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.

Computers in Biology and Medicine 133 (2021) 104364

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

Identification of potential antiviral compounds against SARS-CoV-2 structural and non structural protein targets: A pharmacoinformatics study of the CAS COVID-19 dataset

Rolando García [a,*], Anas Hussain [g], Prasad Koduru [a], Murat Atis [b], Kathleen Wilson [a], Jason Y. Park [a,e,f], Inimary Toby [d], Kimberly Diwa [d], Lavang Vu [d], Samuel Ho [j], Fajar Adnan [i], Ashley Nguyen [d], Andrew Cox [b], Timothy Kirtek [a], Patricia García [h], Yanhui Li [c], Heather Jones [a], Guanglu Shi [a], Allen Green [a], David Rosenbaum [a]

[a] Department of Pathology, University of Texas Southwestern Medical Center, Dallas, TX, USA
[b] Lyda Hill Department of Bioinformatics, University of Texas Southwestern Medical Center, Dallas, TX, USA
[c] Department of Cell Biology, University of Texas Southwestern Medical Center, Dallas, TX, USA
[d] Department of Biology, University of Dallas, Dallas, TX, USA
[e] Department of Pathology, Children's Medical Center, Dallas, TX, USA
[f] Eugene McDermott Center for Human Growth and Development, University of Texas, Southwestern Medical Center, Dallas, TX, USA
[g] Deccan College of Medical Sciences, Hyderabad, India
[h] NCH Corporation, Irving, TX, USA
[i] Department of Computer Science, University of Dallas, Dallas, TX, USA
[j] Harvard University, Department of Biology, Cambridge, MA, USA

ARTICLE INFO

Keywords:
SARS-CoV-2
COVID-19
Molecular docking
Molecular dynamics

ABSTRACT

SARS-CoV-2 is a newly discovered virus which causes COVID-19 (coronavirus disease of 2019), initially documented as a human pathogen in 2019 in the city of Wuhan China, has now quickly spread across the globe with an urgency to develop effective treatments for the virus and emerging variants. Therefore, to identify potential therapeutics, an antiviral catalogue of compounds from the CAS registry, a division of the American Chemical Society was evaluated using a pharmacoinformatics approach. A total of 49,431 compounds were initially recovered. After a biological and chemical curation, only 23,575 remained. A machine learning approach was then used to identify potential compounds as inhibitors of SARS-CoV-2 based on a training dataset of molecular descriptors and fingerprints of known reported compounds to have favorable interactions with SARS-CoV-2. This approach identified 178 compounds, however, a molecular docking analysis revealed only 39 compounds with strong binding to active sites. Downstream molecular analysis of four of these compounds revealed various non-covalent interactions along with simultaneous modulation between ligand and protein active site pockets. The pharmacological profiles of these compounds showed potential drug-likeness properties. Our work provides a list of candidate anti-viral compounds that may be used as a guide for further investigation and therapeutic development against SARS-CoV-2.

1. Introduction

The new SARS-CoV-2 coronavirus, responsible for causing COVID-19, was initially documented as a human pathogen in December 2019 in the city of Wuhan, Hubei province in China [1]. The virus has quickly spread across the globe, and as of December 2020, there were 119,988,220 cases reported with 2,655,612 fatalities (John Hopkins Coronavirus Resource Center 3/14/2021). Infection by the SARS-CoV-2 virus, a single-stranded RNA virus, results in a wide spectrum of illnesses from an asymptomatic carrier state to mild and severe cold-like symptoms to a fatal pneumonia. Multiple vaccines against the SARS-CoV-2 virus are available in several countries, including three in the United States [2,3]. However, concerns related to the timeline of widespread and global vaccination as well as questions about continued vaccine efficacy

* Corresponding author. , Department of Pathology Bio-center at UT Southwestern Medical Center 2330 Inwood Rd Suite EB3.202-B, Dallas, TX, 75235, USA.
E-mail address: rolando.garcia@utsouthwestern.edu (R. García).

https://doi.org/10.1016/j.compbiomed.2021.104364
Received 14 December 2020; Received in revised form 26 March 2021; Accepted 29 March 2021
Available online 19 April 2021

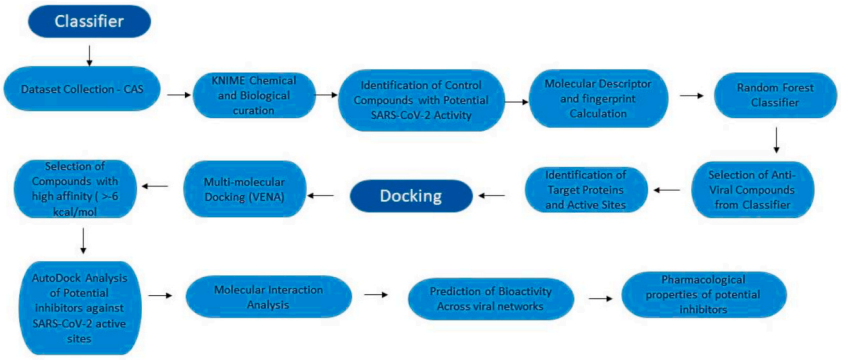

Fig. 1. Process for identifying SARS-CoV-2 inhibitors from CAS dataset.

CAS Index	ETA_AlphaP	ETA_dAlpha_B	ETA_Epsilon_1	ETA_Epsilon_3	ETA_Epsilon_4	ETA_Epsilon_5	ETA_dEpsilon_A
1002335-20-7	0.46869	0.03131	0.81438	0.43608	0.79497	0.95583	0.3783
1009113-51-2	0.44833	0.05167	0.74647	0.43793	0.72095	0.91389	0.30853
100926-23-6	0.46833	0.03167	0.54961	0.44815	0.54961	0.76667	0.10146
100926-24-7	0.4746	0.0254	0.51905	0.44737	0.51905	0.74722	0.07168
100926-41-8	0.47778	0.02222	0.54861	0.45	0.5131	0.74198	0.09861
101468-16-0	0.47778	0.02222	0.53867	0.45	0.5131	0.74198	0.08867
1016774-02-9	0.47949	0.02051	0.59028	0.44054	0.52717	0.80976	0.14974
1024260-90-9	0.46205	0.03795	0.80029	0.43913	0.76253	0.91286	0.36116
1026002-29-8	0.45139	0.04861	0.69921	0.43714	0.64931	0.87816	0.26206
10303-95-4	0.45679	0.04321	0.65229	0.43671	0.61521	0.87957	0.21558
1052524-91-0	0.46236	0.03764	0.65576	0.43933	0.59204	0.84352	0.21643
1052530-88-7	0.45333	0.04667	0.68148	0.43699	0.6366	0.87222	0.2445
1107652-05-0	0.45833	0.04167	0.67024	0.43913	0.6049	0.86036	0.23111

Fig. 2. Snapshot and representative illustration of descriptor calculations. First column describes all compounds (CAS antiviral compounds + controls), while remaining columns show seven of the 46 molecular descriptors with calculated fingerprints.

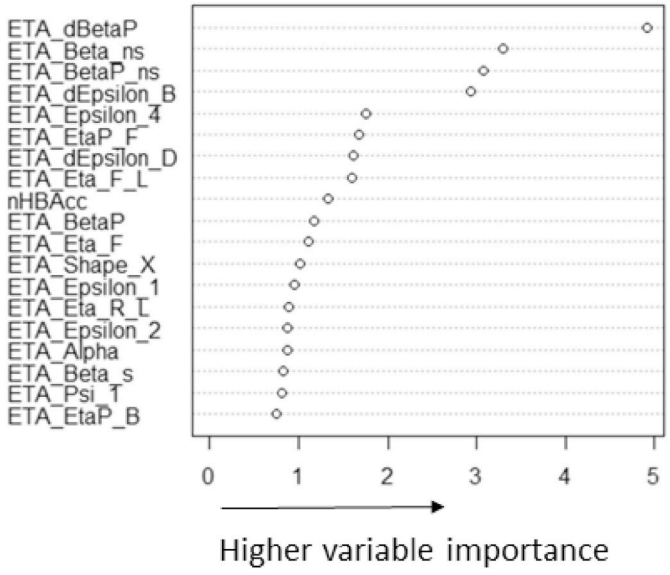

Fig. 3. Random Forest tree predicted plot of molecular descriptors. The plot shows the measure of relative unsaturation content relative to molecular size (ETA_dBetaP) as the most important classifier of potential antiviral compounds.

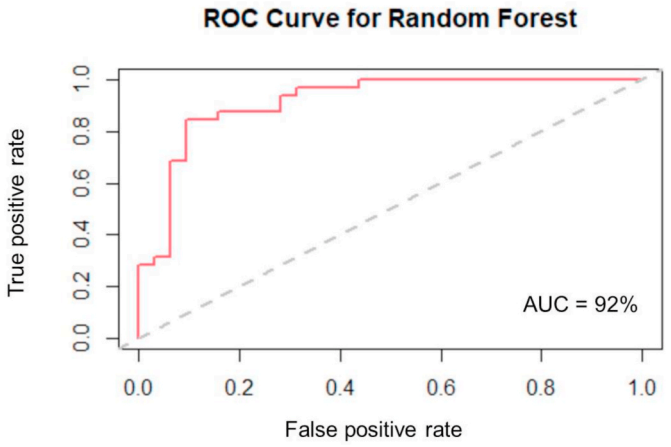

Fig. 4. Receiver operating curve for classifier model developed to identify compounds as potential inhibitors of SARS-CoV-2 target proteins. The area under the curve (AUC) of the classifier showed a good predictive ability of 0.92 (92%).

against newly emerging SARS-CoV-2 variants (e.g. UK and South African) continue to highlight need for development of COVID-19 treatments in parallel to vaccination efforts [4,5]. SARS-CoV-2 belongs to the beta coronavirus genus, which also includes severe acute respiratory syndrome coronavirus (SARS-CoV) and the Middle East respiratory syndrome coronavirus (MERS-CoV). Rapid genomic sequencing of SARS-CoV-2 has enabled comparative analysis between the novel virus and those responsible for previous pandemics [6]. Due to significant homology between the viruses, previously curated knowledge generated through studies with SARS-CoV and MERS-CoV can be used in an attempt to find potential drug targets for SARS-CoV-2 [7]. A tremendous amount of effort has been placed in finding therapeutics for the various coronaviruses. Since the original SARS-CoV emerged in 2002, an effort has been made to target various viral structures and proteins including helicase, protease, endonuclease, exoribonuclease, methyltransferase, and non-structural proteins (NSPs). Researchers have continued to use

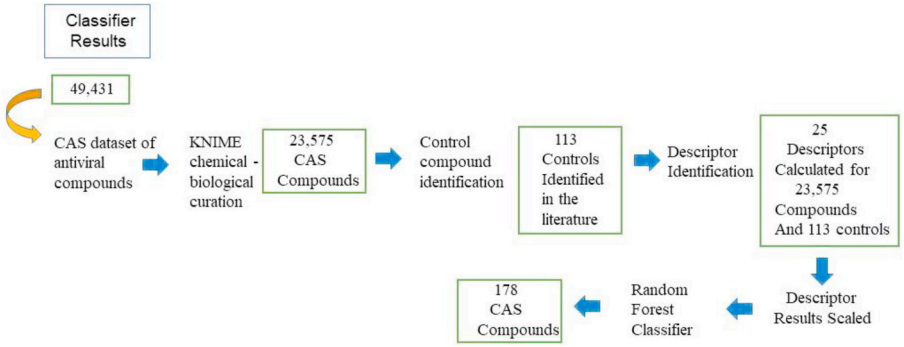

Fig. 5. Summary of results from the classifier model workflow.

Table 1
Active sites of SARS-CoV-2 target proteins used for multi docking.

Target Protein	Active Site Amino Acid Residue	PMID
SBD	LEU455, ALA475, PHE486, GLN493, ARG403, GLY496, ASN501 PHE456, TYR489	32568012 33190011 33184600 33657325 33299995 32623480
NSP12	ASP618, LEU758, SER759, ASP760 LEU758, SER759, ASP760, ASP761, LYS798, LIS813, SER814 LIS813, SER814	32283108 32812340 32438371 32346490
NSP13	LYS288, SER289, ARG567, GLN404 GLU375, ASP374	32346490
NSP15	HIS235, HIS250, LYS290, THR341, TYR343, SER294	32552462
Mpro	CYS 145, HIS41 (catalytic dyad)	32534187
ACE-2	SER19, ASP38, ASN90, GLU329, GLU35, LYS353, LYS31	32225175

Key: PMID, PubMED identifier.

traditional methods to determine antiviral activity of compounds, but these processes can be slow and cumbersome. For these reasons, many researchers have now turned to virtual screening using genomic and structural models. Past efforts have shown that using molecular docking studies as an initial step is useful for screening the most promising antiviral, antibacterial, and antiprotozoal compounds [8,9]. In April 2020, CAS, a division of the American Chemistry Society, released a database containing 49,431 chemical substances assembled from the CAS REGISTRY that have antiviral activity reported in published literature or are structurally similar to known antivirals. In an attempt to find potential anti-viral compounds as inhibitors of SARS-CoV-2, a pharmacoinformatics approach including a classifier model coupled with a multi molecular docking and dynamics analysis was performed.

2. Materials and methods

To identify potential antiviral compounds as inhibitors of SARS-CoV-2, we obtained the CAS dataset of antiviral chemical compounds available at https://www.cas.org/covid-19-antiviral-compounds-dataset. All compounds were converted to Protein Data Bank (PDB) and AutoDock (PDBQT) format for subsequent analysis using the open source Babel package available at http://openbabel.org. The initial data-set of antiviral compounds in SDF format was subjected to chemical and biological curation. The Konstanz Information Miner (KNIME) workflow (https://www.knime.org/) was employed to perform these curations. We use the SDF reader node in the KNIME workflow to read chemical and biological properties of antiviral compounds. For chemical curation, modules in the KNIME workflow included the following for inorganic and organo-metallic removal: SDF reader used to read the input file,

Table 2
Scores generated by multi molecular AutoDock Vina of the top binding compounds to the active site of SARS-CoV-2 targets.

ID	CAS registry Number	Compound Number	Binding Energy (kcal/mol) (Vina)
SBD			
1	1809249-37-3	Control	−6.4
2	1002336-72-2	65	−7.1
NSP12			
3	687967-00-6	137	−12.8
4	1052524-91-0	11	−12
5	911195-33-0	169	−12
6	957794-14-8	177	−11.6
7	774511-55-6	148	−10.9
8	909082-43-5	166	−10.9
9	134637-14-2	34	−9.3
10	190141-36-7	76	−9.2
11	918934-15-3	170	−9.1
12	1399805-48-1	40	−8.7
13	1637769-10-8	61	−8.7
NSP13			
14	687967-00-6	137	−14.2
15	774511-55-6	148	−12.4
16	957794-14-8	177	−12.4
17	768322-56-1	147	−12.0
18	1107652-06-1	14	−11.2
19	710272-33-6	140	−11.2
20	1107652-05-0	13	−11.2
21	911195-33-0	169	−11.0
22	1250937-05-3	25	−10.9
23	122931-60-6	20	−9.8
24	1399805-48-1	40	−9.6
Mpro			
33	1389335-41-4	37	−11.8
34	1817756-68-5	70	−9.5
35	1009113-51-2	2	−9.4
36	1399805-48-1	40	−9.0
37	172645-16-8	64	−9.0
38	134637-14-2	34	−8.6
39	193898-65-6	79	−8.3
40	1883270-61-8	73	−8.2
41	100926-24-7	4	−7.9
42	2256068-99-0	92	−7.9
ACE-2			
43	1399805-48-1	40	−6.5

element filter (removes inorganic and organo-metallic compounds), connectivity (removes mixtures), RDKit Salt Stripper (removes salts), RDKit Optimize Geometry (geometric optimization of screened compounds), RDKit Structure Normalizer (standardizes compounds), RDKit Add Hs (adding of hydrogen), and the SDF writer (generates an output file of screened compounds in SDF format). To perform biological curation, The Duplicate Analysis Workflow using the 3D D-Similarity module was performed to identify duplicate molecules in the dataset. An activity cliff analysis using the Automated Matched Pairs module computes matched molecular pairs and understands molecular activity. A

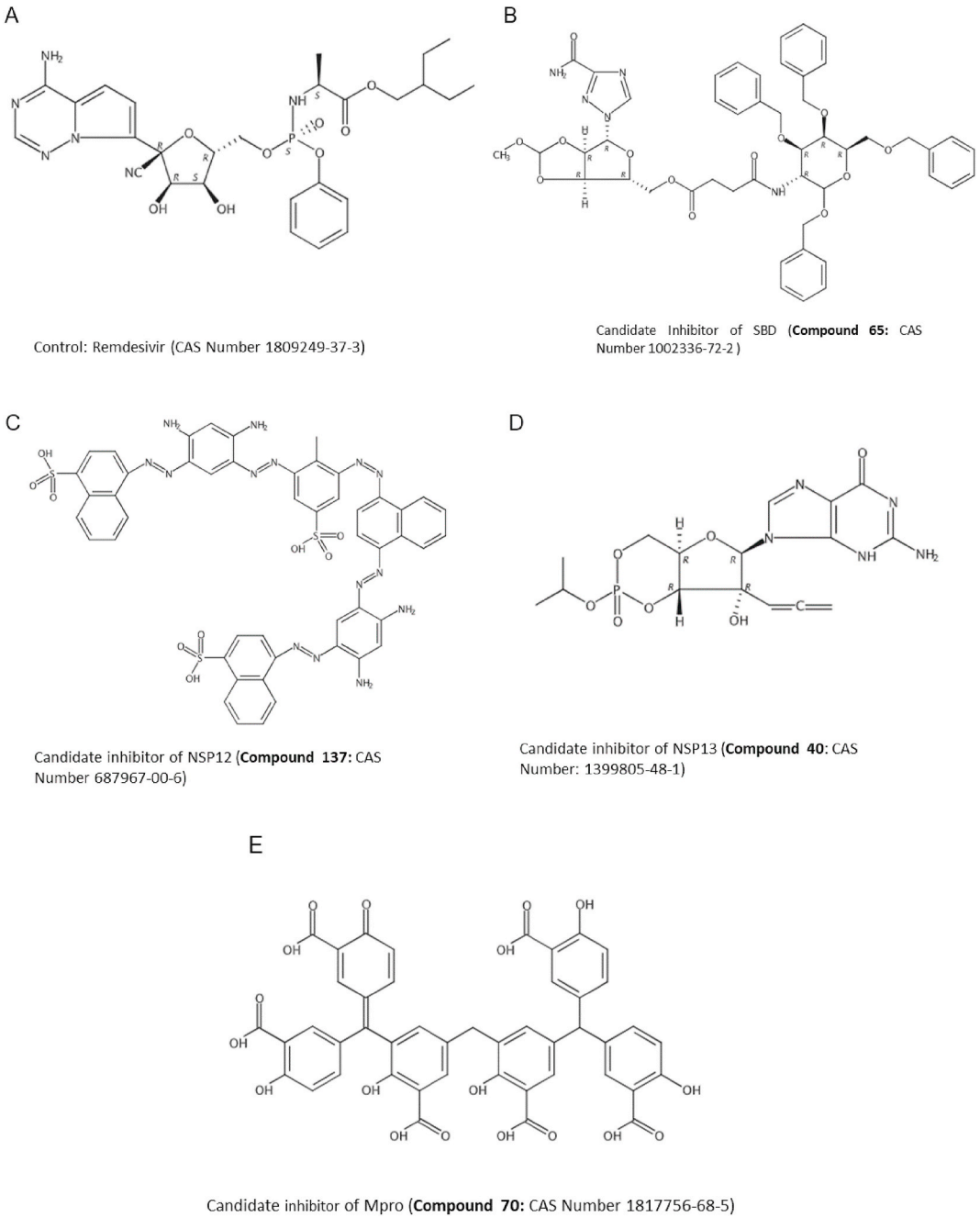

Fig. 6. Chemical Structure of antiviral compounds from the CAS registry as potential inhibitors of SARS-CoV-2 target protein active sites.

careful and manual curation of compounds with similar structure and activity values were then removed. The chemical and biological curation is well documented by Ambure and colleagues [10]. To establish a list of standard or control compounds (i.e. reported potential compounds with favorable interactions against SARS-CoV-2), a search of the literature was performed. These latter set of compounds served to classify prospective inhibitors of SARS-CoV-2 from the CAS dataset. For developing a classification model, a set of molecular descriptors (or indices) for the CAS dataset and controls were calculated using PaDEL, a software package that calculates molecular descriptors and fingerprints available at http://padel.nus.edu.sg/software/padeldescriptor. The generated fingerprints of all compounds from each descriptor were scaled. To obtain the best fit of anti-viral compounds against SARS- CoV-2 from the optimal set of descriptors, a random forest classifier in the R environment [11] was implemented. Control and experimental biomolecules and their accompanying descriptors and fingerprint properties were obtained and divided into testing and training data sets. The training data set consisted of all 32 control molecules and an additional 32 experimental molecules, which were selected from the larger dataset by implementing the Kennard-Stone algorithm with Euclidean distance. The testing data set consisted of 23,520 experimental molecules. A receiver operating curve (ROC) and the area under the curve (AUC) was

Table 3
SMILE and molecular weight of candidate compounds with inhibitory potential for SARS-CoV-2.

Ligand	SMILE Format	Molecular Weight
65	C(OC(=O)CCC(=O)N[C@@H]1[C@@H](OCc2ccccc2)[C@@H](OCc2ccccc2)[C@H](OC1OCc1ccccc1)COCc1ccccc1)[C@@H]1[C@@H]2[C@H]([C@@H](O1)n1nc(nc1)C(=O)N)OC(O2)OC	907.96
137	N(=N\c1c(C)c(/N=N/c2cc(/N=N/c3c4c(c(cc3)S(=O)(=O)[O-])cccc4)c(cc2N)N)cc(c1)S(=O)(=O)[O-])\c1c2c(c(cc1)/N=N/c1cc(/N=N\c3c4c(c(cc3)S(=O)(=O)[O-])cccc4)c(cc1N)N)cccc2	1063.11
40	[C@]123O[C@]1(OC1=COPO[C@@]21[C@H]3C#C)n1c2c(nc1)c(=O)[nH]c(N)n2	425.33
70	C(=C\1/C=C(C(=O)O)C(=O)C=C1)(\c1cc(Cc2cc(C(c3cc(C(=O)O)c(cc3)O)c3cc(C(=O)O)c(cc3)O)cc(c2O)C(=O)O)cc(c1O)C(=O)O)/c1cc(C(=O)O)c(cc1)O	858.71

calculated to determine the reliability of the classifier. Next, four SARS-CoV-2 structural and non-structural protein targets, including NSP12 (RNA polymerase, ID: 6NUR), NSP13 (helicase, ID: 6ZSL), main protease (Mpro, ID:7BQY), spike protein (spike binding region – SBD, ID: 6LZG) from the protein databank (https://www.rcsb.org) were evaluated against potential anti-viral compounds. The angiotensin converting enzyme-2 (ACE-2) protein on the host, essential for attachment to the SBD region that leads to SARS-CoV-2 entry into the host was also evaluated (ID: 7DF4). Moreover, the emerging more transmissible South African SARS-CoV-2 variants with the following amino acid residue substitutions in the SBD region of the spike protein: LYS417ASN, GLU484LYS and ASN501TYR) were also evaluated. Target proteins were refined using AutoDock Tools by deleting water molecules, adding polar hydrogens, Kollman charges, computing the Gasteiger charges and assigning the ADT4 type atoms. The active sites of the target proteins were highlighted by using the grid box of AutoDock. Thereafter, a multi molecular docking analysis between anti-viral compounds and target proteins using AutoDock Vina [12] was subsequently performed. Only those compounds with an affinity to the target protein of less than −7.5 kcal/mol were retained for downstream analysis (binding affinities of less than −6 kcal/mol are generally considered significant values for binding). Thereafter, each of these compounds were manually curated to determine the area of the active site they covered by using Pymol (http://pymol.org). Those ligands that covered a significant area in the active site were subsequently prepared by the LigPrep and Epik modules (Schrödinger Release 2020-4: Schrödinger, LLC, New York, NY) for docking. The AutoDock and MGL Tools were then used to dock these ligands against SARS-CoV-2 active site target structure and non-structure proteins. Subsequently, the Protein-Ligand Interaction Profiler available at https://projects.biotec.tu-dresden.de/plip-web/plip was used to identify the molecular noncovalent interactions of proteins and their ligands. Remdesivir obtained from the PubChem database (pubchem.ncbi.nlm.nih.gov), officially approved as an antiviral to treat SARS CoV-2 patients, was used as a control for docking and molecular interactions studies. To validate the molecular interactions, a molecular dynamic (MD) simulation was then performed. The ligand structures were parametrized, and input files were generated by using CHARM-GUI [13–15] on the docking structures for compounds 65 and remdesivir. The parametrized structures were solvated in a periodic box of TIP3 molecules. The sizes of TIP3 boxes were 90 × 90 × 90 $Å^3$ and 85 × 85 × 85 $Å^3$ for remdesivir-protein and compound 65-protein structures, respectively. Sodium and chloride ions were added to neutralized and obtain a salt concentration of 0.15 mol/L. CHARMM36 force field was used for protein [16]. The simulations were conducted with the NAMD 2.12 package [17]. After 50,000 minimization MD steps with Conjugate gradient algorithm at 0 K, subsequent 600,000 steps equilibration simulation with constrained protein and ligand was applied for both structure at 303.15 K. At the final production MD simulation, NPT ensemble was applied by using Nose-Hoover Langevin piston pressure control [18] at 303.15 K and 1.01325 bars for 20 ns simulation time. 2 fs integration step is applied through all simulation steps. To evaluate the interaction energies and RMSDs, VMD 1.9 software [19] was used. Thereafter, simultaneous modulation of SARS-CoV-2 target proteins was assessed using the virus-associated disease-specific chemo-genomics knowledge (Virus-CKB) available at https://www.cbligand.org/g/virus-ckb. Lastly, pharmacological characteristics, including absorption, distribution, metabolism, excretion, toxicity, drug likeness properties of potential inhibitors of SARS-CoV-2 were explored using the ADMET platform (http://admet.scbdd.com/home/index/). The general workflow illustrating steps taken in screening of the CAS dataset in search of potential inhibitor compounds of SARS-CoV-2 target proteins is shown in Fig. 1.

3. Results

A total of 49,431 compounds were initially obtained from the CAS registry. After the KNIME chemical and biological curation, 23,575 compounds remained. To identify potential compounds as possible inhibitors of SARS-CoV-2 and to establish a set of standards or controls to compare against the CAS dataset, a literature search was performed to identify anti-viral compounds with potential activity against COVID-19. A list of 113 antiviral compounds was selected (see Supplemental Table 1). These compounds served as controls for subsequent classification of the CAS dataset. To construct the classification model, molecular descriptors (topo-chemical atom indices) were generated for each of the compounds (CAS dataset + controls) by using the PaDEL-molecular descriptor software. A total of 25 molecular descriptors (see Supplemental Table 2) and their molecular fingerprints for each of the compounds was calculated and their values subsequently scaled (see Fig. 2).

Thereafter, a matrix composed of 25 descriptors, 23,575 CAS compounds and 113 controls was constructed to perform a random forest tree classifier in the R environment. Fig. 3 shows the random forest tree predicted plot of molecular descriptors and their importance in the classification of antiviral compounds against SARS-CoV-2, while Fig. 4 shows a 0.92 reliability (area under the curve – AUC of the ROC curve) at predicting compounds as inhibitors of SARS-CoV-2.

Based on these findings, the classifier identified 178 antiviral compounds as potential inhibitors of SARS-CoV-2 (Fig. 5).

To validate the compounds obtained from the classifier, a multi molecular docking analysis using AutoDock Vina was performed on 178 compounds against SARS-CoV-2 target proteins: SBD, NSP12, NSP13 (helicase) and Main protease (Mpro). Table 1 shows the protein active site residues that were targeted for multi docking.

Ligands with binding energy less than −6 kcal/mol to SARS-CoV-2 target protein active sites were then selected. A total of 39 compounds plus four duplicates (40, 137, 177, 148) were identified (Table 2).

From these group of compounds, a list of 4 (ID: 40, 65, 70 and 137) were selected based on visual inspection of the active site and the area ligands covered. Selected compounds underwent subsequent analysis, including docking, molecular interactions and description of pharmacological properties. Fig. 6 depicts representative chemical formulas of the four potential antiviral compounds to the various active sites of the

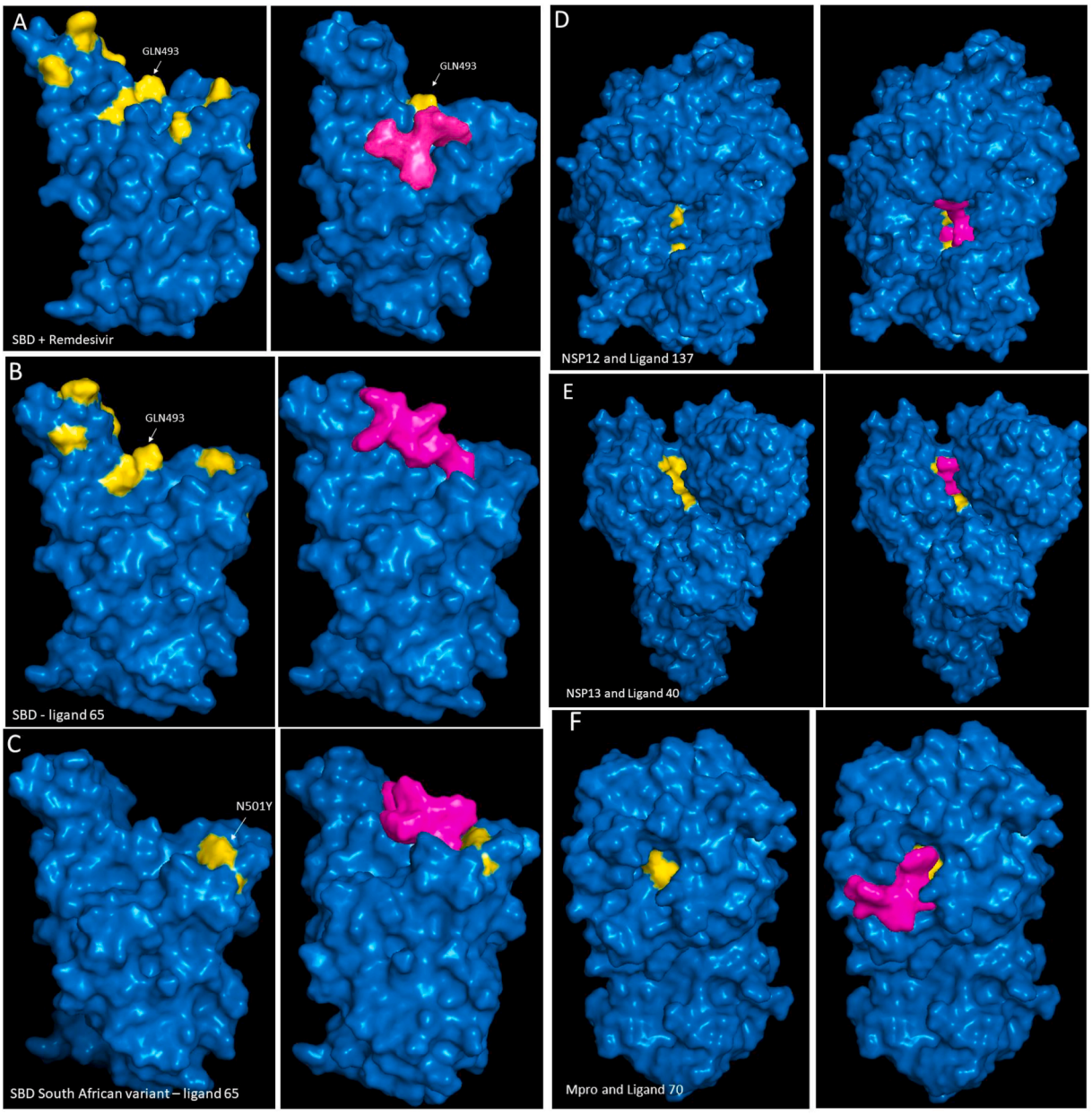

Fig. 7. 3-D surface representation of SARS CoV-2 target proteins and potential inhibitory compounds. Key: blue, SARS-CoV-2 target protein; yellow, target active site; magenta, antiviral compound. Panels (A): The control compound (remdesivir) and its binding relationship to the SBD region. An active site residue in the SBD region GLN493 is labeled in white for reference; (B): ligand 65 predicted binding to the SBD region; (C) ligand 65 binding to the SBD – South African variant with the N501Y (asparagine to tyrosine amino acid residue 501) substitution depicted; (D) ligand 137 bound to NSP12; (E) ligand 40 binding interaction with NSP13; (F) ligand 70 targeting Mpro and panel. Of note, for the South African variant, the ASN417 and LYS484 substitutions are not covered by ligand 65. For the Mpro protein only the dyad active site of HIS45 and CYS145 are shown in yellow.

SARS-CoV-2 proteins, including the control remdesivir, whereas Table 3 describes the SMILE format and molecular weights of the four compounds. The 3D structural surface representation of SARS-CoV-2 target proteins and molecular docking of the potential candidate antiviral compound to each of the targets (including the SBD South African SARS-CoV-2 variant) is illustrated in Fig. 7.

Covalent interactions of the four selected compounds against the active sites of SARS-CoV-2 were then determined using the Biotec molecular profiler, while representative molecular interactions were rendered using Pymol. All molecular interactions and their distance given in angstroms is shown on Table 4, while Fig. 8 illustrates representative molecular interactions between SARS-CoV-2 targets and selected compounds.

To support these interactions, a molecular dynamic simulation was carried out between the control compound (remdesivir) and ligand 65 at the SBD interphase. Both systems were simulated for a fixed period of time (0–20 ns - ns). During this time interval both control and ligand 65 maintain binding affinity to active sites amino acid residues of the SBD protein (see Fig. 9). The binding energies of the two ligands to the SBD region at different time intervals are described on Table 5, while contact

Table 4
Molecular interactions between selected compounds and SARS-CoV-2 active site target proteins.

Target	Compound	Interaction	Amino acid	Distance (Angstroms)
SBD	Standard	Hydrogen	ARG403	3.00
			GLN493	2.70
			GLN493	2.09
			GLY496	3.43
			PHE497	3.06
			GLN498	3.03
			ASN501	2.02
			ASN501	3.91
		Hydrophobic	GLU406	3.92
			LYS417	3.30
			TYR453	3.45
			TYR495	3.41
SBD	65	Hydrogen	ARG403	2.10
			ARG403	2.93
			GLN493	3.01
			GLN493	3.16
			GLY496	2.03
			PHE497	3.20
			GLN498	3.49
			ASN501	2.27
			TYR505	1.70
		Hydrophobic	TYR489	3.59
			PHE490	3.72
SBD_SA	65	Hydrogen	ARG403	3.22
			TYR453	2.38
			GLY496	2.23
			TYR505	2.72
			TYR505	2.85
		Hydrophobic	ILE418	3.93
			TYR453	3.91
			TYR453	3.18
			LEU455	3.77
			GLN493	3.60
			TYR495	3.75
			PHE497	3.94
			TYR501	3.85
			TYR501	3.29
			TYR505	3.02
			TYR505	3.66
			TYR505	3.30
NSP12	137	Hydrogen bond	LYS621	3.67
			CYS622	2.51
			LYS798	2.97
		Hydrophobic	ASP618	3.53
			LYS798	3.59
			TRP800	3.17
		Salt bridges	ASP618	5.00
NSP13	20	Hydrogen bond	GLY285	2.25
			GLY287	3.17
			LYS288	2.45
			SER289	1.92
			ASP374	3.46
			GLU375	3.32
		Hydrophobic	GLU375	3.95
		Salt bridges	LYS288	3.97
			ARG567	4.52
Mpro	70	Hydrogen bond	GLU166	2.66
			GLU166	3.11
		Hydrophobic	HIS41	3.98
			MET165	3.17
			PRO168	3.33
			GLU166	3.61
			ASP187	3.70
		Salt bridges	HIS41	5.43
			HIS41	5.44
			HIS41	5.35
ACE-2	40	Hydrogen	GLU35	3.60
		Hydrophobic	HIS34	3.97
			GLU37	3.70
		Cation interaction	GLU35	3.79
			ASP38	4.42
			LYS353	3.34
			LYS353	4.60

Key: SBD, spike binding domain; SBD_SA, spike binding domain South African variant.

residues at different time intervals are outlined on Table 6.

Ligand binding to active site amino acid residues in the SBD region at different time intervals are bold in black.

The selected list of ligands (40, 65, 70 and 137) were subsequently evaluated with the knowledge base Virus-CKB to determine whether these compounds showed simultaneous modulation of viral pathways in the life cycle of SARS-CoV-2. The analysis predicted all ligands with combined modulation of viral protein targets (e.g. ligand 40 to ACE-2: −9.25 kcal/mol; ligand 65 to ACE-2: −11.26 kcal/mol; ligand 70 to ACE-2: −10.15 kcal/mol, methyl transferase: −8.6 kcal/mol, Mpro: −8.77 kcal/mol, and NSP12: −8.77 kcal/mol; ligand 137 to ACE-2: −11 kcal/mol, Mpro: −9.21 kcal/mol, NSP12: −9.72 kcal/mol, methyl transferase: −10.77 kcal/mol and papain like protease: −8.68 kcal/mol). A spider plot outlining the various viral interactions pathways between selected ligands and targets across viral networks is shown in Fig. 10.

The generated interactions by the spider plot were then evaluated to determine whether ligands bind to SARS CoV-2 target active site amino acid residues. This evaluation revealed only ligand 137 and 40 with a potential drug combination effect to SARS-CoV-2 active site target proteins. Table 7 shows the newly identified interactions. Ligand 137 was also evaluated for the SBD South African variant.

Thereafter, pharmacological properties of these compounds were evaluated using the ADMET Lab server. Properties evaluated included: solubility by the logP distribution coefficient with suggested values of 0–3; volume distribution score (VD), <0.07L/kg confined to blood bound to plasma protein, 0.07–0.7L/kg evenly distributed, >0.7L/kg bound to tissue components (e.g., protein, lipid), blood-brain barrier (BBB) with a suggested value of greater than 0.1 indicating BBB permeation; absorption was evaluated by the Papp (Caco-2 Permeability) with a suggested value greater than −5.15, human intestinal absorption (HIA) including two categories: Category 1 > 30% HIA; Category 0, HIA <30%; elimination by the T1/2 half-life time value (range: >8 h: high; < 8 h: moderate; <3 h: low) and toxicity by the LD50 of acute toxicity score (High-toxicity: 1–50 mg/kg; Toxicity: 51–500 mg/kg; low-toxicity: 501–5000 mg/kg); metabolism and drug likeness assessed by the Lipinski's rule of 5 with a suggested value score of 2 or greater. The pharmacological properties evaluated of the selected compounds are shown on Table 8.

Based on the Lipinski's rule of 5, ligand 65 and 40 showed the highest drug-likeness activity. Ligands 65, 137 and 70 were predicted to have poor aqueous solubility, whereas compound 40 showed poor lipid bilayer permeability. All ligands displayed blood-brain barrier permeation except for compound 70. All set of ligands showed poor absorption, however compound 40 was predicted to have greater than 30% human intestinal absorption. In terms of toxicity and elimination, all showed low toxicity and demonstrated optimal half life time.

4. Discussion

The ongoing SARS-CoV-2 (COVID-19) global epidemic outbreak, infecting people worldwide, has fast track both vaccine and drug therapeutics. In this context, computational methods may be used to decrease the time of drug discovery and development. Indeed, results from in-silico studies have advance ranking of lead compounds and reduce both time and selection of poor lead molecules in the laboratory. These studies have contributed to the development of a series of pharmaco-therapeutic drugs and has proven an effective tool for drug discovery [20–23]. In a similar manner, computational methods have the advantage of testing a wide spectrum of compounds that may be model against specific target areas. For example, a recent work by Wu and colleagues reported compounds from different sources including flavonoids, anti-bacterial, anti-HIV and anti-fungal with activity against

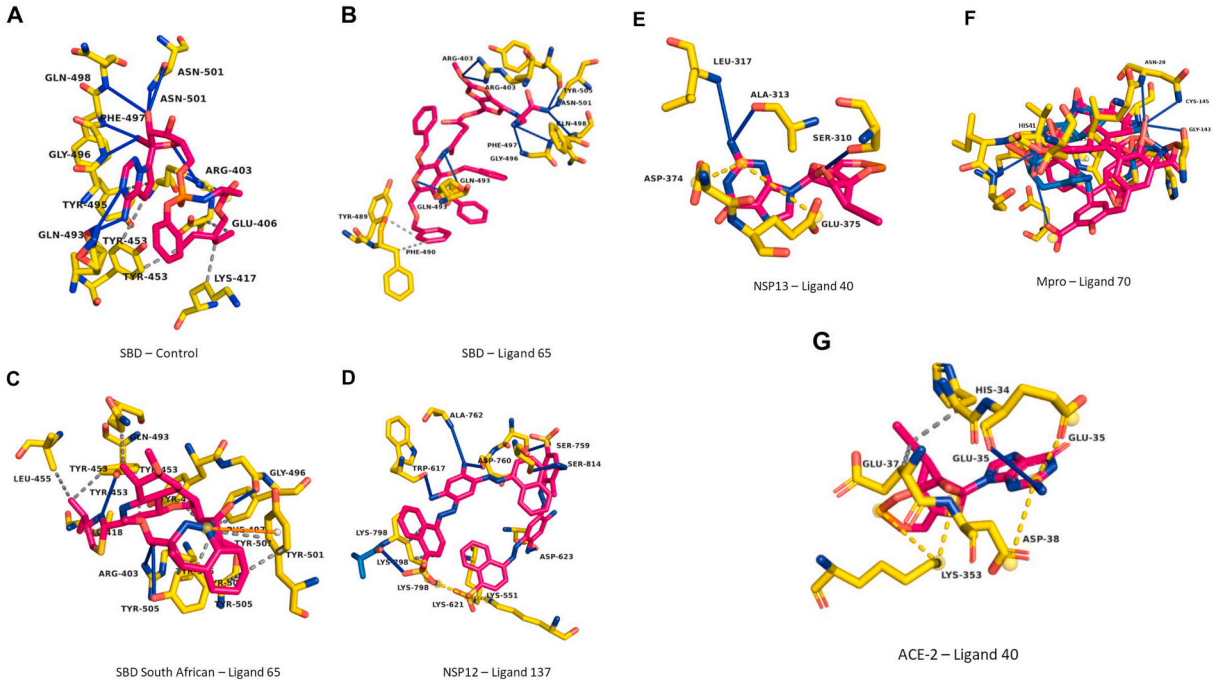

Fig. 8. Molecular interactions between active site amino acid residues and anti-viral compounds. (A) SBD and control ligand remdesivir, (B) SBD with compound 65, (C) SBD South African variant interactions with ligand 65 shows hydrophobic interaction between TYR501 and ligand depicted by gray dashed lines, (D) NSP12 interaction with ligand 137, (E) NSP13 and ligand 40, (F) Mpro and ligand 70, (G) ACE-2 interactions with ligand 40. Molecular interactions represented by solid blue (hydrogen bonding) and dashed lines (gray: hydrophobic; yellow, cation interaction; orange, salt bridge).

SARS-CoV-2 [24]. Likewise, Natesh et al. described culinary compounds with potential to inhibit main protease, spike protein and ACE2 receptors of SARS-CoV-2 [25]. In a similar framework, Feng et al. used pattern recognition computing to identify ten potential anti-viral compounds from the CAS COVID-19 dataset [26]. Similarly, during the first months of the COVID-19 outbreak, machine learning approaches have been applied to a wide spectrum of areas in an effort to better understand and contain the virus, including genomics, prediction of patient outcomes and infections, assay development and potential drug discovery as inhibitors of COVID-19 [27–32].

In this study, computational methods including a random forest classifier along with molecular docking, interaction profiles and molecular simulation were used to identify potential anti-viral compounds against SARS-CoV-2 structure and non-structure proteins. The random forest classifier showed 0.92 (AUC of the ROC curve) reliability at identifying potential inhibitors and selected 178 compounds with potential activity against SARS-CoV-2 from a large dataset of antiviral compounds; it also provided key chemical important features for prioritization of anti-viral compounds and identified key molecular fingerprints as most important classifiers. Binding affinity of these compounds were subsequently validated and confirmed using a multi-molecular docking analysis. Of these, only 39 were validated to bind active sites of viral structural (SBD spike region) and non-structural proteins involve in RNA replication (NSP12-RNA polymerase, NSP13-helicase, and Mpro-main protease). Although the latter compounds showed coverage to SARS-CoV-2 active sites, only four (compounds 40, 65, 70 and 137) were selected for subsequent downstream analysis (molecular interactions and pharmacological properties) based on visual inspection of the active sites covered. The selected compounds showed high affinities to structure and non-structure SARS-CoV-2 targets and formed non-covalent interactions with key active site amino acid residues. All but one, to the best of our knowledge, are exclusively cited as anti-viral agents in the CAS dataset. Only compound 70 (C45 H30 O18, Benzoic acid, 5-[bis(3-carboxy-4-hydroxyphenyl)methyl]-3-[[3-carboxy-5-[(3-carboxy-4-hydroxyphenyl)(3-carboxy-4-oxo-2,5-cyclohexadien-1-ylidene)methyl]-4-hydroxyphenyl]methyl]-2-hydroxy, CAS #1817756-68-5) and its derivatives have varied applications (e.g. treating conditions including nociceptive pain, cancer, age-related macular degeneration, and hemolytic uremic syndrome) [33–36]. To support the molecular interactions identified, molecular dynamics (MD) of a control molecule (remdesivir) and ligand 65 (C48H53 N5 O13, D-Galactopyranoside, phenylmethyl 2-[(3-carboxy-1-oxopropyl)amino]-2-deoxy-3,4,6-tris-O-(phenylmethyl)-, 5'-ester with 1-[2, 3-O-(methoxymethylene)-β-D-ribofuranosyl]-1H-1,2,4-triazole-3-carboxamide, CAS #1002336-72-2) were simulated against the SBD region of the spike protein. These results show the interactions of remdesivir and ligand 65 against active site amino acid residues (ARG403, GLN 493, GLY496, TRY489, ASN 501) of the SBD spike protein region were maintained with strong binding affinity throughout the entire trajectory time interval. Only LEU455 not initially identified by docking and molecular profiles was later described by the MD simulation. We recognize that only one control and one compound were used to confirm the predicted molecular interactions by MD simulation and additional in-vitro and in-vivo experimental studies are warranted to evaluate the bioactivity effects of these compounds on target proteins. Nonetheless, our initial observations provide a foundation to pursue these candidate compounds as potential therapeutics against SARS-CoV-2.

Although the single-drug approach of selected compounds as inhibitors of SARS-CoV-2 may effectively target viral active pockets, it

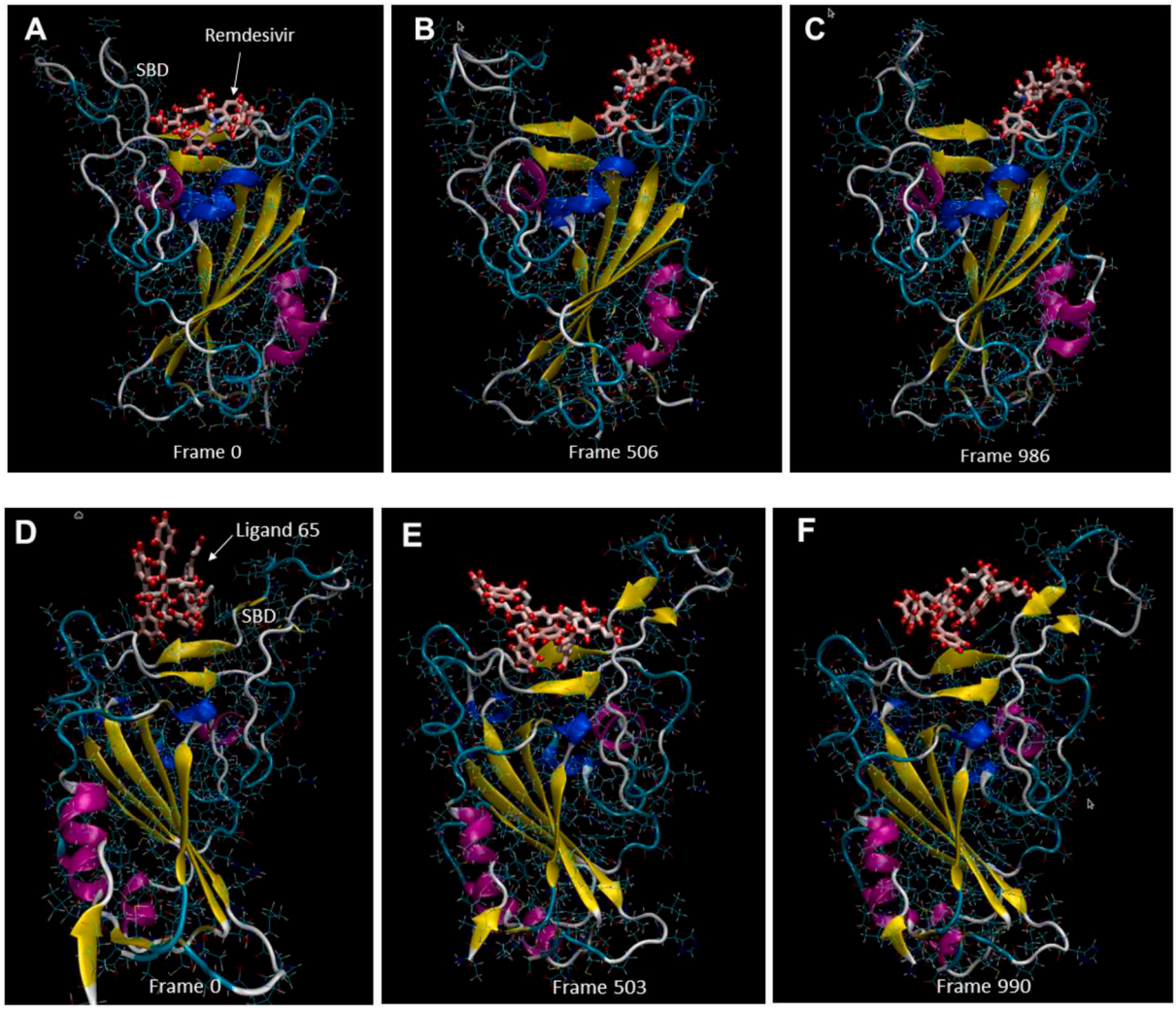

Fig. 9. Molecular dynamic simulation of the control and ligand 65 against the SBD region. Panels A + B + C show the control remdesivir binding to the SBD region at multiple time interval frames (0; 506, 10.14 ns; and 986, 19.74 ns). Panels D + E + F illustrate ligand 65 binding to the SBD region at interval frames (0; 503, 10.08 ns; and 990, 19.82 ns).

Table 5
Energies of Remdesivir and ligand 65 to the SBD active site regions at different time intervals.

Remdesivir:					
Frame	Time (ns)	Elec	VdW	Nonbond	Total
0	0	−18.804	−36.5253	−55.3293	−55.3293
506	10.14	−15.3262	−29.2573	−44.5835	−44.5835
986	19.74	−15.4416	−30.4299	−45.8715	−45.8715
Ligand 65					
Frame	Time (ns)	Elec	VdW	Nonbond	Total
0	0	+4.4186	−30.7731	−26.3545	−26.3545
503	10.08	−9.6287	−30.4741	−40.1028	−40.1028
990	19.82	−15.0934	−31.3019	−46.3953	−46.3953

Key: Energies are in kcal/mol; ns, nanoseconds; Elec, electrostatic; VdW, Vanderwall; Nonbond, nonbonded.

From Table 5, the total energy suggests remdesivir with an initial stronger binding energy to the SBD region, however, ligand 65 showed a stronger binding energy at the end of the 20 ns simulation.

may not be enough to arrest the life cycle of the virus and both multi-target or combinations of drugs may be needed to treat COVID-19. This is consistent with current clinical trials [37]. In this alternative approach, we applied the four selected compounds to the virus chemo-genomics platform described by Feng and colleagues to identify potential drug treatment combinations against SARS-CoV-2. Several combinations were identified that target structural and non-structural targets simultaneously, however, only two (ligand 40, C16H20 N5 O7 P, Guanosine, 2'-C-1,2-propadien-1-yl-, cyclic 3′,5'-(1-methylmethyl phosphate), CAS #1399805-48-1 and ligand 137, C49H38 N14 O9 S3, 1-Naphthalenesulfonic acid, 4-[[2,4-diamino-5-[[4-[[3-[[2,4-diamino-5-[(4-sulfo-1-naphthalenyl)azo]phenyl]azo]-2-methyl-5-sulfophenyl]azo]-1-naphthalenyl]azo]phenyl]azo], CAS #687967-00-6) were confirmed by docking and molecular interactions to bind to active site pockets of SARS-CoV-2. Of these, compound 40 showed binding to both NSP13 and ACE-2, while compound 137 showed affinity to NSP12 and the SBD region of the spike protein. Interestingly, compound 137 was able to bind with high affinity to the South African variant amino acid residues ASN417 and LYS484. These latter two compounds possibly generating a drug combined effect treatment strategy to both structure and replication apparatus of the virus and possibly inhibiting the newly emerged South African variant to gain access to the host may prove interesting therapeutic candidates to pursue.

In terms of the pharmacological profiles, selected compounds demonstrated acceptable drug-likeness metrics, especially compounds 40 and 65 as inhibitors of NSP13/ACE-2 and the SBD region each. Although the metric for drug likeness of compounds 70 and 137 were not optimal, it should be noted that clinical drugs may have

Table 6
Contact amino acid residues between ligands and the SBD region at different time (nanoseconds).

Ligand	Time Intervals		
Remdesivir			
	Time = 0	Time = 10.14	Time = 19.74
	Frame 0	Frame 506	Frame 986
	ARG-403	**ARG-403**	**ARG-403**
	ASP-405	GLY-447	GLY-447
	GLU-406	TYR-449	TYR-449
	LYS-417	TYR-453	TYR-453
	ILE-418	SER-494	SER-494
	ASN-422	TYR-495	TYR-495
	TYR-453	**GLY-496**	**GLY-496**
	ARG-454	PHE-497	PHE-497
	LEU-455	GLN-498	GLN-498
	GLN-493	THR-500	THR-500
	SER-494	**ASN-501**	**ASN-501**
	TYR-495	GLY-502	GLY-502
	GLY-496	GLY-504	TYR-505
	PHE-497	TYR-505	GLN-506
	GLN-498	GLN-506	
	ASN-501		
	TYR-505		
Ligand 65			
	Time = 0	Time = 10.08	Time = 19.82
	Frame 0	Frame 503	Frame 990
	TYR-449	TYR-351	TYR-449
	ASN-450	TYR-449	LEU-452
	LEU-452	ASN-450	**LEU-455**
	LEU-455	TYR-451	PHE-456
	GLU-484	LEU-452	ILE-472
	TYR-489	THR-470	TYR-473
	PHE-490	GLU-471	GLY-485
	LEU-492	ILE-472	CYS-488
	GLN-493	PHE-490	**TYR-489**
	SER-494	LEU-492	PHE-490
	TYR-495	**GLN-493**	LEU-492
	GLY-496	SER-494	**GLN-493**
	GLN-498		SER-494

characteristics that result them to score low on indices of drug likeness. For example, one of the compounds initially identified as a candidate inhibitor of NSP13, ligand 25 CAS #1250937-05-3, showed low drug likeness metrics, however, it has been reported as a clinical drug to inhibit viral nonstructural protein NS5A [38]. Likewise, drugs with low drug likeness scores may exhibit acceptable pharmacokinetic profiles, a prerequisite for an effective clinical drug [39]. In regards to toxicity, all selected compounds showed low level of toxicity. Although, the present work lacks in-vitro and in-vivo investigations to ascertain selected compounds toxicity and bioactivity against SARS-CoV-2, we believe our initial observations may be used as appropriate guide for the next phases of drug discovery to fast track drug development and therapeutics against SARS-CoV-2 (COVID-19).

5. Conclusion

With the global spread of SARS-CoV-2 and with the roll-out of the first approved vaccines, we anticipate prevention and decrease number of fatalities due to SARS-CoV-2 transmission. Nonetheless, challenges will still remain, mainly emerging variants that are more transmissible. Undoubtedly, advances in anti-viral drug development and clinical trials will lead to our arsenal of effective therapeutics against SARS-CoV-2. In this study, we have proposed use of a machine learning tool along molecular docking studies and molecular interactions to identify potential therapeutics against SARS-CoV-2. Our findings from surface 3-D structures between SARS-CoV-2 target sites and anti-viral compounds, along with molecular interactions and simulations suggests selected compounds may serve as a guide for the next phases of drug discovery to treat COVID-19. Though only a selected list of compounds was detailed for further docking and molecular interactions, the remaining compounds identified here and their affinity to the active site of viral targets suggests these compounds may also be used as a guide for further investigation. Lastly, the potential of these compounds (40, 65, 70, 137), and mainly ligand 40 and 137 to simultaneously modulate structure and non-structure protein in SARS-CoV-2 and target variants may provide

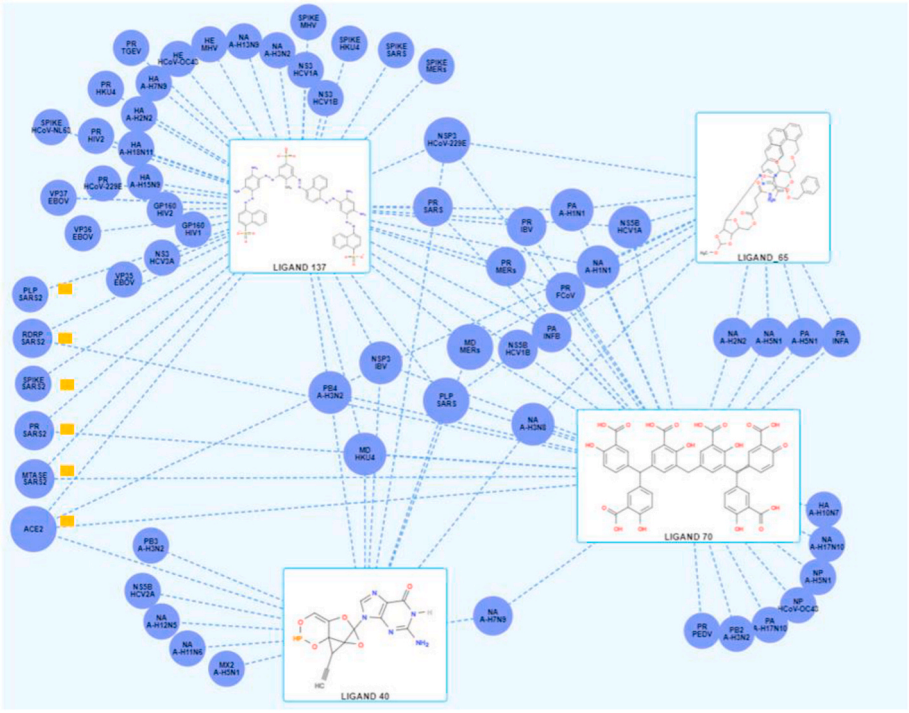

Fig. 10. Spider plot of selected compounds (40, 65, 70, 137) against viral targets using the knowledgebase Virus-CKB. The blue circle with yellow highlights represents the predicted targets (Spike, spike binding protein; ACE2, angiotensin converting enzyme-2; MTA SARS2: methyl transferase; PR, Mpro and RDRP, NSP12 – RNA dependent RNA polymerase; PLP, papain-like protease) of selected compounds and dashed lines represent their interactions.

Table 7
Assessment of ligands with potential drug combination effect against SARS-CoV-2 active site amino acid residues.

Target	Compound	Interaction	Amino acid Residue	Distance (Angstroms)	Binding Energy Auto-Dock Kcal/mol
SBD	137				−8.6
		Hydrogen	**ARG403**	2.56	
			TYR421	1.73	
			TYR453	2.45	
			TYR453	3.20	
			ALA475	3.39	
			TYR489	2.18	
			GLY496	2.90	
			GLY496	3.13	
			GLN498	3.05	
			ASN501	2.93	
		Hydrophobic	**LEU455**	3.27	
			PHE456	3.41	
			PHE456	3.10	
			ALA475	3.84	
			TYR489	3.74	
			GLN493	3.11	
			TYR495	3.04	
			GLN498	3.82	
			TYR505	3.49	
			TYR505	3.98	
SBD - SA		Hydrogen	**ASN417**	1.76	−12.41
			LYS484	3.54	
			LEU492	2.55	
			GLN493	2.28	
			SER494	3.09	
		Hydrophobic	449TYR	3.13	
			455LEU	3.06	
			455LEU	3.39	
			456PHE	3.87	
			493GLN	3.88	
			493GLN	3.53	
ACE-2	40				−6.73
		Hydrogen	GLU35	3.60	
		Hydrophobic	HIS34	3.97	
			GLU37	3.70	
		Salt bridge	GLU35	3.79	
			ASP38	4.42	
			LYS353	3.34	
			LYS353	4.60	

Key: SBD-SA, spike binding region South African Variant. Active site amino acid residues and South African variants are bold in black.

guidance to design and deliver effective therapeutics to treat COVID-19.

CRediT authorship contribution statement

Rolando García: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration. **Anas Hussain:** Methodology, Validation, Formal analysis, Investigation, Data curation, Writing – original draft. **Prasad Koduru:** Writing – review & editing. **Murat Atis:** Formal analysis. **Kathleen Wilson:** Writing – review & editing. **Jason Y. Park:** Writing – review & editing. **Inimary Toby:** Methodology, Formal analysis, Data curation. **Kimberly Diwa:** Methodology, Formal analysis. **Lavang Vu:** Formal analysis. **Samuel Ho:** Formal analysis. **Fajar Adnan:** Formal analysis. **Ashley Nguyen:** Formal analysis. **Andrew Cox:** Formal analysis. **Timothy Kirtek:** Formal analysis. **Patricia García:** Formal analysis. **Yanhui Li:** Formal analysis. **Heather Jones:** Formal analysis. **Guanglu Shi:** Formal analysis. **Allen Green:** Writing – original draft. **David Rosenbaum:** Writing – original draft, All authors have read and agreed to the published version of the manuscript.

Table 8
ADMET pharmacokinetic evaluation of ligands with potential activity against SARS-CoV-2 target structures.

Properties	Compound	65	137	40
ADMET/Lab				
Solubility (LogP) (suggested: 0–3)	4.1	12.6	−0.972	5.43
BBB (Suggested: > 0.1)	0.53	0.70	0.819	0.05
Absorption (suggested: higher than −5.15)	−5.24	−6.3	−5.22	−6.46
HIA category Suggested: 1	0	0	1	0
Volume Distribution (suggested: 0.04–20 L/kg)	−4.73	−1.01	−0.54	−0.96
Elimination (half-life time, suggested > 0.5hr)	2	2.3	1.15	2.62
Toxicity Suggested (>500 mg/kg)	1256.2	1707.5	714.6	846.9
Rule of 5 (Drug likeness) Suggested (Minimum 2)	3	1	3	0
Metabolism		CYP1A2	CYP1A2	
Substrate	CYP3A4			CYP2C9
Inhibitor				

Declaration of competing interest

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

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2021.104364.

Funding

This research received no external funding.

Disclosures

The authors have no financial relationships relevant to this article to disclose.

Funding sources

None.

References

[1] D. Benvenuto, et al., The 2019-new coronavirus epidemic: evidence for virus evolution, J. Med. Virol. 92 (4) (2020) 455–459.
[2] FDA moderna COVID-19 vaccine emergency use authorization (EUA) letter of authorization, December 18, https://www.fda.gov/media/144636/download, 2020.
[3] FDA pfizer COVID-19 vaccine emergency use authorization (EUA) letter of authorization, reissued December 23, https://www.fda.gov/media/144412/download, 2020.
[4] T. Koyama, D. Weeraratne, J.L. Snowdon, L. Parida, Emergence of drift variants that may affect COVID-19 vaccine development and antibody treatment, Pathogens 9 (5) (2020) 324, https://doi.org/10.3390/pathogens9050324.
[5] J. Wise, Covid-19: new coronavirus variant is identified in UK, BMJ 371 (2020), https://doi.org/10.1136/bmj.m4857.
[6] A. Wu, et al., Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China, Cell Host Microbe 27 (3) (2020) 325–328.
[7] J. He, et al., Potential of coronavirus 3C-like protease inhibitors for the development of new anti-SARS-CoV-2 drugs: insights from structures of protease and inhibitors, Int. J. Antimicrob. Agents 56 (2) (2020), 106055.
[8] B. Naik, et al., High throughput virtual screening reveals SARS-CoV-2 multi-target binding natural compounds to lead instant therapy for COVID-19 treatment, Int. J. Biol. Macromol. 160 (2020) 1–17.
[9] X. Lin, X. Li, X. Lin, A review on applications of computational methods in drug screening and design, Molecules 25 (6) (2020).

[10] P. Ambure, et al., Identifying natural compounds as multi-target-directed ligands against Alzheimer's disease: an in silico approach, J. Biomol. Struct. Dyn. 37 (5) (2019) 1282–1306.
[11] Team, R.c., R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. URL http://www.R-project.org/. 2013.
[12] S.D. Handoko, et al., QuickVina: accelerating AutoDock Vina using gradient-based heuristics for global optimization, IEEE ACM Trans. Comput. Biol. Bioinf 9 (5) (2012) 1266–1272.
[13] S. Jo, T. Kim, V.G. Iyer, W. Im, CHARMM-GUI: a web-based graphical user interface for CHARMM, J. Comput. Chem. 29 (2008) 1859–1865.
[14] B.R. Brooks, C.L. Brooks III, A.D. MacKerell Jr., et al., CHARMM: the biomolecular simulation program, J. Comput. Chem. 30 (2009) 1545–1614.
[15] J. Lee, X. Cheng, J.M. Swails, M. S, et al., CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field, J. Chem. Theor. Comput. 12 (2016) 405–413.
[16] R.B. Best, X. Zhu, J. Shim, et al., Optimization of the additive CHARMM all atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain χ(1) and χ(2) dihedral angles, J. Chem. Theor. Comput. 8 (2012) 3257–3273.
[17] Phillips, et al., J. Comput. Chem. 26 (2005) 1781–1802.
[18] G.J. Martyna, D.J. Tobias, M.L. Klein, Constant pressure molecular dynamics algorithms, J. Chem. Phys. 101 (1994) 4177–4189.
[19] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38.
[20] H.J. Boehm, et al., Novel inhibitors of DNA gyrase: 3D structure based biased needle screening, hit validation by biophysical methods, and 3D guided optimization. A promising alternative to random screening, J. Med. Chem. 43 (14) (2000) 2664–2674.
[21] S.J. Lu, F.C. Chong, Combining molecular docking and molecular dynamics to predict the binding modes of flavonoid derivatives with the neuraminidase of the 2009 H1N1 influenza A virus, Int. J. Mol. Sci. 13 (4) (2012) 4496–4507.
[22] J. Bajorath, Chemoinformatics and Computational Chemical Biology, Springer, 2011.
[23] M.U. Mirza, et al., Perspectives towards antiviral drug discovery against Ebola virus, J. Med. Virol. 91 (12) (2019) 2029–2048.
[24] C. Wu, et al., Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods, Acta Pharm. Sin. B 10 (5) (2020) 766–788.
[25] J. Natesh, et al., Culinary spice bioactives as potential therapeutics against SARS-CoV-2: computational investigation, Comput. Biol. Med. 128 (2020), 104102.
[26] Zhing Feng, Maozi Chen, Ying Xue, et al., MCCS: a novel recognition pattern-based method for fast track discovery of anti-SARS-CoV-2 drugs, Briefings Bioinf. (2020) 1–17.
[27] G.S. Randhawa, et al., Machine learning using intrinsic genomic signatures for rapid classification of novel pathogens: COVID-19 case study, PloS One 15 (4) (2020) e0232391.
[28] A. Alimadadi, et al., Artificial intelligence and machine learning to fight COVID-19, Physiol. Genom. 52 (4) (2020) 200–202.
[29] A. Banerjee, et al., Use of machine learning and artificial intelligence to predict SARS-CoV-2 infection from full blood counts in a population, Int. Immunopharm. 86 (2020), 106705.
[30] A.B. Gussow, et al., Genomic determinants of pathogenicity in SARS-CoV-2 and other human coronaviruses, Proc. Natl. Acad. Sci. U. S. A. 117 (26) (2020) 15193–15199.
[31] A. Tarnok, Machine learning, COVID-19 (2019-nCoV), and multi-OMICS, Cytometry 97 (3) (2020) 215–216.
[32] C. Mirabelli, et al., Morphological Cell Profiling of SARS-CoV-2 Infection Identifies Drug Repurposing Candidates for COVID-19, bioRxiv, 2020.
[33] Patrick L. McGeer, Moonhee Lee, Krista Kennedy, Edith G. McGeer, Douglas N. Bell, Methods for treating and preventing nociceptive pain using a compound selected from diacetyl salicylic acid, aurin tricarboxylic acid, aurin quadracarboxylic acid and aurin hexacarboxylic acid, From U.S. Pat. Appl. Publ., 2018. US 20180207180 A1 Jul 26, 2018.
[34] Patrick L. McGeer, Moonhee Lee, Methods and Kits for Prevention and Treatment of Cancer Comprising Aurin Tricarboxylic Acid, Aurin Quadracarboxylic Acid, Aurin Hexacarboxylic Acid and Aurin Tricarboxylic Acid Complex, From U.S Pat. Appl. Publ., 2020. US 10653651 May 19, 2020.
[35] Patrick L. McGeer, Moonhee Lee, Treatment of Age-Related Macular Degeneration, From U.S Pat. Appl. Publ., 2013. US 20130035392 Feb 7, 2013.
[36] Patrick L. McGeer, Moonhee Lee, Treatment of Atypical Hemolytic Uremic Syndrome, From U.S Pat. Appl. Publ., 2013. US 20130035392 Feb 7, 2013.
[37] ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2020 Feb. Identifier NCT04261270, A Randomized,Open,Controlled Clinical Study to Evaluate the Efficacy of ASC09F and Ritonavir for 2019-nCoV Pneumonia; 2020 Feb 7 [cited 2020 Feb 7]; Available from: www. https://clinicaltrials.gov/ct2/show/NCT04261270.
[38] M. Belema, Preparation of biphenylacetylenes, biphenylethers, terphenyls and related derivatives end-capped with amino acid or peptide derivatives as hepatitis C virus inhibitors, PCT Int. Appl. (2012) 317. WO 2012039717 A1 20120329, 2012.
[39] T.J. Ritchie, S.J. Macdonald, How drug-like are 'ugly' drugs: do drug-likeness metrics predict ADME behavior in humans? Drug Discov. Today 4 (19) (2012) 489–495.