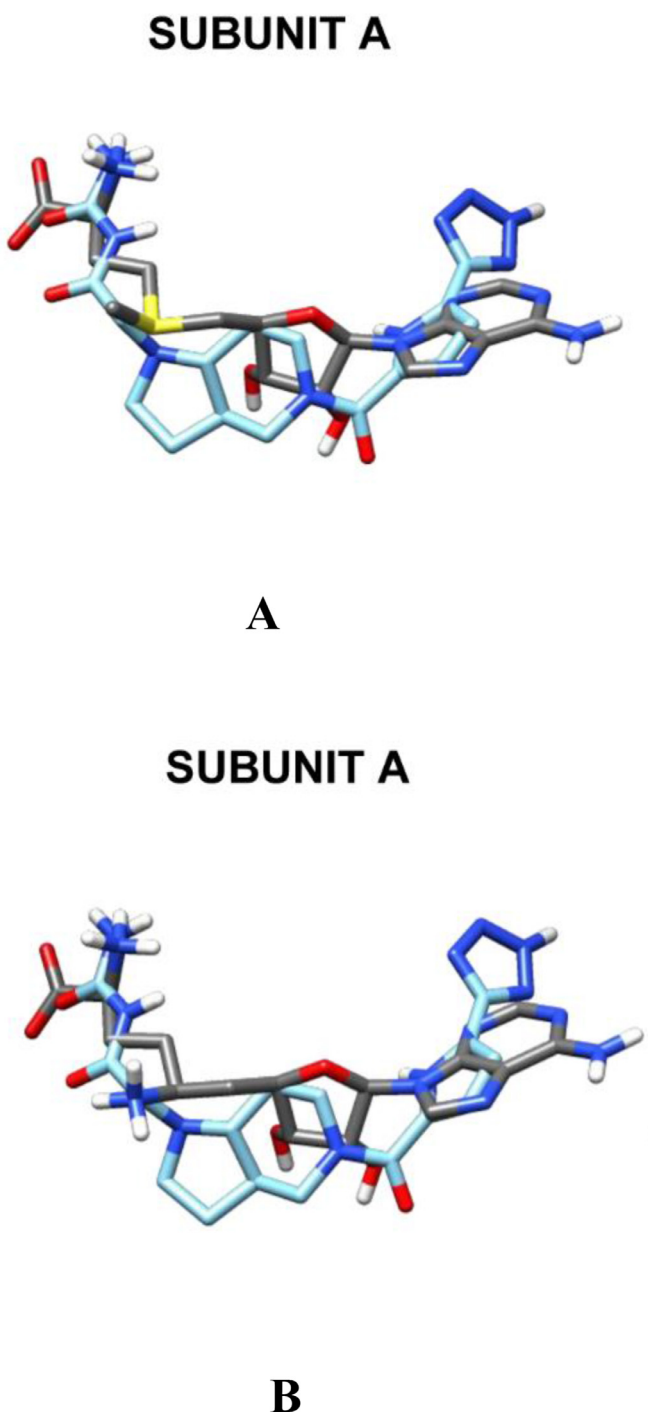


Fig. 9. A) Comparative structures of docked SAM (in gray) with the best screening result **1581896711** in the subunit A of NSP10-16 (PDB ID: 6W75). **B)** The same for PDB ID: 6WKQ.

3. Results and discussion

In this study, we performed a virtual screening using Autodock Vina with the purpose of filtering thousands of molecules in order to find those that bind to NSP10/16. The area under the curve (AUC) is a widely used metric for analyzing the overall performance of virtual screening, because it summarizes in a number whether the screening methodology can differentiate by scoring whether compounds are active or not [38,39]. In our study, AUC=0.886, indicating that Autodock Vina performed satisfactorily in the ranking of the compounds, that is, experimentally active molecules with good IC_{50} (for example, IC_{50}= 0.3+-0.05 and 0.008+-0.0006) [31] had good energies (-8.9 and -8.8 kcal/mol, respectively). Global metrics of this curve also allows us to infer that the early recognition of active compounds was very satisfying (BEDROC = 0.576) Fig. 1.

As a second part of docking validation, redocking analyses was performed with S-adenosylmethionine (SAM) and Sinefungin (SFG). Energy values and RMSD indicates the successful of the algorithm to study this system (Table 2).

We can observe in Fig. 2C that SAM interacts mostly via hydrogen bonds in the two homologue sites (9, in both analogous subunits of the protein dimer) with the following amino acids: TYR 6845, GLY 6869, GLY 6879, ASN 6841, GLY 6871, TYR 6930, CYS 6913, ASP 6912 differing only in the hydrogen bond of one of the hydroxyls of the five-membered cycle, which can interact with ASP 6897 only in the C subunit. The protonated amine can interact ionically with an asparagine (ASP 6928), which also shows a possible ionic interaction with the positive sulfur in both subunits (A and C).

Fig. 3C shows the main interactions of the sinefungin (SFG) with the receptor. SFG can mostly make hydrogen bonds (9, in both analogous subunits of the protein dimer) with the following amino acids: TYR 6845, GLY 6869, GLY 6879, ASN 6841, ASP 6897, TYR 6930, CYS 6913, ASP 6912, differing solely in the hydrogen bond with ASP 6928 only in subunit A and a hydrogen bond with GLY 6871 in subunit C. The protonated amine can ionically interact with an asparagine (ASP 6928), which also shows a possible ionic interaction with positive sulfur in both subunits (A and C). It can be seen that the amino acids that are responsible for the interaction between ligand (either SAM or SGF) and macromolecule are the same (Figs. 2 and 3). Thus, a promising molecule should probably interact with the same amino acids listed above.

After the virtual screening by docking with the 248 molecules from the Caatinga database, two compounds were selected as most promisor results (filtered by energy less than or equal to -8.7 kcal/mol): **15** (5′-hydroxyflavone) and **98** (amentoflavone), as we can see in Fig. 4. Both are secondary metabolites extracted from *Caesalpinia pyramidalis*, popularly known as Catingueira [34,40]. Table 3 lists the results obtained by dockings.

It is noteworthy that some of the interactions between molecules **15** and **98** with the NSP10-NSP16 dimer occur between the same amino acids that interact with the crystallographic coenzyme (Figs. 2C, 5A and B), an indication that both are promising for further experimental study. The same is seen when compared to the known antagonist, SFG (Figs. 2C, 5A, and B). Therefore, it is noticeable that there was no sudden variation in the conformations of the structures, which shows a pattern in the way the molecules stabilize in the site (Figs. 6 and 7).

Regarding the ZINC database, since the number of molecules was too large, docking was performed on the A subunit of the crystallography code 6W75 of the protein dimer NSP10-NSP16. After screening with the 28,926 molecules, only those that had energy better than or equal to -10.0 kcal/mol were analyzed at the C-subunit site of 6W75, as well as at both sites of 6WKQ. Results are shown in Table 4 and the top 5 molecules are shown in Fig. 8.

It is noteworthy that the arrangement of the top 3 molecules in the binding site was comparable to the coenzyme ligand (Fig. 9), with the possibility of hydrogen bond formation at the same amino acids (Figs. 2C, and 10A-D), an indication that it is a promising analog for further pharmacological testing. The same is perceived when compared to the known antagonist, SFG (Figs. 3C and 10A-D). Despite the good interaction profile, most of the best molecules present more than 10 hydrogen bond acceptors in their structure,

Fig. 10. A) 2D diagram of intermolecular interactions between the 1581896711 from ZINC in the NSP16-NSP10 protein dimer. B), C) and D) the same respectively for 1581897071, 1593194102, 1606009988.

which violates Lipinski's rule of five [37]. However, bioisosteric strategies can be used to exploit the molecular bases found, by modifying them through substitutions that do not interfere with the protein affinity profile Fig. 11.

Regarding the search for SAM and SFG analogs in PubChem, after screening with the 1069 molecules, molecules with energy better or equal to the energy obtained in the redocking were selected (-8.7 kcal/mol). For SAM, the selected analogs were **20592073** and **45268044**, both with the best energy of -8.8 kcal/mol. In addition, for the SFG analogs, 25 results were initially selected. All molecules whose isomers were not explicit in the database were excluded, resulting in 17 results (2 SAM analogs and 15 SFG analogs, illustrated in 2D in Fig. 12).

Table 5 lists the results obtained after filtering. The analogs highlighted with an asterisk are those whose energies at all

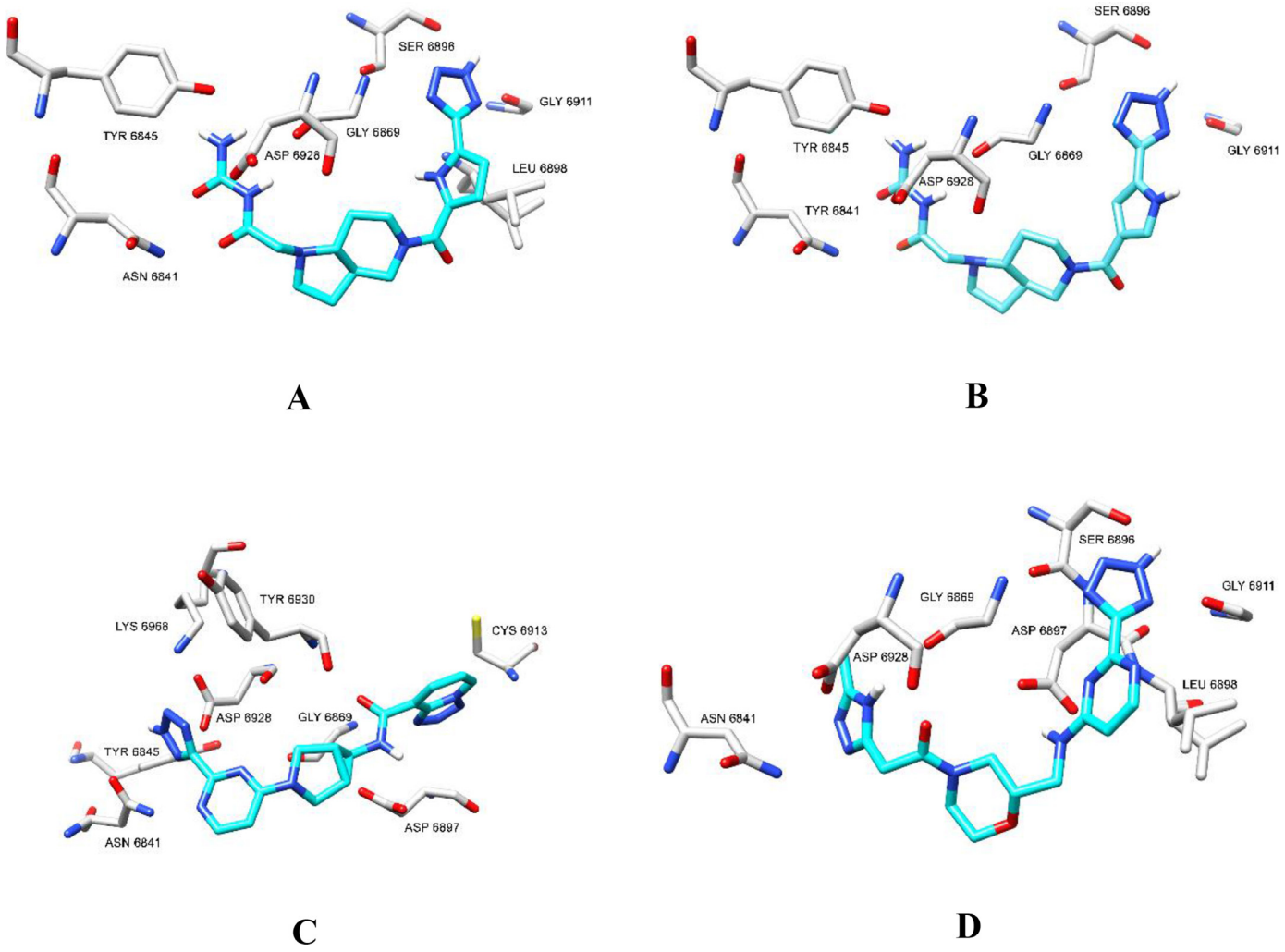


Fig. 11. A) 3D visualization of the binding site (on 6W75, subunit A) with the amino acids that can possibly interact with the molecule. B), C) and D) the same respectively for 1581897071, 1593194102 and 1606009988.

sites were equal to or above -8.7 kcal/mol, and the molecules **152457420, 52921643**, and **132561820** had the best results within this group. Molecules with nitrogen-substituted heterocycles have the most diverse health applications, having already presented anti-cancer [41] anti-neurodegenerative disease [42,43] activities, and furthermore, are commonly used to combat viruses and bacteria [44–47]. The presence of the fluorine atom in molecule **152457420** certainly influenced its better energy compared to the others, which have a similar structure. Some studies show that a fluorine substitution is an approach that is being increasingly used to improve the interaction profile of the ligand [48,49].

It is noted that the arrangement of the 3 best molecules in the binding site was comparable to the native ligand (Fig. 13) making hydrogen bonds at the same amino acids (Fig. 2C), an indication that it is a promising analog for further pharmacological testing (Figs. 14 and 15). The same is observed when compared to the known antagonist, SFG (Fig. 3C).

For the best molecules in this study, all from the ZINC database, ADMET calculations were conducted using the SwissADMET online platform (http://www.swissadme.ch), and some parameters were analyzed. Log P is a classic descriptor to represent the lipophilicity of a substance, and the consensus of the Log P represents the arithmetic mean of the value obtained by 5 methods (XLOP3, WLOGP, MLOGP, SILICOS-IT, iLOGP). It has been a common practice performed in order to increase the accuracy of simulations [50].

Two other characteristics that are crucial to understanding the probable pharmacokinetic behavior are the ability to be passively absorbed from the gastrointestinal (GI) tract and to cross the blood-brain barrier (BBB), which were also evaluated. PAINS (pan assay interference compounds) refers to the presence of structural alerts (frequent hitters or promiscuous compounds) [51] Water solubility is an important property, since it facilitates and guides drug development activities, so, was also evaluated. SwissADMET uses three predictors for this calculation [50]. Finally, synthetic accessibility is a score that indicates how easily that compound can be obtained by synthesis [52].

As shown in Table 6, the best energy compounds are polar, can be synthesized with some ease, without structural alerts, and do not cross the blood-brain barrier, characteristics that are considered positive for the purpose that the molecules are being studied. Due to the low level of gastrointestinal absorption, studies to test an adequate drug delivery should be performed, or even molecular modifications that do not interfere significantly in its interaction with the protein. In addition, it has good water solubility acording to all three predictions used by the platform. So, one can infer that besides having a good activity for the target, the molecules also

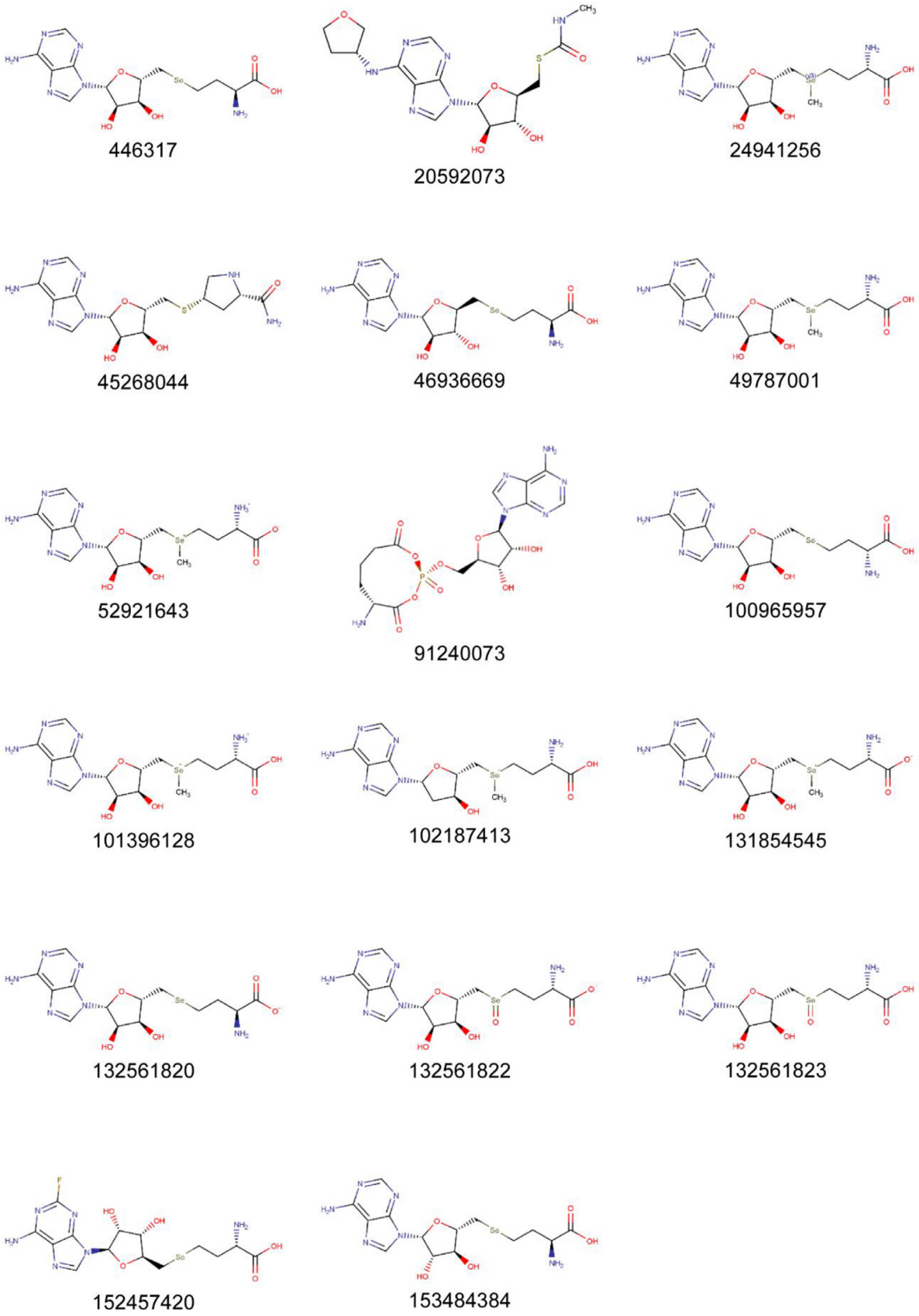


Fig. 12. 2D structures of the selected molecules after docking of SAM and SFG analogs.

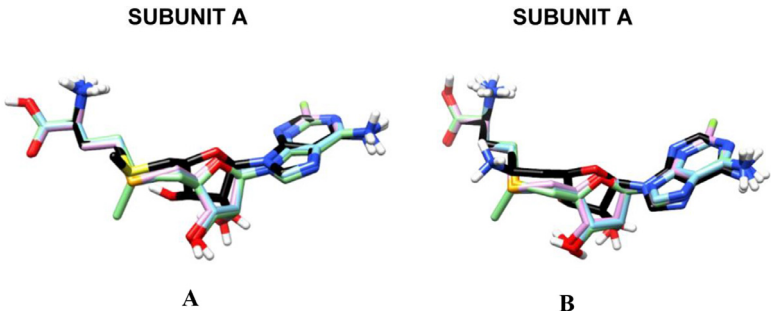


Fig. 13. A) Comparative SAM overlay (in black) with the top three results, 132561820, 152457420, and 52921643 (6W75). B) Comparative overlay of SFG (in black) with the top three results, 132561820, 152457420, and 52921643 (6WKQ).

Fig. 14. A) 2D diagram of intermolecular interactions between the molecule 52921643 and the subunit A of the protein dimer NSP16-NSP10. B) Same for 132561820. C) Same for 152457420.

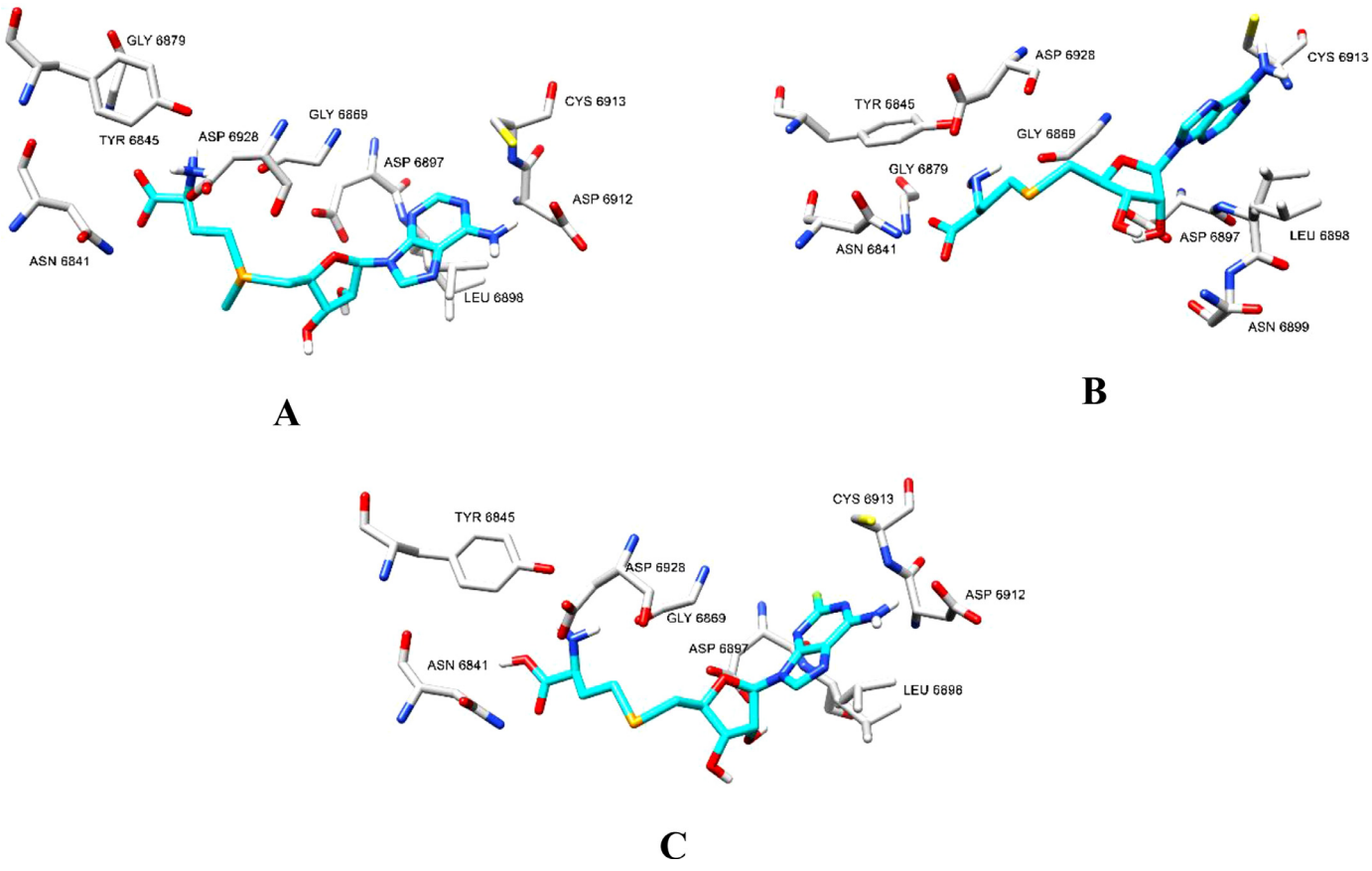


Fig. 15. A) 3D visualization of the binding site (on 6W75, subunit A) with the amino acids that can possibly interact with the molecule. B) Same for 132561820. C) Same for 152457420.

have some ADMET parameters that favor future studies for drug development.

4. Conclusions

The virtual screening procedure allowed the identification of molecules analogous to S-adenosylmethionine (enzyme cofactor) and sinefungin (inhibitor described in the literature), among other molecular scaffolds, with high affinity for the NSP16-NSP10 protein dimer receptor of SARS-CoV-2. Future pharmacological assays are encouraged, starting with the results obtained with ZINC database, which showed better results, such as better affinity with the macromolecule, do not cross the blood-brain barrier and can be easily synthesized. Secondary metabolities from Caatinga Brazilian Biome also showed interesting affinity results. This study, therefore, contributes to the discovery of potential *hits* for the control of infectious processes caused by SARS-CoV-2, including the use of combination therapies with other compounds already reported to date.

Authors' statement

João P Agra-Gomes: Investigation, Methodology, Writing - Review & Editing; Larissa O Rocha: Investigation, Methodology, Writing; Cíntia EY Leal: Methodology, Writing - Review & Editing; Edilson B Alencar Filho: Conceptualization, Investigation, Methodology, Writing - Original Draft preparation.

Ethical statement/conflict of interest

The authors state that research ethical precepts have been observed and that there is no conflict of interest.

Declaration of Competing Interest

None.

Acknowledgments

The authors are grateful to the Foundation of Support to the Science and Technology of the State of Pernambuco – FACEPE (APQ-0423-1.06/14), National Council for Scientific and Technological Development – CNPQ (financial support to students), Coordination of Improvement of Higher Level Personnel – CAPES (financial for Master degree student), Federal University of São Francisco Valley - UNIVASF, for financial support, and National High-performance Processing Center of Federal University of Ceará - CENAPAD-UFC for some computational facilities. Molecular graphics and analyses performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311.

References

[1] C. Sohrabi, et al., Impact of the coronavirus (COVID-19) pandemic on scientific research and implications for clinical academic training – a review, Int. J. Surg. 86 (Feb. 2021) 57–63, doi:10.1016/J.IJSU.2020.12.008.
[2] D. Huremović, Brief history of pandemics (pandemics throughout history), Psychiatry Pandem. (2019) 7–35, doi:10.1007/978-3-030-15346-5_2.
[3] Johns Hopkins UniversityCOVID-19 Map - Johns Hopkins Coronavirus Resource Center, Johns Hopkins Coronavirus Resource Center, 2021.
[4] M. Hasöksüz, S. Kiliç, F. Saraç, Coronaviruses and sars-cov-2, Turk. J. Med. Sci. 50 (SI-1) (2020) 549–556, doi:10.3906/sag-2004-127.
[5] L. Zhang, et al., Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved a-ketoamide inhibitors, Science 368 (6489) (2020) 409–412, doi:10.1126/science.abb3405.

[6] S.R. Weiss, S. Navas-Martin, Coronavirus pathogenesis and the emerging pathogen severe acute respiratory syndrome coronavirus, Microbiol. Mol. Biol. Rev. 69 (4) (2005) 635–664, doi:10.1128/mmbr.69.4.635-664.2005.

[7] N. Vithani, et al., SARS-CoV-2 Nsp16 activation mechanism and a cryptic pocket with pan-coronavirus antiviral potential, Biophys. J. 120 (14) (Jul. 2021) 2880–2889, doi:10.1016/J.BPJ.2021.03.024.

[8] H. Yapici-Eser, Y.E. Koroglu, O. Oztop-Cakmak, O. Keskin, A. Gursoy, Y. Gursoy-Ozdemir, Neuropsychiatric symptoms of COVID-19 explained by SARS-CoV-2 proteins' mimicry of human protein interactions, Front. Hum. Neurosci. 15 (March) (2021) 1–16, doi:10.3389/fnhum.2021.656313.

[9] Y. Chen, et al., Biochemical and structural insights into the mechanisms of SARS coronavirus RNA Ribose 2′-O-Methylation by nsp16/nsp10 protein complex, PLoS Pathog. 7 (10) (Oct. 2011) e1002294, doi:10.1371/journal.ppat.1002294.

[10] M. Rosas Lemus et al., The crystal structure of nsp10-nsp16 heterodimer from SARS-CoV-2 in complex with S-adenosylmethionine, bioRxiv. 2020.04.17.047498 (2020), doi:10.1101/2020.04.17.047498.

[11] S. Perveen, et al., A high-throughput RNA displacement assay for screening SARS-CoV-2 nsp10-nsp16 complex toward developing therapeutics for COVID-19, SLAS Discov. 26 (5) (2021) 620–627, doi:10.1177/2472555220985040.

[12] E. Decroly, F. Ferron, J. Lescar, B. Canard, Conventional and unconventional mechanisms for capping viral mRNA, Nat. Rev. Microbiol. 10 (1) (Jan. 2012) 51–65, doi:10.1038/nrmicro2675.

[13] P. Morales, et al., Interfering with mRNA Methylation by the 2′O-Methyltransferase (NSP16) from SARS-CoV-2 to Tackle the COVID-19 Disease, Catalysts 10 (9) (2020) 1023, doi:10.3390/catal10091023.

[14] S.H. Seyedi, M.S. Alhagh, M. Ahmadizad, N. Ardalan, E.H. Koushki, C. Farshadfar, Structural screening into the recognition of a potent inhibitor against non-structural protein 16 : a molecular simulation to inhibit SARS-CoV-2 infection, J. Biomol. Struct. Dyn. 0 (0) (2021) 1–16, doi:10.1080/07391102.2021.2001374.

[15] A.K. Maurya, N. Mishra, In silico validation of coumarin derivatives as potential inhibitors against Main Protease, NSP10 /NSP16-Methyltransferase, Phosphatase and Endoribonuclease of SARS CoV-2, J. Biomol. Struct. Dyn. 0 (0) (2020) 1–16, doi:10.1080/07391102.2020.1808075.

[16] Y. Huang, C. Yang, X. Xu, W. Xu, S. Liu, Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19, Acta Pharmacologica Sinica 2020 41:9 vol. 41 (9) (Aug. 2020) 1141–1149, doi:10.1038/s41401-020-0485-4.

[17] M. Yadav, S. Dhagat, J.S. Eswari, Emerging strategies on in silico drug development against COVID-19: challenges and opportunities, Eur. J. Pharm. Sci. 155 (Dec. 2020) 105522, doi:10.1016/J.EJPS.2020.105522.

[18] R.V.C. Guido, A.D. Andricopulo, G. Oliva, Planejamento de fármacos, biotecnologia e química medicinal: aplicações em doenças infecciosas, Estudos Avançados 24 (70) (2010) 81–98, doi:10.1590/S0103-40142010000300006.

[19] D.B. Kitchen, H. Decornez, J.R. Furr, J. Bajorath, Docking and scoring in virtual screening for drug discovery: Methods and applications, Nat. Rev. Drug Discov. 3 (11) (2004) 935–949, doi:10.1038/nrd1549.

[20] R.T. Borchardt, L.E. Eiden, B.S. Wu, C.O. Rutledge, Sinefungin, a potent inhibitor of S-adenosylmethionine: Protein O-methyltransferase, Biochem. Biophys. Res. Commun. 89 (3) (Aug. 1979) 919–924, doi:10.1016/0006-291X(79)91866-7.

[21] Z.A. Shyr, Y.-S. Cheng, D.C. Lo, W. Zheng, Drug combination therapy for emerging viral diseases, Drug Discov. Today 26 (10) (Oct. 2021) 2367–2376, doi:10.1016/j.drudis.2021.05.008.

[22] H. Yuan, Q. Ma, L. Ye, G. Piao, The traditional medicine and modern medicine from natural products, Molecules 21 (5) (Apr. 2016) 559, doi:10.3390/molecules21050559.

[23] R.C. Forzza, et al., *Catálogo de plantas e fungos do Brasil - Vol. 1*, vol. 2, Instituto de Pesquisa Jardim Botânico do Rio de Janeiro (2010), doi:10.7476/9788560035083.

[24] J.G. De Melo et al., Antiproliferative Activity, Antioxidant Capacity and Tannin Content in Plants of Semi-Arid Northeastern Brazil, Molecules 15 (12) (2010) 8534-42, doi:10.3390/molecules15128534

[25] A. A. Roque, R. M Rocha, M.I.B Loiola, Uso e diversidade de plantas medicinais da Caatinga na comunidade rural de Laginhas, município de Caicó, Rio Grande do Norte (nordeste do Brasil), Rev. Bras. Plantas Med. 12 (1) (2010), doi:10.1590/S1516-05722010000100006.

[26] C.G. Silva, M.G.V. Marinho, M.F.A. Lucena, G.M. Costa, Levantamento etnobotânico de plantas medicinais em área de Caatinga na comunidade do Sítio Nazaré, município de Milagres, Ceará, Brasil, Rev. Bras. Plantas Med. 17 (1) (2015), doi:10.1590/1983-084X/12_055.

[27] R.C. Silva, H.F. Freitas, J.M. Campos, N.M. Kimani, C.H.T.P. Silva, R.S. Borges, S.S.R. Pita, C.B.R. Santos, Natural Products-Based Drug Design against SARS-CoV-2 Mpro 3CLpro, Int. J. Mol. Sci. 22 (21) (2021) 11739, doi:10.3390/ijms222111739.

[28] C. Empereur-mot, J. Zagury, M. Montes, Screening Explorer – An interactive tool for the analysis of screening results Screening Explorer – An Interactive Tool for the Analysis of Screening Results, J. Chem. Inf. Mod. 56 (12) (2016) 2281–2286, doi:10.1021/acs.jcim.6b00283.

[29] M.M. Mysinger, M. Carchia, J.J. Irwin, B.K. Shoichet, Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking, J. Med. Chem. 55 (14) (2012) 6582–6594, doi:10.1021/jm300687e.

[30] C. Debarnot, I. Imbert, B. Selisko, E.J. Snijder, B. Canard, In Vitro Reconstitution of SARS-Coronavirus mRNA Cap Methylation, PLOS Pathog. 6 (4) (2010) e1000863, doi:10.1371/journal.ppat.1000863.

[31] R. Nencka, J. Silhan, M. Klima, T. Otava, H. Kocek, P. Krafcikova, E. Boura, Critical Reviews and Perspectives Coronaviral RNA-methyltransferases: function structure and inhibition, Nucleic Acids Res 50 (2) (2022) 635–650, doi:10.1093/nar/gkab1279.

[32] T.Otava, M. Šála, F. Li, J. Fanfrlík, K. Devkota, S. Perveen, I. Chau, P. Pakarian, P. Hobza, M. Vedadi, E. Boura, R. Nencka, The Structure-Based Design of SARS-CoV 2 nsp14 Methyltransferase Ligands Yields Nanomolar Inhibitors, ACS Infec. Dis. 7 (8) (2021) 2214-2220, doi:10.1021/acsinfecdis.1c00131.

[33] O. Bobiļeva, R. Bobrovs, I. Kaņepe, L. Patetko, G. Kalniņš, M. Šišovs, A.L. Bula, S. Grīnberga, M. Borodušķis, A. Ramata-Stunda, N. Rostoks, A. Jirgensons, K. Tārs, K. Jaudzems, Potent SARS-CoV 2 mRNA Cap Methyltransferase Inhibitors by Bioisosteric Replacement of Methionine in SAM Cosubstrate, ACS Med. Chem. Lett. 12 (7) (2021) 10–15, doi:10.1021/acsmedchemlett.1c00140.

[34] G.N. Maia, Caatinga : árvores e arbustos e suas utilidades, 2a ed., Printcolor Gráfica e Editora, Fortaleza, 2012.

[35] 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 (2) (2009) p. NA-NA, doi:10.1002/jcc.21334.

[36] N.M. O'Boyle, M. Banck, C.A. James, C. Morley, T. Vandermeersch, G.R. Hutchison, Open babel: an open chemical toolbox, J. Cheminformat. 3 (10) (Oct. 2011) 1–14, doi:10.1186/1758-2946-3-33.

[37] C.A. Lipinski, Lead- and drug-like compounds: The rule-of-five revolution, Drug Discov. Today 1 (4) (Dec. 2004) 337–341 Drug Discov Today Technol, doi:10.1016/j.ddtec.2004.11.007.

[38] E. Habib, B. Maia, L.C. Assis, T.A. De Oliveira, A. Marques, A.G. Taranto, Structure-Based Virtual Screening : From Classical to Artificial Intelligence, Front. Chem. 8 (343) (2020), doi:10.3389/fchem.2020.00343.

[39] N. Triballeau, F. Acher, I. Brabet, J. Pin, H. Bertrand, Virtual Screening Workflow Development Guided by the ' Receiver Operating Characteristic ' Curve Approach. Application to High-Throughput Docking on Metabotropic Glutamate Receptor Subtype 4, J. Med. Chem. 48 (7) (2005) 2534–2547, doi:10.1021/jm049092j.

[40] M.V. Bahia, J.P. David, J.M. David, Occurrence of biflavones in leaves of Caesalpinia pyramidalis specimens, Química Nova 33 (6) (2010) 1297–1300, doi:10.1590/S0100-40422010000600015.

[41] M.R. Aouad, M.A. Almehmadi, N. Rezki, F.F. Al-blewi, M. Messali, I. Ali, Design, click synthesis, anticancer screening and docking studies of novel benzothiazole-1,2,3-triazoles appended with some bioactive benzofused heterocycles, J. Mol. Struct. 1188 (2019) 153–164, doi:10.1016/j.molstruc.2019.04.005.

[42] P. Sridhar, M. Alagumuthu, B. Ram, S. Arumugam, S.R. Reddy, Drugs against neurodegenerative diseases: design and synthesis of 6-amino–substituted Imidazo[1,2-b]pyridazines as acetylcholinesterase inhibitors, ChemistrySelect 2 (2) (2017) 842–847, doi:10.1002/slct.201601353.

[43] R.K. Sharma, M. Singh, K. Ghimeray, P. Juneja, G. Dev, S. Pulavarthi, S.R. Reddy, R.S. Akundi, Imidazopyridazine Acetylcholinesterase Inhibitors Display Potent Anti-Proliferative Effects in the Human Neuroblastoma Cell-Line, IMR-32, Mol 26 (17) (2021) 5319–5338, doi:10.3390/molecules26175319.

[44] S. Sudharsana, et al., Molecular docking and simulation studies of 3-(1-chloropiperidin-4-yl)-6-fluoro benzisoxazole 2 against VP26 and VP28 proteins of white spot syndrome virus, J. Fish Dis. 39 (10) (2016) 1231–1238, doi:10.1111/jfd.12454.

[45] G. Grace Victoria, S. Rajasekhara Reddy, Recent advances in the synthesis of organic chloramines and their insights into health care, New J. Chem. 45 (19) (2021) 8386–8408, doi:10.1039/d1nj01086g.

[46] C.B. Rajashekar Reddy, S. Dinesh, N. Anusha, T. Itami, S. Rajasekhara Reddy, R. Sudhakaran, Antiviral activity of 3-(1-chloropiperidin-4-yl)-6-fluoro benzisoxazole 2 against White spot syndrome virus in Freshwater crab, Paratelphusa hydrodomous, Aquacult. Res. 47 (8) (2016) 2677–2681, doi:10.1111/are.12704.

[47] P. Sridhar, M. Alagumuthu, S. Arumugam, S.R. Reddy, Synthesis of quinoline acetohydrazide-hydrazone derivatives evaluated as DNA gyrase inhibitors and potent antimicrobial agents, RSC Adv. 6 (69) (2016) 64460–64468, doi:10.1039/c6ra09891f.

[48] E.P. Gillis, K.J. Eastman, M.D. Hill, D.J. Donnelly, N.A. Meanwell, Applications of Fluorine in Medicinal Chemistry, J. Med. Chem. 58 (21) (Nov. 2015) 8315–8359, doi:10.1021/acs.jmedchem.5b00258.

[49] W.K. Hagmann, The Many Roles for Fluorine in Medicinal Chemistry, J. Med. Chem. 51 (15) (Aug. 2008) 4359–4369, doi:10.1021/jm800219f.

[50] A. Daina, O. Michielin, V. Zoete, SwissADME : a free web tool to evaluate pharmacokinetics, drug- likeness and medicinal chemistry friendliness of small molecules, Nat. Publish. Gr. (October 2016) (2017) 1–13, doi:10.1038/srep42717.

[51] J.B. Baell, G.A. Holloway, New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays, J. Med. Chem. 53 (7) (Apr. 2010) 2719–2740, doi:10.1021/JM901137J/SUPPL_FILE/JM901137J_SI_001.PDF.

[52] P. Ertl, A. Schuffenhauer, Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions, J. Cheminfo. 11 (2009) 1–11, doi:10.1186/1758-2946-1-8.