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) 104359

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

Identification of novel compounds against three targets of SARS CoV-2 coronavirus by combined virtual screening and supervised machine learning

Onat Kadioglu [a,1], Mohamed Saeed [a,1], Henry Johannes Greten [b], Thomas Efferth [a,*]

[a] Department of Pharmaceutical Biology, Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg University, Mainz, Germany
[b] Heidelberg Clinic of Integrative Diagnostics, Heidelberg, Germany

ARTICLE INFO

Keywords:
Artificial intelligence
Chemotherapy
COVID-19
Infectious diseases
Natural products

ABSTRACT

Coronavirus disease 2019 (COVID-19) is a major threat worldwide due to its fast spreading. As yet, there are no established drugs available. Speeding up drug discovery is urgently required. We applied a workflow of combined *in silico* methods (virtual drug screening, molecular docking and supervised machine learning algorithms) to identify novel drug candidates against COVID-19. We constructed chemical libraries consisting of FDA-approved drugs for drug repositioning and of natural compound datasets from literature mining and the ZINC database to select compounds interacting with SARS-CoV-2 target proteins (spike protein, nucleocapsid protein, and 2′-o-ribose methyltransferase). Supported by the supercomputer MOGON, candidate compounds were predicted as presumable SARS-CoV-2 inhibitors. Interestingly, several approved drugs against hepatitis C virus (HCV), another enveloped (−) ssRNA virus (paritaprevir, simeprevir and velpatasvir) as well as drugs against transmissible diseases, against cancer, or other diseases were identified as candidates against SARS-CoV-2. This result is supported by reports that anti-HCV compounds are also active against Middle East Respiratory Virus Syndrome (MERS) coronavirus. The candidate compounds identified by us may help to speed up the drug development against SARS-CoV-2.

1. Introduction

In the Chinese city of Wuhan, Hubei province, several cases of novel, SARS-like, severe pneumonia occurred in December 2019, as confirmed by the Chinese Center for Disease Control and Prevention and the China Office of the World Health Organization on December 31, 2019. Sequencing of the complete genome on January 13th, 2020 showed that it was a novel coronavirus (GenBank No. MN908947). The official name is SARS-CoV-2. The previous, preliminary names were 2019-nCoV or Wuhan virus. The disease caused by SARS-CoV-2 has been termed Coronavirus disease 2019 (COVID-19) [1], which has been declared by the World Health Organization (WHO) as a global pandemic.

SARS-CoV-2 is an enveloped positive-sense single-stranded RNA virus (ssRNA) consisting of 29,903 nucleotides and two untranslated sequences of 254 and 229 nucleotides at the 5′- and 3′-ends, respectively (GenBank No. MN908947) [2]. The putative genes code for a surface spike glycoprotein, an envelope membrane glycoprotein, a nucleocapsid phosphoprotein, a replicase complex and five other proteins, which compare to SARS-CoV and other coronaviruses. Comparable to SARS-CoV, the novel SARS-CoV-2 enters human cells via binding of the viral spike protein to the human angiotensin-converting enzyme 2 (ACE2) [3,4]. Some coronaviruses also express hemagglutinin esterase on the surface, which is a shorter spike-like protein.

Primary infective hosts were supposed to be traded as foods at the Huanan Fish and Seafood market in Wuhan, since several of the very first patients worked on this market. High sequence similarities of SARS-CoV-2 to coronaviruses in the Malayan pangolin (*Sunda pangolins*) [5] and bats (*Rhinolophi sinicus*) [3,6] suggest that the virus might be

Abbreviations: ACE2, angiotensin converting enzyme 2; AUC, area under the curve; COVID-19, coronavirus disease 2019; FDA, Food and Drug Administration; HIV, human immunodeficiency virus; LBE, lowest binding energy; MERS, Middle East Respiratory Syndrome; ROC, receiver operating characteristic; SARS, severe acute respiratory syndrome; ssRNA, single-stranded RNA virus; WHO, World Health Organization.

* Corresponding author. Staudinger Weg 5, 55128, Mainz, Germany.
E-mail address: efferth@uni-mainz.de (T. Efferth).

[1] both authors contributed equally.

https://doi.org/10.1016/j.compbiomed.2021.104359
Received 24 November 2020; Received in revised form 24 March 2021; Accepted 24 March 2021
Available online 30 March 2021

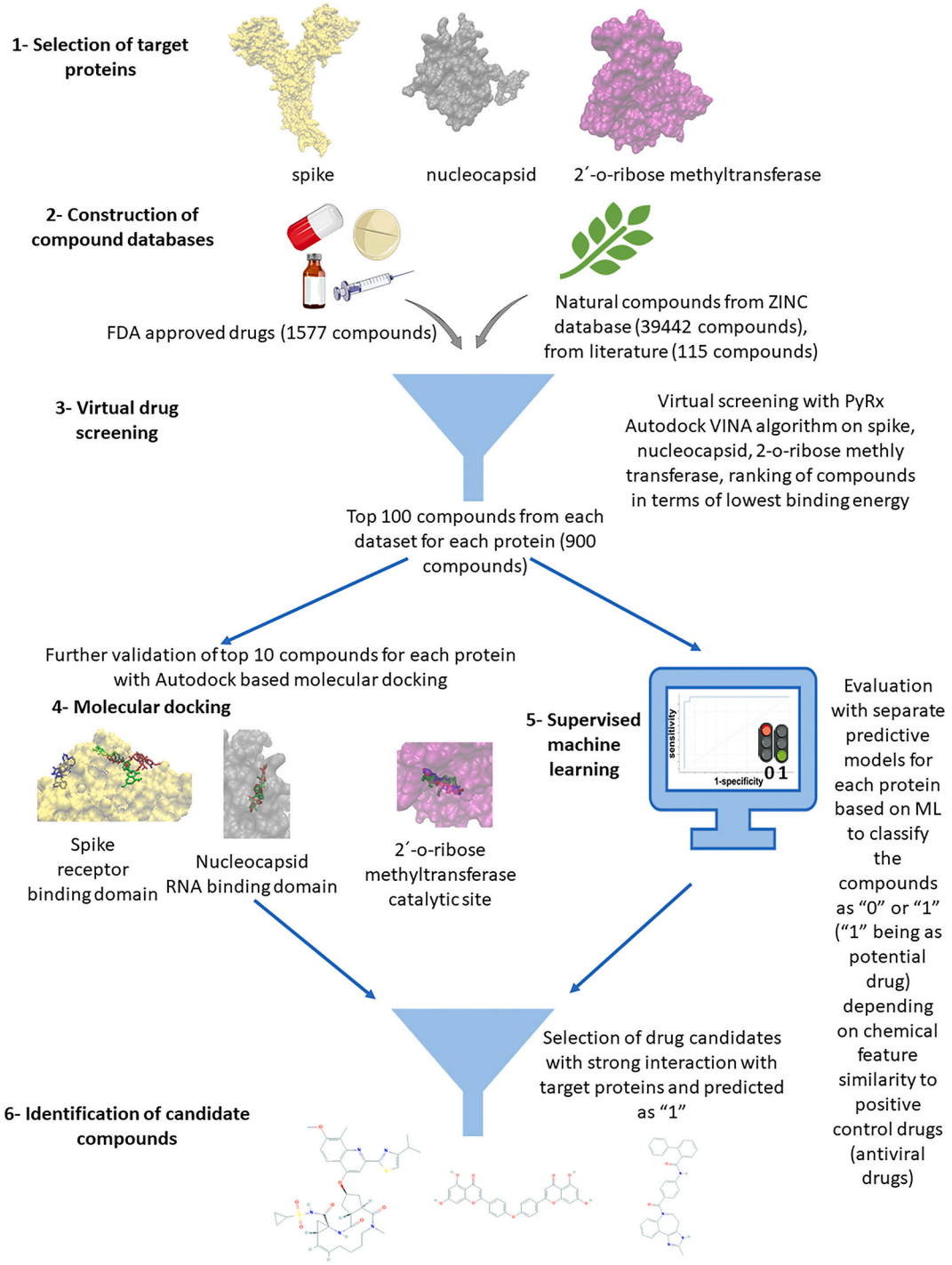

Fig. 1. Flowchart of the *in silico* strategy to identify drug candidates against SARS-CoV-2.

transmitted from these animals to human hosts, although other hypotheses have also been put forward.

Some coronaviruses (*e.g.* HCoV-229E, -NL63, -OC43, and -HKU1) usually cause respiratory infections and circulate worldwide in human populations [7]. Other coronavirus species (*e.g.* SARS-CoV, MERS-CoV, SARS-CoV-2) are rare and reveal higher mortality rates. In SARS-CoV-2 and MERS-CoV, more males than females are affected. Typical symptoms of SARS-CoV, SARS-CoV-2 and MERS-CoV include fever, dry cough, dyspnea, loss of tasting sense, muscle pain and other symptoms [8]. As of March 23, 2021, more than 124 million people were infected and more than 2.7 million deaths occurred. (https://www.worldometers.info/coronavirus/).

As of yet, there are no drugs to treat or prevent SARS-CoV-2. Some preliminary experiences with individual healing trials or animal experiments using anti-retroviral drugs (*e.g.* remdesivir, lopinavir, ritonavir, oseltamivir) and also alternative approaches from traditional Chinese medicine have been reported [9–12]. For instance, clinical trials are running and some of them have been already published for remdesivir, lopinavir and ritonavir [13,14]. The current clinical treatment is largely based on symptom-based therapies [11,15]. Therefore, strategies for the rapid identification of drug candidates are urgently required.

The concept of drug repurposing (or repositioning) came into the spotlight for several reasons [16]. As it became apparent that drugs approved for one disease, may also exert activity for other indications,

Table 1
Positive and negative control drugs to generate training and test sets for the supervised machine learning algorithms.

Training set			Test set		
Molecule Name	Class	LBE	Molecule Name	Class	LBE
Spike protein			**Spike protein**		
Atazanavir	1	−7.50	Indinavir	1	−8.20
Bevirimat	1	−7.20	Grazoprevir	1	−8.30
Calanolide A	1	−8.60	Elbasvir	1	−8.70
Capravirine	1	−7.00	Dolutegravir	1	−8.00
Cobicistat	1	−7.70	Delavirdine	1	−7.00
Lopinavir	1	−8.30	Darunavir	1	−7.90
Maraviroc	1	−8.20	Dapivirine	1	−8.20
Nelfinavir	1	−8.10	Daclatasvir	1	−8.70
Nevirapine	1	−7.10	Acetylcholine	0	−4.40
Ombitasvir	1	−8.80	Mechlorethamine	0	−3.40
Raltegravir	1	−7.50	Succinylcholine	0	−4.40
Rilpivirine	1	−7.30	Disulfiram	0	−3.80
Ritonavir	1	−8.10	Methimazole	0	−3.80
Saquinavir	1	−8.20	Dimercaprol	0	−3.50
Tipranavir	1	−7.70	Dalfampridine	0	−4.40
Velpatasvir	1	−9.80	Tolbutamide	0	−5.50
Acepromazine	0	−7.00	Naproxen	0	−6.90
Acetaminophen	0	−5.60	Mephentermine	0	−5.20
Acetylsalicylic acid	0	−6.00			
Amiodarone	0	−6.40			
Amphetamine	0	−5.50			
Bretylium	0	−5.50			
Captodiame	0	−6.10			
Carbachol	0	−4.10			
Cetylpyridinium	0	−5.30			
Choline	0	−3.90			
Colestipol	0	−4.60			
Dinoprostone	0	−4.10			
Dopamine	0	−5.60			
Etilefrine	0	−5.70			
Fluvoxamine	0	−5.80			
Ibuprofen	0	−6.40			
Loxoprofen	0	−6.70			
Methacholine	0	−4.40			
Methenamine	0	−4.80			
Orlistat	0	−4.30			
Nucleocapsid protein Training set			**Nucleocapsid protein Test set**		
Molecule Name	**class**	**LBE**	**Molecule Name**	**class**	**LBE**
Abacavir	1	−7.00	CalanolideA	1	−8.40
Bevirimat	1	−8.40	Cobicistat	1	−7.20
Capravirine	1	−8.50	Daclatasvir	1	−8.50
Darunavir	1	−7.70	Dapivirine	1	−7.90
Delavirdine	1	−8.00	Indinavir	1	−8.40
Dolutegravir	1	−7.70	Maraviroc	1	−8.20
Elbasvir	1	−8.60	Nelfinavir	1	−7.80
Grazoprevir	1	−7.70	Nevirapine	1	−7.60
Ombitasvir	1	−7.50	Acetylcholine	0	−3.80
Raltegravir	1	−7.60	Carbachol	0	−3.90
Remdesivir	1	−7.10	Cetylpyridinium	0	−4.60
Rilpivirine	1	−7.80	Choline	0	−3.30
Saquinavir	1	−9.40	Colestipol	0	−4.30
Suramin	1	−8.40	Dinoprostone	0	−6.60
Tipranavir	1	−7.80	Mechlorethamine	0	−3.60
Velpatasvir	1	−8.80	Methacholine	0	−4.00
Acepromazine	0	−6.50	Naproxen	0	−6.50
Acetaminophen	0	−4.90	Orlistat	0	−4.80
Acetylsalicylic acid	0	−5.10			
Amiodarone	0	−7.00			
Amphetamine	0	−5.40			
Bretylium	0	−4.90			
Captodiame	0	−5.90			
Dalfampridine	0	−4.10			
Dimercaprol	0	−2.80			
Disulfiram	0	−4.20			
Dopamine	0	−5.20			
Etilefrine	0	−5.30			
Fluvoxamine	0	−4.70			
Ibuprofen	0	−6.10			
Loxoprofen	0	−6.40			
Mephentermine	0	−5.20			
Methenamine	0	−3.90			
Methimazole	0	−3.70			
Succinylcholine	0	−4.20			
Tolbutamide	0	−6.60			
2′-o-ribose methyl transferase Training set			**2′-o-ribose methyl transferase Test set**		
Molecule Name	**class**	**LBE**	**Molecule Name**	**class**	**LBE**
Abacavir	1	−7.20	Elbasvir	1	−8.70
Atazanavir	1	−7.20	Dolutegravir	1	−9.00
Bevirimat	1	−9.80	Delavirdine	1	−8.90
Calanolide A	1	−8.50	Darunavir	1	−8.00
Capravirine	1	−7.10	Ritonavir	1	−8.10
Cobicistat	1	−8.20	Rilpivirine	1	−7.90
Daclatasvir	1	−9.70	Remdesivir	1	−7.60
Dapivirine	1	−8.30	Raltegravir	1	−10.30
Grazoprevir	1	−7.80	Ombitasvir	1	−10.00
Indinavir	1	−8.60	Acetylcholine	0	−4.00
Lopinavir	1	−7.40	Mechlorethamine	0	−3.30
Maraviroc	1	−8.40	Succinylcholine	0	−5.00
Nelfinavir	1	−7.60	Disulfiram	0	−4.00
Saquinavir	1	−9.30	Methimazole	0	−3.50
Suramin	1	−9.60	Dimercaprol	0	−3.00
Tipranavir	1	−8.90	Dalfampridine	0	−3.90
Velpatasvir	1	−9.20	Tolbutamide	0	−6.60
Zanamivir	1	−7.00	Naproxen	0	−6.90
Acepromazine	0	−6.20	Captodiame	0	−5.50
Acetaminophen	0	−5.50			
Acetylsalicylic acid	0	−6.00			
Amiodarone	0	−6.50			
Amphetamine	0	−4.60			
Bretylium	0	−4.90			
Carbachol	0	−4.20			
Cetylpyridinium	0	−4.10			
Choline	0	−3.30			
Colestipol	0	−4.60			
Dinoprostone	0	−5.50			
Dopamine	0	−5.80			
Etilefrine	0	−6.10			
Fluvoxamine	0	−6.20			
Ibuprofen	0	−6.20			
Loxoprofen	0	−6.90			
Mephentermine	0	−5.40			
Methacholine	0	−4.30			
Methenamine	0	−4.00			
Orlistat	0	−5.40			

1, positive control drug; 0, negative control drug.
LBE, lowest binding energy (kcal/mol).

FDA (https://www.fda.gov/) approved drugs became attractive as source for new drug development. A considerable advantage of old drugs in terms of time and costs for drug development is that their toxicity profile and pharmacokinetics are well-known in human beings. As the number of FDA-approved drugs is continuously decreasing during the past three decades, drug repurposing may speed up the marketing of new drugs. The dimension of drug development is, however, much broader in a sense that natural products (antibiotics, marine compounds, phytochemicals) represent a large chemical basis for drug development. Natural products serve as chemical scaffolds for derivatization to come up with novel compounds with improved pharmacological features. As a matter of fact, surveys of the National Cancer Institute, USA, repeatedly demonstrated that three quarters of drugs for all diseases worldwide during the past half century were in the one way or another based on natural resources [17,18]. Hence, chemical scaffolds from natural sources are indispensable for drug development.

Another dimension has been recently added by combining virtual drug screening methods with machine learning approaches for the

Table 2
Performance parameters of the established prediction models for spike protein, nucleocapsid protein, and 2′-*O*-ribose-methyltransferase.

	TP	TN	FP	FN	Sensitivity	Specificity	Overall predictive accuracy	Precision	AUC
Learning									
Spike protein (neural network)	16	19	1	0	1.000	0.950	0.972	0.941	0.994
Nucleocapsid protein (neural network)	15	19	1	1	0.938	0.950	0.944	0.938	0.997
2-o-ribose-methyltransferase (naïve bayes)	16	18	2	2	0.889	0.900	0.895	0.889	0.978
External validation									
Spike protein	8	10	0	0	1.000	1.000	1.000	1.000	
Nucleocapsid protein	8	10	0	0	1.000	1.000	1.000	1.000	
2-o-ribose-methyltransferase	9	10	0	0	1.000	1.000	1.000	1.000	

TP, true positive; TN, true negative; FP, false positive; FN, false negative; AUC, area under the curve.

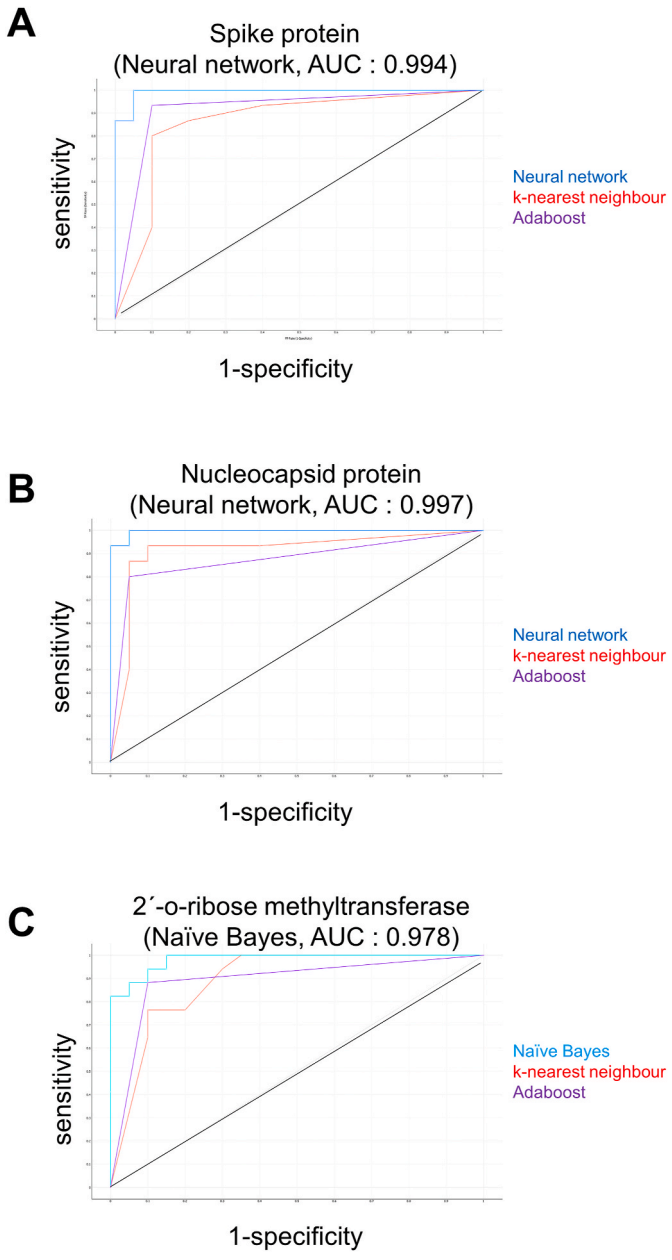

Fig. 2. Receiver operating characteristic (ROC) curves for spike protein (A), nucleocapsid protein (B), 2′-o-ribose-methyltransferase (C).

development of new drugs [19,20], overcoming multidrug resistance [21], and applications in precision medicine to select drugs for individualized therapies [22,23].

The aim of the present study was to identify candidate drugs using a combined approach of virtual drug screening, molecular docking and supervised machine learning techniques. For this purpose, we used a library of FDA-approved drugs to investigate their potential for repurposing as anti-SARS-CoV-2 drugs as well as two chemical libraries with natural products.

2. Materials and methods

2.1. Workflow

A flowchart of our *in silico* strategy to identify drug candidates against SARS-CoV-2 is shown in Fig. 1. The workflow consisted of 6 steps:

(1) Selection of target proteins: Homology modeling of target proteins: The amino acid sequence of the target proteins from SARS-CoV virus were translated into the sequences of the corresponding SARS-CoV-2 proteins. The available crystal structures of SARS-CoV spike protein (PDB: 5XLR), nucleocapsid protein (PDB: 2OFZ, 6IEX, 2CJR), and 2′-o-ribose methyltransferase (PDB: 3R24) were taken as templates to generate 3D homology models of the three SARS-CoV-2 proteins. Furthermore, the available crystal structures of SARS-CoV-2 spike protein ACE2 receptor binding domain (PDB ID: 7BZ5), nucleocapsid protein RNA binding domain (PDB ID: 6VYO), and 2′-o-ribose methyltransferase catalytic site (PDB ID: 7L6T) were retrieved from PDB database. The reported pharmacophores in the literature were verified as the receptor binding domain of spike glycoprotein, the RNA binding domain of nucleocapsid protein, and the catalytic site of 2′-o-ribose methyltransferase) [24–26].
(2) Construction of compound databases: (A) 1577 FDA-approved drugs (taken from ZINC database), (B) 39,442 natural products (taken from ZINC database) and (C) 115 natural products (taken from literature) were included in the study. Clinically established anti-viral drugs were chosen as presumable positive controls and clinically established drugs without antiviral activity were taken as presumable negative controls. All compounds were prepared in three-dimensional sdf format.
(3) Virtual drug screening: All compounds were subjected to PyRx AutoDock VINA (blind docking mode) to generate ranking lists with compounds binding with high affinity to the three target proteins of SARS-CoV-2.
(4) Molecular docking: The top 100 compounds from chemical libraries (A), (B) and (C) were analyzed for their ability to bind to the relevant pharmacophores of the three targets (ACE2 interaction site of spike protein, RNA-binding site of nucleocapsid protein and catalytic site of 2'o-ribose methyl transferase). Compounds with the best binding energies were then subjected to AutoDock VINA and AutoDock 4.2.6 (both in defined docking mode) to identify the amino acid residues involved in drug-

Table 3
Virtual screening (obtained by AutoDock VINA), molecular docking (obtained by AutoDock 4.2.6) results and ROC probability of compounds binding to spike protein. Top 10 compounds are shown, each from FDA-approved drugs, natural compounds taken from literature and ZINC database. Binding affinities are expressed as lowest binding energies (LBE) in kcal/mol obtained. The ROC probabilities are based on the model obtained from positive and negative control drugs. Those compounds are labeled in bold, where VINA and AutoDock revealed binding energies < -7 kcal/mol. Amino acid residues forming hydrogen bonds are labeled in bold.

Dataset	ROC probability	VINA defined	AutoDock defined	Interacting amino acid residues
		LBE	LBE	
FDA-approved drugs:				
Simeprevir	0.993	−8.73 ± 0.06	−10.09 ± 0.10	Asn334, Leu335, Cys336, Pro337, Phe338, **Gly339**, Phe342, Asn343, Cys361, Asp364, Val367, Leu368, Phe374
Paritaprevir	0.997	−9.27 ± 0.15	−10.04 ± 0.06	Tyr449, Leu452, Leu455, Phe456, Glu484, Tyr489, Phe490, Leu492, **Gln493**, **Ser494**, Tyr495, Gly496
Velpatasvir	0.999	−8.57 ± 0.12	−9.11 ± 0.09	Leu335, Phe338, Gly339, Phe342, Asn343, Val362, Asp364, Val367, Leu368, Ser371, Ser373, Phe374, Pro527, **Lys528**
Rifapentine	0.998	−8.80 ± 0.17	−8.84 ± 0.03	**Arg355,** Gly381, Val382, Phe392, Tyr396, Phe429, **Thr430,** Phe464, Ser514, **Phe515**, Glu516, Leu517, Leu518
Eribulin	0.999	−8.43 ± 0.15	−8.82 ± 0.08	Phe456, **Glu484**, Gly485, Phe486, Tyr489, Phe490
Teniposide	0.997	−8.63 ± 0.06	−8.46 ± 0.08	Leu335, Phe338, Gly339, Phe342, Asn343, **Asp364**, Leu368, Phe374, Trp436, Leu441
Trabectedin	0.995	−8.11 ± 0.10	−7.91 ± 0.04	Leu335, Cys336, Pro337, Phe338, Gly339, Phe342, **Asn343**, Ala363, **Asp364**, Val367, **Ser371**, Ser373, Phe374
Ivermectin	0.999	−9.07 ± 0.06	−7.49 ± 0.04	Trp353, Arg355, Tyr396, **Asp428**, Phe429, **Thr430**, Lys462, **Pro463**, Phe464, Glu465, Glu516
Ledipasvir	0.999	−8.53 ± 0.15	−7.31 ± 0.09	Tyr449, Leu452, **Glu484**, Gly485, Cys488, Phe490, Leu492, Gln493, Ser494
Nystatin	0.994	−8.53 ± 0.15	−6.83 ± 0.07	Pro426, **Asp427**, **Asp428**, **Arg457**, Ser459, Lys462, Pro463, **Glu465**, Arg466
Natural compounds from literature:				
Euphol	1.000	−7.93 ± 0.06	−9.13 ± 0.10	Arg454, Arg457, Lys458, **Asp467**, Ser469, Tyr473, Gln474, Ala475
Loniflavone	0.996	−9.23 ± 0.12	−8.59 ± 0.03	Leu335, Cys336, Phe338, **Val362**, Ala363, **Asp364,** Val367, Leu368, Ser371, Ser373, Phe374, Trp436, **Asn437**, Ser438, **Asn440,** Leu441
Amyrin	0.996	−9.03 ± 0.06	−8.46 ± 0.03	Arg355, Pro426, Asp428, Thr430, Phe464, Ser514, Phe515, Glu516
Procyanidin	0.999	−8.77 ± 0.15	−7.37 ± 0.02	Tyr396, Pro426, **Asp428**, Phe429, **Thr430, Pro463**, Phe464, Ser514, Phe515, Glu516, **Leu517**
Crinine	0.999	−7.77 ± 0.06	−6.84 ± 0.08	**Arg454**, Arg457, Lys458, Ser459, **Ser469**, Glu471, Tyr473, Pro491
Quercetin	0.993	−7.87 ± 0.06	−6.77 ± 0.09	**Asp467**, Ser469, Glu471, Tyr473, Gln474, Ala475, Pro491
IlexsaponinB2	1.000	−8.43 ± 0.06	−6.70 ± 0.15	**Ala372**, **Tyr369, Ser375, Cys379**, Tyr380, Gly381, Phe344
Strictinin	0.998	−7.83 ± 0.12	−6.26 ± 0.07	**Arg454, Phe456**, Lys458, **Arg466**, Asp467, **Ile468**, Ser469, **Ile472**, Tyr473
Quercetin-3-*o*-rutinoside	0.999	−8.23 ± 0.05	−5.12 ± 0.11	**Cys480**, Agr454, **Phe456**, Lys458, **Gly471**, Ile472, Tyr473, Gln474
Punicalagin	1.000	−8.11 ± 0.11	−4.22 ± 0.07	**Phe377, Lys378**, Cys379, Tyr380, **Gly381, Val382,** Ser383, Pro384, **Arg408**
Natural compounds from ZINC database:				
ZINC000252515584; (1R,3S,6S,7E,13S,16R,17R,21S,22S)-28-Hydroxy-17-[(2R,4R,5S,6R)-4-hydroxy-5-[(2S,4R,5R,6R)-5-hydroxy-4-(2-methoxy-6-methylbenzoyl)oxy-6-methyloxan-2-yl]oxy-6-methyloxan-2-yl]oxy-3,22-dimethyl-23,26-dioxo-24,27-dioxapentacyclo [23.2.1.01,6.013,22.016,21]octacosa-4,7,14,25 (28)-tetraene-4-carboxylic acid	0.999	−9.57 ± 0.09	−10.21 ± 0.08	Trp353, Asn354, Arg355, Pro426, **Asp428**, Thr430, Pro463, Phe464, Glu465, Arg466, Ser514
	0.997			

(*continued on next page*)

Table 3 (*continued*)

Dataset	ROC probability	VINA defined	AutoDock defined	Interacting amino acid residues
		LBE	LBE	
ZINC000027215482; (1R,4S,7S)-4-benzyl-9-[(1R,4S,7R)-4-benzyl-3,6-dioxo-2,5,16-triazatetracyclo [7.7.0.0^{2},^{7}.0^{1o},^{1s}]hexadeca-10,12,14-trien-9-yl]-2,5,16-triazatetracyclo [7.7.0.0^{2},^{7}.0^{1o},^{1s}]hexadeca-10 (15),11,13-triene-3,6-dione		−9.83 ± 0.06	−9.64 ± 0.09	Arg457, Lys458, Ser459, Asn460, Lys462, Ser469, Tyr473, **Gln474**, Gly476
ZINC000027215486; (1R,4S,7S)-4-benzyl-9-[(1R,4S,7S)-4-benzyl-3,6-dioxo-2,5,16-triazatetracyclo [7.7.0.0^{2},^{7}.0^{1o},^{1s}]hexadeca-10,12,14-trien-9-yl]-2,5,16-triazatetracyclo [7.7.0.0^{2},^{7}.0^{1o},^{1s}]hexadeca-10 (15),11,13-triene-3,6-dione	0.997	−9.38 ± 0.07	−8.82 ± 0.11	Phe342, Asn343, **Ala344**, Thr345, Ser371, Ser373, Phe374, Trp436, Asn440, **Arg509**
ZINC000095788030; 5-[(1R,3aS,3bR,7S,9aR,11aR)-7-{[(2R,3R,4R,5S,6R)-3,4-dihydroxy-6-(hydroxymethyl)-5-{[(2S,3R,4R,5R,6S)-3,4,5-trihydroxy-6-methyloxan-2-yl]oxy}oxan-2-yl]oxy}-3a-hydroxy-9a,11a-dimethyl-1H,2H,3H,3aH,3bH,4H,5H,7H,8H,9-H,9aH,9bH,10H,11H,11aH-cyclopenta [*a*]phenanthren-1-yl]-2H-pyran-2-one	0.993	−9.08 ± 0.08	−8.13 ± 0.10	Phe338, Gly339, **Phe342**, **Asn343**, Ala344, **Asp364**, Val367, Leu368, Ser371, Ser373, Phe374, Trp436, Arg509
ZINC000150343123	0.999	−8.70 ± 0.10	−7.82 ± 0.07	Arg454, Phe456, Arg457, Lys458, Ser469, **Glu471**, **Ile472**, Tyr473, Gln474, Pro479, Cys480, Asn481, Pro491
ZINC000255205550; (4aS,6aS,6bR,10S,12aR,14bS)-10-{[(2R,3R,4R,5S,6R)-3-acetamido-4,5-dihydroxy-6-({[(2R,3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy}methyl)oxan-2-yl]oxy}-2,2,6a,6b,9,9,12a-heptamethyl-1,2,3,4,4a,5,6,6a,6b,7,8,8a,9,10,11,12,12a,12b,13,14b-icosahydropicene-4a-carboxylic acid	0.998	−8.93 ± 0.06	−7.64 ± 0.07	Trp353, Arg355, Tyr396, Asp428, Phe429, **Thr430**, Phe464, Arg466, Ser514
ZINC000514287935; 6-[1-(9a,11a-dimethyl-9-oxo-7-)[3,4,5-trihydroxy-6-()[3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy)methyl)oxan-2-yl]oxy)-1H,2H,3H,3aH,3bH,4H,6H,7H,8H,9H,9aH,9bH,10H,11H,11aH-cyclopenta [*a*]phenanthren-1-yl)-1-hydroxyethyl]-3,4-dimethyl-5,6-dihydro-2H-pyran-2-one	0.999	−8.85 ± 0.05	−7.12 ± 0.09	Arg355, Tyr396, Asp428, Phe429, Thr430, **Arg457**, Pro463, Phe464, Glu465, Arg466, **Asp467**, **Ile468**, Ser514, Phe515
ZINC000253500823; 15'-[(5-{[3,4-dihydroxy-6-(hydroxymethyl)-5-{[3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy}oxan-2-yl]oxy}-3-hydroxy-4-methoxy-6-methyloxan-2-yl)oxy]-7′-hydroxy-8′,12′-dimethyl-6′-oxaspiro [oxolane-3,5′-pentacyclo [9.8.0.0^{1},^{7}.0^{4},^{8}.0^{12},^{17}]nonadecan]-5-one	0.999	−9.06 ± 0.10	−6.93 ± 0.06	Leu335, Phe338, Gly339, Phe342, Asn343, Asp364, Val367, Leu368, Ser371, **Ser373**, Phe374, Trp436, **Asn440**, Leu441
ZINC000253389471; (4aR,5R,6aS,6bR,10S,12aR)-10-{[(2R,3R,4R,5S,6R)-6-({[(2S,3R,4S,5S)-4,5-dihydroxy-3-{[(2S,3R,4S,5R)-3,4,5-trihydroxyoxan-2-yl]oxy}oxan-2-yl]oxy}methyl)-3-acetamido-4,5-dihydroxyoxan-2-yl]oxy}-5-hydroxy-2,2,6a,6b,9,9,12a-heptamethyl-1,2,3,4,4a,5,6,6a,6b,7,8,8a,9,10,11,12,12a,12b,13,14b-icosahydropicene-4a-carboxylic acid	0.999	−9.20 ± 0.10	−6.22 ± 0.12	Leu368, Tyr369, **Ser371**, **Ala372**, Phe374, Ser375, Phe377, Lys378, Cys379, Tyr380, Pro384
ZINC000253500685; 4-[(7S,9aS,11aR)-3a-hydroxy-7-[(4-methoxy-6-methyl-5-{[3,4,5-trihydroxy-6-({[3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy}methyl)oxan-2-yl]oxy}oxan-2-yl)oxy]-9a,11a-dimethyl-hexadecahydro-1H-cyclopenta [*a*]phenanthren-1-yl]-2,5-dihydrofuran-2-one	0.995	−9.67 ± 0.06	−5.94 ± 0.16	Gly381, Pro426, **Asp428**, **Thr430**, Ser514, Phe515, Glu516, Leu517

binding. 3D illustrations of drug-protein interactions were prepared using VMD.

(5) Study of drug-likeliness by supervised machine learning: The clinically established positive and negative control drugs (see step 2) were used to generate prediction models for drug-likeliness of test compounds based on 12 chemical descriptors. These predictions were applied to the top 100 compounds of libraries (A), (B), and (C).

(6) Identification of candidate compounds: Compounds with lowest binding energies of $<$ -7 kcal/mol (from step 4) and probability values of R $>$ 0.995 (from step 5) were proposed as candidate compounds with activity against SARS-CoV-2.

2.2. Virtual screening with AutoDock VINA

Three sets of compounds were considered for the virtual screening on three proteins (spike protein, nucleocapsid protein, and 2′-o-ribose methyltransferase). FDA-approved drugs (1577 compounds), natural compounds from the ZINC database (39,442 compounds), and natural compounds mined from the literature with antiviral activity (115 compounds) [27–31]. Furthermore, antiviral drugs were selected as presumable positive control drugs (27 compounds) and non-cytotoxic antidiabetic, antidepressants, cardiovascular agents, non-steroidal anti-inflammatory drugs (NSAIDs) and proton pump inhibitors were selected as presumable negative control drugs (30 compounds) from the DrugBank database (https://www.drugbank.ca/). As described before, the threshold was set as −7 kcal/mol to consider the affinity of a chemical compound to its target protein as being strong [32]. The positive control drugs revealed binding energies of $\leq$ -7 kcal/mol, while negative control drugs bound with affinities of $>$ -7 kcal/mol to the three targets (Table 1). The test compounds have been subjected to an automated and comprising molecular docking campaign by using the AutoDock VINA algorithm based PyRx software (https://pyrx.sourceforge.io/) (blind docking mode) and the high-performance supercomputer MOGON (Johannes Gutenberg University, Mainz).

2.3. Molecular docking

After the selection of compounds with strong interaction with target proteins, further validation was performed with molecular docking. For this purpose, the Lamarckian algorithm of AutoDock VINA was chosen (defined docking mode), and the AutoDock 4.2.6 (http://autodock.scripps.edu/). Lamarckian algorithm was used to analyze the docking poses and binding energies as described before [21,33]. The ligand moved around the rigid protein with 250 runs and 25,000,000 energy evaluations for each cycle. The amino acids of the target proteins binding to the ligands were also determined by AutoDock 4.2.6. Compound-protein interactions were visualized by using the Visual Molecular Dynamics (VMD) software. (Theoretical and Computational Biophysics group at the Beckman Institute, University of Illinois at Urbana-Champaign, IL, USA) (https://www.ks.uiuc.edu/Research/vmd/).

2.4. Supervised machine learning

In order to build separate predictive models for each protein to identify potential drugs against SARS-CoV-2 and considering recent clinical reports that some COVID-19 patients were treated with antiviral drugs [34–37], we used the above mentioned presumable positive control and negative control drugs. After random selection was applied,

Table 4
Virtual screening (obtained by AutoDock VINA), molecular docking (obtained by AutoDock 4.2.6) results and ROC probability of compounds binding to nucleocapsid protein. Details see Table 3.

Dataset	ROC probability	VINA defined LBE	AutoDock defined LBE	Interacting amino acid residues
FDA-approved drugs:				
paritaprevir	0.999	−11.30 ± 0.06	−13.50 ± 0.40	**Gly170, Ala173,** Leu161, Gln163, Thr165, Pro168, Lys169, Phe171, Tyr172
trypan blue	0.999	−10.80 ± 0.05	−12.17 ± 0.07	Gln70, Val72, Ile74, **Leu161**, Pro162, Gln163, **Thr165**, Leu167
simeprevir	0.999	−10.60 ± 0.1	−11.58 ± 0.50	Glu62, Leu161, Gln163, Thr165, **Leu167**, Pro168, Lys169, Phe171, Ala173
dihydroergotamine	0.999	−10.50 ± 0.1	−11.46 ± 0.20	**Ala173**, Leu161, Pro162, Gln163, Leu167, Phe171
conivaptan	0.995	−10.90 ± 0.06	−10.795 ± 0.3	Pro73, Ile74, Asn75, Thr76, Ser78, Pro80, Gln83
ergotamine	0.993	−10.50 ± 0.2	−10.61 ± 0.50	Leu161, Pro162, Thr165, Leu167, Phe171
venetoclax	0.995	−10.60 ± 0.1	−10.50 ± 0.60	**Leu167**, Pro162, **Gln163,** Thr165, Thr166, Tyr17, Ala173
idarubicin	0.990	−10.10 ± 0.1	−8.35 ± 0.30	**Gly69, Ala134**, Phe66, Arg68, Tyr123, Trp132, Val133, **Glu136**
ivermectin	0.999	−10.60 ± 0.1	−8.33 ± 0.40	**Arg89**, Gln58, Lys61, Arg88, Thr91, Asp128, Gly129, Ile130
nystatin	0.999	−10.30 ± 0.1	−7.45 ± 0.60	**Thr76**, Ile74, Gln160, Pro162, Gln163, Tyr172, Ala173
Natural compounds from literature:				
ilexsaponinB1	0.999	−9.40 ± 0.20	−9.55 ± 0.30	**Asn75**, Thr135, Pro73, Pro163, Gln163, Leu167
ilexsaponinB2	1.000	−9.20 ± 0.10	−9.50 ± 0.50	**Asn75,** Pro162, Gln163, Gly164, Leu167, Tyr172, Ala173
procyanidin	0.999	−8.90 ± 0.10	−8.47 ± 0.60	**Gln163**, Leu161, Gly164, Leu167, Phe171, Tyr172, **Ala173**
crinine	0.999	−8.80 ± 0.10	−7.74 ± 0.30	**His59, Thr54, Asn77,** Val158, Ser78
strictinin	0.999	−9.40 ± 0.10	−7.31 ± 0.30	**Asn75, Thr165**, Gln163, Gly164, Leu167, Tyr172, **Ala173**
ilexsaponinB3	1.000	−8.60 ± 0.10	−6.54 ± 0.60	**Gly164, Asn75**, Ile74, Pro73, Thr76, Ser78, Pro80, Gln83
rutin	0.999	−9.10 ± 0.10	−6.34 ± 0.10	**Gly170,** Pro162, Gln163, Gly164, Leu167, Tyr172, Ala173
forsythiaside	0.999	−9.20 ± 0.10	−4.56 ± 0.20	**Tyr123**, Leu113, Gly114, Ala119, Gly120, Pro122
punicalagin	1.000	−9.50 ± 0.10	−4.54 ± 0.30	**Leu167, Gln163,** Gly164, Thr165, **Thr166**, Tyr172, Ala173
tirucallinA	1.000	−9.60 ± 0.10	−4.44 ± 0.60	**Leu167, Gln163,** Gly164, **Thr165**, Thr166, Tyr172, **Ala173**
Natural compounds from ZINC database:				
ZINC000027215482; (1R,4S,7*S*)-4-benzyl-9-[(1R,4S,7S)-4-benzyl-3,6-dioxo-2,5,16-triazatetracyclo [7.7.0.0^{2},^{7}.0^{1o},^{1s}]hexadeca-10,12,14-trien-9-yl]-2,5,16-triazatetracyclo [7.7.0.0^{2},^{7}.0^{1o},^{1s}]hexadeca-10 (15),11,13-triene-3,6-dione	1.000	−10.47 ± 0.12	−10.89 ± 0.03	Thr76, Ser78, Ser79, Pro80, Leu161, Pro162, **Gln163**, Gly164, Thr165, Leu167, Phe171, Tyr172, Ala173
ZINC000252515584; (1R,3S,6S,7E,13S,16R,17R,21S,22*S*)-28-hydroxy-17-[(2R,4R,5S,6R)-4-hydroxy-5-[(2S,4R,5R,6R)-5-hydroxy-4-(2-methoxy-6-methylbenzoyl)oxy-6-methyloxan-2-yl]oxy-6-methyloxan-2-yl]oxy-3,22-dimethyl-23,26-dioxo-24,27-dioxapentacyclo [23.2.1.01,6.013,22.016,21] octacosa-4,7,14,25 (28)-tetraene-4-carboxylic acid	1.000	−11.00 ± 0.10	−10.87 ± 0.05	**Asn75**, Thr76, Gln163, **Gly164**, Thr166, **Leu167**, Tyr172
ZINC000004149988; 2-(2,2,2-trifluoroacetyl)-*N*-[(1S,2S,5S,6S,9S,16S)-5,9,16-trimethyl-4-oxo-3-oxa-12-thia-14-azatetracyclo [7.7.0.0^{2},^{6}.0^{11},^{1s}]hexadeca-11 (15),13-dien-13-yl]-1,2,3,4-tetrahydroisoquinoline-7-sulfonamide	0.999	−10.20 ± 0.10	−10.30 ± 0.05	Gly69, Gln70, Gly71, **Val72**, Pro73, Ile74, Asn75, Thr76, Pro80, Gln83, Thr135, Pro162, Gln163, Gly164
ZINC000004097766; Albanol A	0.999	−10.33 ± 0.06	−10.16 ± 0.05	Ile74, **Asn75**, Ser78, Gln83, Leu161, **Pro162**, Thr165, **Leu167**, Phe171, Tyr172, Ala173
ZINC000004098521; Hinokiflavone	0.995	−10.57 ± 0.06	−9.84 ± 0.05	**Gly69**, Gln70, Val72, **Asn75**, Thr76, Ser78, Pro80, Gln83, Thr135, Pro162, **Leu167**, Phe171, Tyr172, Ala173
ZINC000005433649; (3R,3aS,6S,6aR)-6-[(4-{[1,1′-biphenyl]-4-yl}pyrimidin-2-yl)amino]-hexahydrofuro [3,2-*b*]furan-3-yl *N*-(naphthalen-1-yl)carbamate	0.999	−10.88 ± 0.08	−9.39 ± 0.07	Val72, Pro73, Ile74, Asn75, Thr76, Gln83, Thr135, Leu161, **Gln163**, Thr165, Leu167, Ala173
ZINC000253504770; 4-(7-)[5-()5-[(4,5-dihydroxy-6-methyloxan-2-yl)oxy]-4-hydroxy-6-methyloxan-2-yl)oxy)-4-hydroxy-6-methyloxan-2-yl]oxy)-3a,11-dihydroxy-9a,11a-dimethyl-hexadecahydro-1H-cyclopenta [*a*]phenanthren-1-yl)-2,5-dihydrofuran-2-one	0.999	−10.37 ± 0.12	−9.17 ± 0.06	Gly69, Val72, Ile74, Asn75, Thr76, Gln83, Thr135, Leu159, **Leu161**, Pro162, **Leu167**
ZINC000015675926; 5-[(6S)-5-(4-fluorobenzoyl)-1H,4H,5H,6H,7H-imidazo [4,5-*c*]pyridin-6-yl]-3-(naphthalen-2-yl)-1,2,4-oxadiazole	0.994 0.999	−10.95 ± 0.05	−9.06 ± 0.06 −8.62 ± 0.04	Asn75, Thr76, Ser78, Gln83, Leu161, Gln163, Thr165, Leu167, Tyr172, Ala173

(*continued on next page*)

Table 4 (*continued*)

Dataset	ROC probability	VINA defined LBE	AutoDock defined LBE	Interacting amino acid residues
ZINC000253504766; 4-(7-)[5-()5-[(4,5-dihydroxy-6-methyloxan-2-yl)oxy]-4-hydroxy-6-methyloxan-2-yl)oxy)-4-hydroxy-6-methyloxan-2-yl]oxy)-3a,11-dihydroxy-9a,11a-dimethyl-hexadecahydro-1H-cyclopenta [*a*]phenanthren-1-yl)-2,5-dihydrofuran-2-one		−10.80 ± 0.10		Gln70, **Val72**, Ile74, Asn75, Thr76, Pro80, Gln83, Thr135, Pro162, Gln163, Gly164, Thr165, Leu167, Tyr172, **Ala173**
ZINC000005434062; (3R,3aS,6S,6aR)-6-{[4-(3-fluorophenyl)pyrimidin-2-yl] amino}-hexahydrofuro [3,2-*b*]furan-3-yl *N*-(naphthalen-1-yl)carbamate	0.998	−10.67 ± 0.12	−8.39 ± 0.05	Ile74, Thr76, **Asn77**, Ile157, **Gln160**, Pro162, Gln163, Thr165, Ala173

16 positive control, 20 negative control drugs were used for the spike protein learning set. For the external validation step, 8 positive control and 10 negative control drugs were used (Table 1). For the nucleocapsid protein learning set, 16 positive control, 20 negative control drugs were used. For the external validation step, 8 positive control and 10 negative control drugs were used. For the 2′-o-ribose methyltransferase learning set, 18 positive control, 20 negative control drugs were used. For the external validation step, 9 positive control and 10 negative control drugs were used.

The positive control drug class was labeled as "1" and the negative control drug class was labeled as "0". After the descriptors were calculated by Data Warrior software, the descriptors were selected in a similar manner, as previously reported by us using the SPSS software and considering the correlations of each descriptor with the class (0/1) [21]. After calculation of the 32 chemical descriptors, correlation coefficients between descriptors and correlation of the descriptors with the class (1/0) (potential drug; yes or no) were determined using SPSS statistics software version 23.0.0.3 (IBM, Armonk, NY: IBM Corp, USA). If the correlation with the class (1/0) (potential drug; yes or no) was below 0.1, this descriptor was omitted. Only descriptors correlating with the class (1/0) (potential drug; yes or no) category above 0.1 were selected for further processing. As a next step, descriptors having a pairwise correlation coefficient higher than 0.9 were excluded. By this strategy, relevant descriptors without an issue of over-fitting can be selected. Leave-one-out random sampling was used to build the models. Correlation matrix approach is among the preferred feature selection techniques. By applying the above-mentioned correlation matrix approach, we could eliminate overfitting and select only the relevant descriptors which are positively correlated with the target variable (potential drug (1/0) classification).

The selected descriptors meeting the criteria were as follows: H-acceptors, H-donors, total surface area, relative PSA, molecular complexity, rotatable bonds, ring closures, aromatic atoms, sp3 atoms, symmetric atoms, amides, and aromatic nitrogens. To select the most suited algorithm, we used the Orange software (Ljubljana, Slovenia) (https://orange.biolab.si/). We tested all 11 different algorithms and found that neural network performed better than the other algorithms for nucleocapsid protein and spike protein models, whereas naïve bayes was the best algorithm for 2′-o-ribose methyltransferase model. The performance parameters for each model are summarized in Table 2. The top 100 compounds based on lowest binding energy (LBE) from each virtual screening output on three proteins were selected to evaluate their classes with our prediction model. The receiver operating characteristic (ROC) curves of 3 out of 11 algorithms are depicted in Fig. 2.

2.5. Molecular dynamics

The final docking pose of loniflavone on the spike protein receptor binding domain (PDB ID: 7BZ5) was used for molecular dynamics (MD) simulation with the PlayMolecule software (https://www.playmolecule.org/). At first, parameterization of the ligand was performed with the Parameterize function [38,39]. The spike protein receptor binding domain was prepared for MD simulation with the ProteinPrepare function [40]. The protein-ligand complex was formed with the System-Builder function. MD simulation (with maximum available duration; 8 ns equilibration, 15 ns simulation) was performed with the SimpleRun function [41].

3. Results

After establishing the prediction models for spike protein, nucleocapsid protein, and 2′-o-ribose-methyltransferase using the positive and negative control drugs (Table 1) and virtual drug screening using AutoDock VINA, the top 100 compounds binding to each of the three protein models were selected for further analysis (top 100 from ZINC, top 100 from FDA and top 100 from literature compounds).

At the beginning of our analyses, crystal structures of the 3 target proteins were not available. At the end of the calculations, the corresponding crystal structures have been published. Therefore, we validated our homology model-based calculations with crystal structure-based calculations by using the top 100 FDA-approved drugs and the top 100 natural compounds taken from literature and performed AutoDock VINA virtual screening. We found correlation values of $R = 0.897$ for the spike protein, $R = 0.855$ for the nucleocapsid protein, and $R = 0.906$ for the 2′-*O*-ribose methyltransferase, indicating that the homology models provided reliable results.

We then evaluated their therapeutic probability against SARS-CoV-2 by using our established prediction models with positive and negative control drugs. The compounds were ranked according to their binding energy (yielded from the AutoDock VINA-based virtual screening in blind docking mode). We selected the top 10 compounds from each dataset for each protein model and considered a probability threshold of $R > 0.995$.

Then, these 10 compounds from each dataset were subjected to two further molecular docking programs for verification. PyRx implemented in AutoDock VINA allowed rapid screening in the blind docking mode, *i.e.* the best docking pose on the entire target protein surface was investigated. As a next step, we applied two defined docking modes (AutoDock VINA and AutoDock 4.2.6) based on the Lamarckian algorithm. Here, we defined the docking position at the sites, which are relevant for protein function, *i.e.* the ACE2-interaction site of the spike protein, the RNA-binding site of the nucleocapsid protein, and the catalytic site of 2′-*O*-ribose methyltransferase. In addition, we also identified the amino acid residues involved in compound binding within the defined binding domains. The results for the 10 best compounds of each dataset (FDA-approved drugs, natural compounds selected from literature and ZINC database) are shown in Tables 3–5.

Those compounds which consistently passed binding energy thresholds of < -7 kcal/mol with all three programs (2 × AutoDock VINA and AutoDock 4.2.6) may be considered more suitable for further investigations than the other compounds (Tables 3–5).

In parallel, these sets of each 10 compounds were subjected to supervised machine learning to gain insight into the drug-likeliness of the compounds (ROC probability of being class "1" yielded from the prediction models). Eleven different algorithms available in the Orange software were tested for building the prediction models. The neural network algorithm was the best for the spike and nucleocapsid proteins, while naïve bayes was superior for 2′-o- ribose methyltransferase. Fig. 2 displays 3 out of 11 tested algorithms for illustration. With these prediction models, the test compounds were calculated, and excellent ROC

Table 5
Virtual screening (obtained by AutoDock VINA), molecular docking (obtained by AutoDock 4.2.6) results and ROC probability of compounds binding to 2′-o-ribose-methyltransferase. Details see Table 3.

Dataset	ROC probability	VINA defined LBE	AutoDock defined LBE	Interacting amino acid residues
FDA approved drugs:				
Conivaptan	1.000	−10.83 ± 0.06	−10.94 ± 0.38	Trp6803, Gln6804, Gln6850, Asn6853, Tyr7040, Ser7041, Asp7044
Lifitegrast	1.000	−10.57 ± 0.11	−10.88 ± 0.23	Trp6803, Asn6853, Tyr7040, Lys7047, Lys7051
Dihydroergotamine	1.000	−12.17 ± 0.05	−9.56 ± 0.41	Ser6903, Ala6905, Asp6906, Ser6907, Ser7090, Val7092
Ergotamine	1.000	−12.30 ± 0.10	−9.46 ± 0.07	Leu6892, Ala6905, Asp6906, Ser6907, Thr6908, Val7086, Val7087
Eltrombopag	0.999	−11.03 ± 0.05	−9.19 ± 0.34	**Lys6933**, Thr6934, Asn6936, Lys6939, Lys6944, Glu6945
Ponatinib	1.000	−10.47 ± 0.12	−9.17 ± 0.29	**Ser7039**, Trp6803, Thr6833, Leu7037, Tyr7040, Phe7043
Lumacaftor	0.999	−11.13 ± 0.05	−8.75 ± 0.08	**His6972**, Lys6933, Thr6934, Asn6936, Lys6939, Ser6943, Glu6945
Nilotinib	1.000	−11.30 ± 0.10	−8.43 ± 0.34	Asp6931, Lys6939, Ser6943, Lys6944, **Glu6945**, Ser6973, Tyr7020
Regorafenib	1.000	−10.50 ± 0.10	−7.98 ± 0.33	**Tyr7040**, Trp6803, Gln6850, Asn6853, Ser7041, Lys7047
Aprepitant	1.000	−10.63 ± 0.05	−6.25 ± 0.42	Ser6800, Trp6803, Gln6804, Pro6805, Gln6850, Asp7044
Natural compounds from literature:				
Loniflavone	1.000	−10.80 ± 0.10	−9.69 ± 0.09	Asp6897, **Lue6898, Cys6913**, Met6929, **Tyr6930,** Asp6931, Ser6943, Glu6945
Friedelin	1.000	−9.73 ± 0.06	−8.95 ± 0.02	Thr6833, Leu7037, Ser7039, Tyr7040, Phe7043
TingeninB	0.999	−10.47 ± 0.06	−8.75 ± 0.10	Tyr6803, Gln6804, Gln6850, Tyr7040, Asp7044, Lys7047
Hoslunddiol	1.000	−9.60 ± 0.10	−7.12 ± 0.09	**Tyr7040,** Ser6800, Trp6803, Gln6804, Pro6805, **Asp7044**, Ser7041
Wogonoside	1.000	−9.57 ± 0.11	−6.18 ± 0.16	Asp6931, Thr6934, Lys6939, Sdp6942, Ser6943, Glu6945, Gly6946
Procyanidin	1.000	−9.70 ± 0.10	−5.98 ± 0.49	**Asn6936**, Asp6931, Lys6933, Val6937, Thr6938, Glu6940, **Lys6939**
Baicalin	1.000	−9.80 ± 0.10	−5.95 ± 0.22	**Lys7075**, Leu7010, Arg7014, Ser7074
IlexsaponinB2	1.000	−9.47 ± 0.12	−4.25 ± 0.03	**Lys7075**, Pro6810, **Asn6811**, Tyr6851, Thr6854, Leu6855, Trp6987, **Lys7075**
Punicalagin	1.000	−9.43 ± 0.15	−2.45 ± 0.04	**Ile7088**, Leu6892, Asp6906, Ser6907, Thr6908, **Arg7085,** Val7087
TirucallinA	1.000	−10.17 ± 0.06	0.25 ± 0.06	Asn6862, **Arg6864**, Leu6892, Asp6906, Ser6907, **Val7086**
Natural compounds from ZINC database:				
ZINC000015675938; (6S)–*N*-(4-chlorophenyl)-6-[3-(naphthalen-2-yl)-1,2,4-oxadiazol-5-yl]-1H,4H,5H,6H,7H-imidazo [4,5-*c*]pyridine-5-carboxamide	1.000	−12.03 ± 0.21	−12.04 ± 0.07	Asp6931, Thr6934, Asn6936, Val6937, Thr6938, Lys6939, Glu6940, Asp6942, Lys6944, Glu6945, Gly6946, Phe6947, Ser6973
ZINC000004258894; *N*-[(1R,9S)-11-{[1,1′-biphenyl]-4-carbonyl}-6-oxo-7,11-diazatricyclo [7.3.1.0^{2,7}]trideca-2,4-dien-5-yl]-2-phenylacetamide	1.000	−12.09 ± 0.10	−11.86 ± 0.05	Thr6833, Ser7038, Ser7039, Tyr7040, Ser7041, Phe7043, Asp7044
ZINC000004259861; 3-[(3S,3aR,6S,6aR)-6-{5-[4-(1H-1,2,4-triazol-1-yl) phenoxy]-1H-1,2,3,4-tetrazol-1-yl}-hexahydrofuro [3,2-*b*]furan-3-yl]-1-(4-phenoxyphenyl)urea	1.000	−12.11 ± 0.09	−11.84 ± 0.07	Trp6803, Asp6830, Ser6831, Thr6833, Leu7037, Ser7038, Ser7039, Tyr7040, Phe7043
ZINC000004259794; *N*-[(3S,3aR,6S,6aR)-6-[5-(2H-1,3-benzodioxol-5-yloxy)-1H-1,2,3,4-tetrazol-1-yl]-hexahydrofuro [3,2-*b*]furan-3-yl] naphthalene-2-sulfonamide	1.000	−12.13 ± 0.12	−11.31 ± 0.09	Trp6803, Ala6832, **Thr6833**, Leu7037, Ser7038, Ser7039, Tyr7040, Phe7043
ZINC000008635475; *N*-[4-()[(2R,4S,5R)-5-[1-methyl-3-(naphthalen-2-yl)-1H-pyrazol-5-yl]-1-azabicyclo [2.2.2]octan-2-yl]methyl)sulfamoyl) phenyl]acetamide	1.000	−12.27 ± 0.15	−11.24 ± 0.06	Lys6939, **Glu6940**, Asp6942, Ser6943, Lys6944, Glu6945, Asp7018, Tyr7020
ZINC000004259775; 3-[(3S,3aR,6S,6aR)-6-[5-(2H-1,3-benzodioxol-5-yloxy)-1H-1,2,3,4-tetrazol-1-yl]-hexahydrofuro [3,2-*b*]furan-3-yl]-1-(naphthalen-1-yl)urea	1.000	−12.37 ± 0.21	−11.16 ± 0.13	Gly6869, Gly6871, **Asp6897**, **Asp6912**, Cys6913, Asp6928, **Tyr6930**, Asp6931, Lys6944, Glu6945, Gly6946, Phe6947
ZINC000253407092; 4-[3-((4,5-bis [(4,5-dihydroxy-6-methyltetrahydro-2H-pyran-2-yl)oxy]-6-methyltetrahydro-2H-pyran-2-yl)oxy)-14,16-dihydroxy-10,13-dimethylhexadecahydro-1H-cyclopenta [*a*]phenanthren-17-yl]-2(5H)-furanone	1.000	−12.40 ± 0.30	−11.13 ± 0.10	Gly6869, Ser6872, **Asp6897**, Leu6898, Asn6899, Met6929, Tyr6930, Asp6931, Pro6932, **Lys6933**, Thr6934, Ser6943, Lys6944, Glu6945, Gly6946, Phe6947
ZINC000253504766; 3-[(3S,5S,8S,9R,10S,12R,13S,14S,17S)-3-[(2S,4R,5R,6R)-5-[(2R,4R,5R,6R)-5-[(2S,4R,5S,6R)-4,5-Dihydroxy-6-methyloxan-2-yl]oxy-4-hydroxy-6-methyloxan-2-yl]oxy-4-hydroxy-6-methyloxan-2-yl]oxy-12,14-dihydroxy-10,13-dimethyl-1,2,3,4,5,6,7,8,9,11,12,15,16,17-tetradecahydrocyclopenta [*a*] phenanthren-17-yl]-2H-furan-5-one	1.000	−12.43 ± 0.15	−10.82 ± 0.60	Ser6872, Asp6873, Asp6897, Leu6898, **Asn6899**, Met6929, Tyr6930, Asp6931, Pro6932, Lys6939, Lys6944, Glu6945, Gly6946, **Phe6947**, Phe6948
	1.000			

(*continued on next page*)

Table 5 (*continued*)

Dataset	ROC probability	VINA defined LBE	AutoDock defined LBE	Interacting amino acid residues
ZINC000253504770; 4-(7-)[5-()5-[(4,5-dihydroxy-6-methyloxan-2-yl)oxy]-4-hydroxy-6-methyloxan-2-yl)oxy)-4-hydroxy-6-methyloxan-2-yl]oxy)-3a,11-dihydroxy-9a,11a-dimethyl-hexadecahydro-1H-cyclopenta [*a*] phenanthren-1-yl)-2,5-dihydrofuran-2-one		−12.47 ± 0.12	−10.64 ± 0.16	**Tyr6828**, Gly6829, Asp6830, **Ser6831**, Thr6833, Leu7037, Ser7039, Tyr7040, Phe7043
ZINC000253500684; 4-[(7S,9aS,11aR)-3a-hydroxy-7-[(4-methoxy-6-methyl-5-{[3,4,5-trihydroxy-6-({[3,4,5-trihydroxy-6-(hydroxymethyl) oxan-2-yl]oxy}methyl)oxan-2-yl]oxy}oxan-2-yl)oxy]-9a,11a-dimethyl-hexadecahydro-1H-cyclopenta [*a*]phenanthren-1-yl]-2,5-dihydrofuran-2-one	1.000	−11.98 ± 0.24	−8.23 ± 0.10	Ala6808, Tyr6851, Leu6855, Thr6856, Trp6987, Leu7010, Asp7067, Leu7070, Ser7071, Ser7074, Lys7075

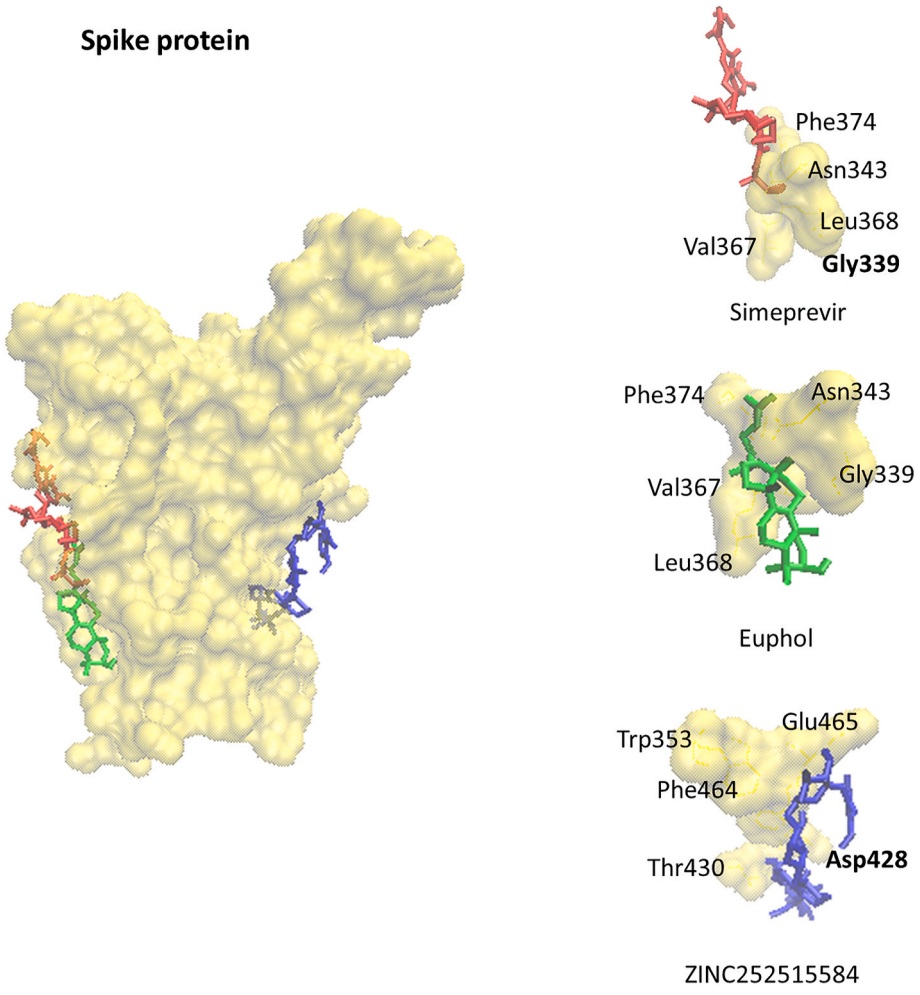

Fig. 3. Docking poses of simeprevir (red), euphol (green) and ZINC252515584 (blue) on spike protein (yellow). Residues forming hydrogen bonds are labeled bold.

probabilities were obtained (Tables 3–5), indicating that the test compounds fulfilled the criteria of drug-likeliness defined by the 12 chemical parameters setting up the predictive models.

Interestingly, among the drugs binding with high affinity to the spike protein were several approved drugs against another enveloped (+) ssRNA virus, the hepatitis C virus (HCV), *i.e.* paritaprevir, simeprevir and velpatasvir), indicating that these drugs may also be effective to treat COVID-19.

Interestingly, some of the compounds shown in Tables 3–5 bound with high affinity not only to one target protein but also to another one. Among the FDA-approved drugs, ivermectin, nystatin, paritaprevir and simeprevir bound to spike protein and nucleocapsid, conivaptan, dihydroergotamine and ergotamine to nucleocapsid protein and 2′-*O*-ribose methyltransferase. Among the natural products, crinine, ilexsaponinB2, procyanidin, punicalagin, strictinin, ZINC000027215482 and ZINC000252515584 bound to spike protein and nucleocapsid, while loniflavone, ilexsaponin B2, procyanidin, punicalagin bound to spike protein and 2′-o-ribose methyltransferase, ilexsaponin B2, procyanidin, punicalagin, tirucallin A, ZINC000253504770 and ZINC000253504766 bound to nucleocapsid protein and 2′-*O*-ribose methyltransferase. These "two-in-one" compounds may be attractive for further drug development.

Finally, as a conclusion from virtual screening, molecular docking and supervised machine learning the top compounds were identified. The target interactions (1) with the spike protein were highest for simeprevir, euphol and ZINC252515584, (2) with the nucleocapsid

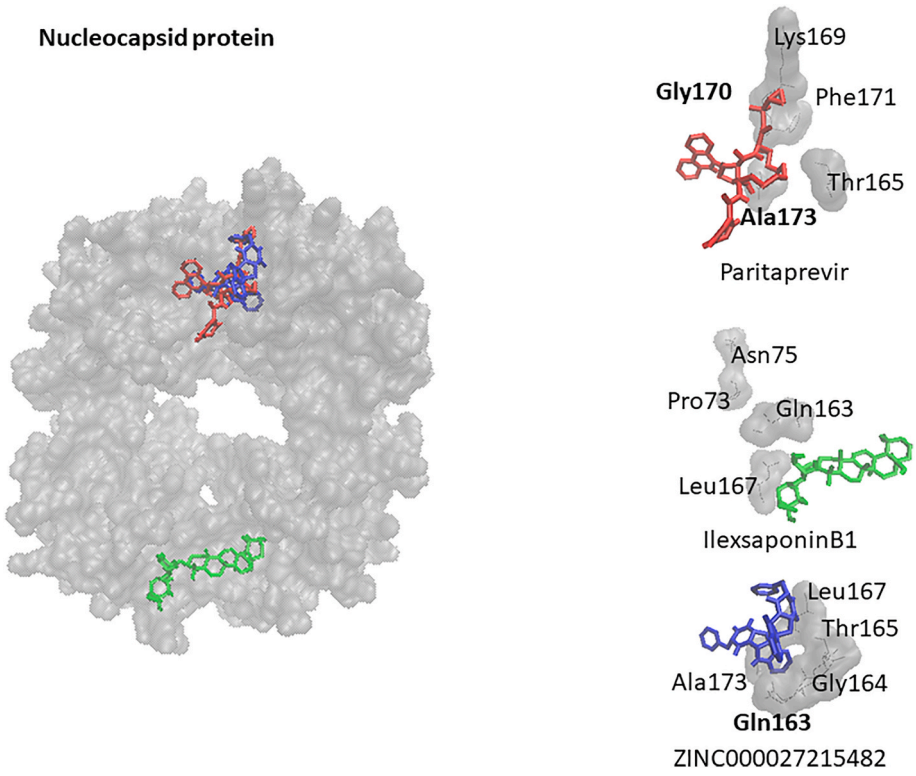

Fig. 4. Docking poses of paritaprevir (red), ilexsaponin B1 (green) and ZINC27215482 (blue) on nucleocapsid protein (gray). Residues forming hydrogen bonds are labeled bold.

protein for paritaprevir, ilexsaponin B1 and ZINC27215482, and with 2′-o-ribose methyltransferase for conivaptan, loniflavone and ZINC15675938. The protein-drug interactions are illustrated in Figs. 3–5.

The stability of the loniflavone docking pose on the spike receptor binding domain was assessed with MD simulation. As can be seen in Fig. 6 and **Supplementary Video**, loniflavone was stably interacting with the protein.

Supplementary video related to this article can be found at https://doi.org/10.1016/j.compbiomed.2021.104359

4. Discussion

COVID-19 rapidly increased to an epidemic in China. Although still mostly restricted to the Hubei province, there is a reasonable threat that the disease may spread all over the world. With 219 countries and territories affected (status: March 23, 2021), it will be difficult to manage the outbreak without drugs and vaccines available. Therefore, there is an urgent requirement for drugs that inhibit SARS-CoV-2. We have selected three important viral proteins as targets for our combined virtual screening/machine learning approach, *i.e.* spike protein, nucleocapsid protein, and 2′-o-ribose-methyltransferase. The spike protein is involved in binding of the virus to cellular receptors of the host. As this protein governs the entry of the virus into the host cell [42], it represents a premier target for the development of drugs and vaccines against coronaviruses [43,44]. The nucleocapsid protein forms complexes with genomic RNA of the virus and plays a crucial role in coronavirus transcription and assembly [45]. It has recently been discussed as valuable target for the development of drugs against coronaviruses [46]. 2′-*O*-ribose methyltransferase is involved in the capping of coronaviral mRNAs and is essential for efficient coronavirus RNA synthesis and processing [47]. We also performed virtual screening with another conserved structural coronaviral protein, *i.e.* the envelope protein, but found only weak binding energies (higher than −7 kcal/mol) of the FDA-approved drugs and natural compounds to the selected three target proteins (data not shown). Therefore, we did not further consider the envelope protein as relevant target for anti-SARS-CoV-2 drugs.

These coronaviral proteins were used as targets for virtual screening (blind and defined docking mode), molecular docking (defined docking mode), and supervised machine learning algorithms (naïve bayes, neural network) using FDA-approved drugs and natural compounds. The drug repurposing approach in the present investigation also brought up interesting results. Several FDA-approved drugs against hepatitis C, bacterial and fungal infections, cancer and other diseases appeared in the top ranks of our virtual screenings. Especially, the anti-hepatitis C drugs (paritaprevir, simeprevir, grazoprevir, and velpatasvir) deserve attention, since the hepatitis C virus is also an enveloped ssRNA virus. There is already an editorial article regarding the possible effect of velpatasvir combined with sofosbuvir (another anti-hepatitis C drug) against SARS-CoV-2 [48]. Hence, it is reasonable to speculate that these drugs may also exert activity against SARS-CoV-2. Interestingly, all of the identified anti-hepatitis C drugs bound to the spike protein in our *in silico* approach.

The validity of our results is supported by a recent study the anti-HCV drug IDX-184 was also active against Middle East Respiratory Syndrome (MERS) coronavirus [49]. Hence, anti-HCV drugs might reveal a general potency against human coronaviruses. The finding that anti-HCV drugs may be active against SARS-CoV-2 is novel and may enlarge the armory of investigational drugs to fight COVID-19. Other anti-retroviral drugs are also under investigation against SARS-CoV-2. These drugs act against enveloped (−) ssRNA viruses (remdesivir against Ebola virus and Marburg virus, oseltamivir against influenza A and B viruses) or enveloped linear, dimeric ssRNA viruses (lopinavir and ritonavir against HIV1 and HIV-2). This is in line with the fact that HCV is also an enveloped (−) ssRNA virus. Hence, it is reasonable to assume that anti-HCV drugs are also valuable to combat SARS-CoV-2.

Many drugs among the FDA-approved drugs and also among the natural product datasets bind with high affinity not only to one target

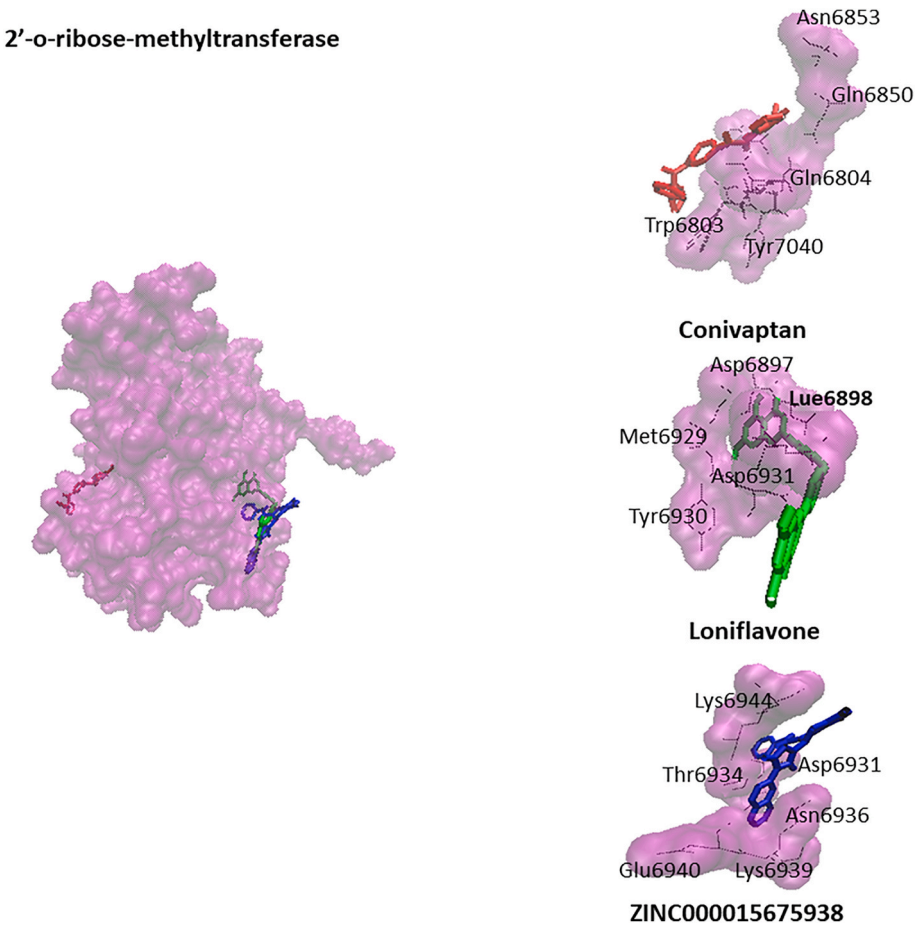

Fig. 5. Docking poses of conivaptan (red), loniflavone (green) and ZINC15675938 (blue) on 2′-o-ribose-methyltransferase (purple). Residues forming hydrogen bonds are labeled bold.

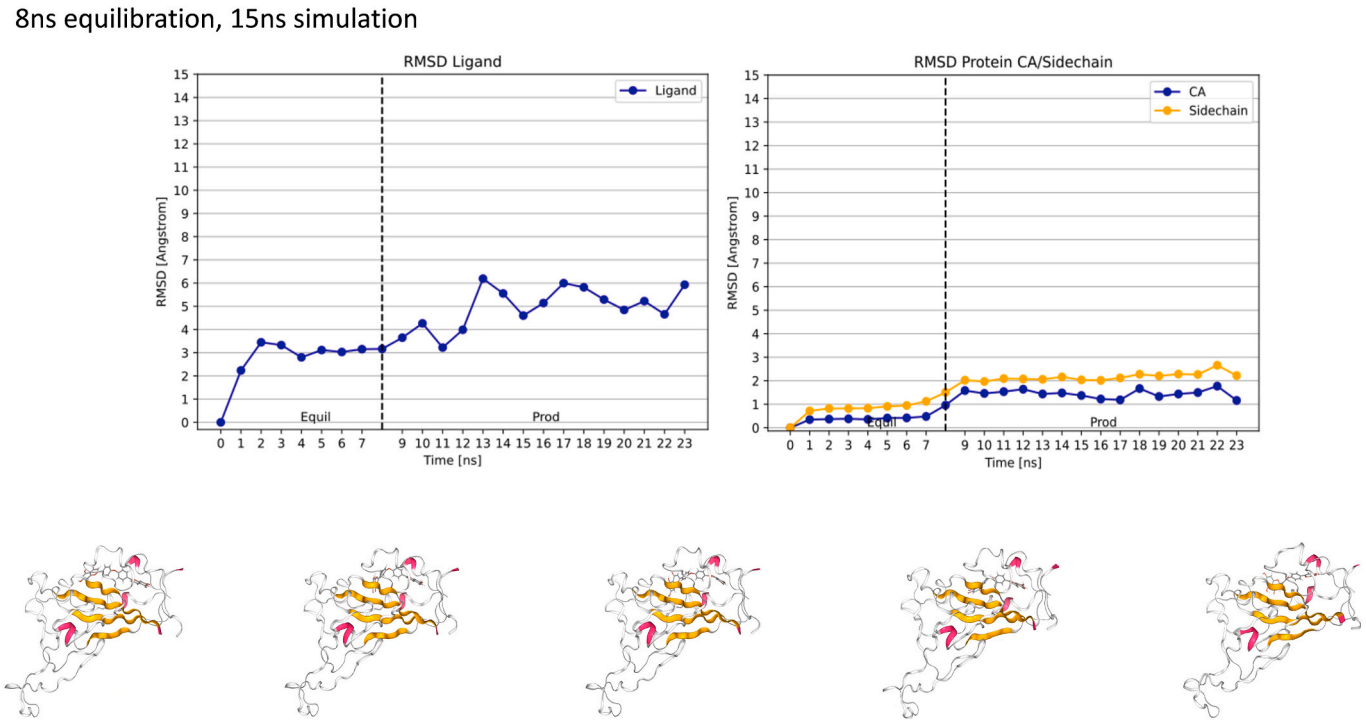

Fig. 6. MD simulation of the loniflavone docking pose on the spike receptor binding domain.

protein but also to another one (conivaptan, paritaprevir, simeprevir, dihydroergotamine, ZINC000027215482, ZINC000252515584, loniflavone, procyanidin). There are already published studies reporting the potential of paritaprevir [50], simeprevir [51], dihydroergotamine [52], procyanidin [53] against SARS-CoV-2. These compounds deserve special attention. Binding of small molecule inhibitors to two targets at the same time may increase the therapeutic efficacy and decrease the probability of development of resistance to one of the targets. Especially, natural products are known to bind to multiple targets [54]. This has been frequently misinterpreted as non-specificity. During evolution of life on earth, chemical weapons of organisms against microbial attack from viruses, bacteria, protozoans or other threats from predators were more successful, if they were multi-specific. Inhibiting several targets at the same time better prevents the development of resistance against single-target drugs. From an evolutionary point of view, this strategy provided better chances for the survival of the fittest. It deserves further exploration, whether the bispecifically binding compounds exert superior activity against SARS-CoV-2.

Furthermore, our results from the drug repurposing approach by using 1577 FDA-approved drugs generally fit together with other well-known drugs from the literature, *e.g.* the anti-malarial artemisinin and its derivatives are also active against viruses, other infectious diseases and cancer [55–58]. Broad-spectrum activities have also been reported for other classes of pharmacological drugs [59], indicating that drug repurposing represents a fertile reservoir to develop drugs to fight COVID-19.

During the past few years, molecular docking has been used for the identification of synthetic and natural drug candidates against targets of MERS-CoV and SARS-CoV such as chymotrypsin-like protease [60–63], mRNA polymerases [64], and helicase [65]. To the best of our knowledge, we are the first describing drug candidates against viral proteins of SARS-CoV-2 by a combined virtual screening/molecular docking/supervised machine learning *in silico* approach. The compounds identified by us deserve further investigation *in vitro* and *in vivo*. The candidate compounds reported in this study may provide a basis to develop novel treatment options against COVID-19.

Declaration of competing interest

The authors declare no conflict of interest.

Acknowledgments

Parts of this research were conducted using the supercomputer MOGON II and/or advisory services offered by Johannes Gutenberg University Mainz (hpc.uni-mainz.de), which is a member of the AHRP (Alliance for High-Performance Computing in Rhineland Palatinate, www.ahrp.info) and the Gauss Alliance e.V. The authors gratefully acknowledge the computing time granted on the supercomputer Mogon at Johannes Gutenberg University Mainz (hpc.uni-mainz.de).

References

[1] World Health Organization Coronavirus disease (COVID-2019), Situation Reports, 2020 accessed on February 23, 2020, https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports. accessed on March 23, 2021.
[2] F. Wu, S. Z, B. Yu, Y.M. Chen, W. Wang, Z.G. Song, Y. Hu, Z.W. Tao, J.H. Tian, Y. Y. Pei, M.L. Yuan, Y.L. Zhang, F.H. Dai, Y. Liu, Q.M. Wang, J.J. Zheng, L. Xu, E. C. Holmes, Y.Z. Zhang, A new coronavirus associated with human respiratory disease in China, Nature 579 (2020) 265–269.
[3] P. Zhou, X.-L. Y, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, H.-R. Si, Y. Zhu, B. Li, C.-L. Huang, H.-D. Chen, J. Chen, Y. Luo, H. Guo, R.-D. Jiang, M.-Q. Liu, Y. Chen, X.-R. Shen, X. Wang, X.-S. Zheng, K. Zhao, Q.-J. Chen, F. Deng, L.-L. Liu, B. Yan, F.-X. Zhan, Y.-Y. Wang, G.-F. Xiao, Z.-L. Shi, Discovery of a Novel Coronavirus Associated with the Recent Pneumonia Outbreak in Humans and its Potential Bat Origin, bioRxiv, 2020.
[4] Y. Chen, et al., Structure analysis of the receptor binding of 2019-nCoV, Biochem. Biophys. Res. Commun. 525 (1) (2020) 135–140.
[5] P. Liu, W. Chen, J.P. Chen, Viral metagenomics revealed Sendai virus and coronavirus infection of Malayan pangolins (Manis javanica), Viruses 11 (11) (2019) 979.
[6] D. Paraskevis, et al., Full-genome evolutionary analysis of the novel corona virus (2019-nCoV) rejects the hypothesis of emergence as a result of a recent recombination event, Infect. Genet. Evol. 79 (2020) 104212.
[7] V.M. Corman, et al., Hosts and sources of endemic human coronaviruses, Adv. Virus Res. 100 (2018) 163–188.
[8] D. Wang, et al., Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China, J. Am. Med. Assoc. 323 (11) (2020) 1061–1069.
[9] Z. Wang, et al., Clinical characteristics and therapeutic procedure for four cases with 2019 novel coronavirus pneumonia receiving combined Chinese and Western medicine treatment, Bioscience Trends 14 (1) (2020) 64–68.
[10] E. de Wit, et al., Prophylactic and therapeutic remdesivir (GS-5734) treatment in the rhesus macaque model of MERS-CoV infection, Proc. Natl. Acad. Sci. U. S. A. 117 (12) (2020) 6771–6776.
[11] Y.M. Arabi, R. Fowler, F.G. Hayden, Critical care management of adults with community-acquired severe respiratory viral infection, Intensive Care Med. 46 (2) (2020 Feb) 315–328, https://doi.org/10.1007/s00134-020-05943-5.
[12] J. Lim, et al., Case of the index patient who caused tertiary transmission of COVID-19 infection in Korea: the application of lopinavir/ritonavir for the treatment of COVID-19 infected pneumonia monitored by quantitative RT-PCR, J. Korean Med. Sci. 35 (6) (2020) e79.
[13] J.T. Denholm, et al., The Australasian COVID-19 Trial (ASCOT) to assess clinical outcomes in hospitalised patients with SARS-CoV-2 infection (COVID-19) treated with lopinavir/ritonavir and/or hydroxychloroquine compared to standard of care: a structured summary of a study protocol for a randomised controlled trial, Trials 21 (1) (2020) 646.
[14] C.D. Spinner, et al., Effect of remdesivir vs standard care on clinical status at 11 days in patients with moderate COVID-19: a randomized clinical trial, J. Am. Med. Assoc. 324 (11) (2020) 1048–1057.
[15] R.S. Wax, M.D. Christian, Practical recommendations for critical care and anesthesiology teams caring for novel coronavirus (2019-nCoV) patients, Can. J. Anaesthesiol. 67 (5) (2020 May) 568–576, https://doi.org/10.1007/s12630-020-01591-x.
[16] T.T. Ashburn, K.B. Thor, Drug repositioning: identifying and developing new uses for existing drugs, Nat. Rev. Drug Discov. 3 (8) (2004) 673–683.
[17] D.J. Newman, G.M. Cragg, Natural products as sources of new drugs from 1981 to 2014, J. Nat. Prod. 79 (3) (2016) 629–661.
[18] D.J. Newman, G.M. Cragg, Natural products as sources of new drugs over the 30 years from 1981 to 2010, J. Nat. Prod. 75 (3) (2012) 311–335.
[19] T. Wang, et al., Quantitative structure-activity relationship: promising advances in drug discovery platforms, Expet Opin. Drug Discov. 10 (12) (2015) 1283–1300.
[20] X. Yang, et al., Concepts of artificial intelligence for computer-assisted drug discovery, Chem. Rev. 119 (18) (2019) 10520–10594.
[21] O. Kadioglu, T. Efferth, A machine learning-based prediction platform for P-glycoprotein modulators and its validation by molecular docking, Cells 8 (10) (2019) 1286.
[22] M.C. Robinson, R.C. Glen, A.A. Lee, Validating the validation: reanalyzing a large-scale comparison of deep learning and machine learning models for bioactivity prediction, J. Comput. Aided Mol. Des. 34 (7) (2020) 717–730.
[23] Y. Chang, et al., Cancer Drug Response Profile scan (CDRscan): a deep learning model that predicts drug effectiveness from cancer genomic signature, Sci. Rep. 8 (1) (2018) 8857.
[24] 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) (2011), e1002294.
[25] M. Galloux, et al., Characterization of a viral phosphoprotein binding site on the surface of the respiratory syncytial nucleoprotein, J. Virol. 86 (16) (2012) 8375–8387.
[26] Y. Yuan, et al., Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains, Nat. Commun. 8 (2017) 15092.
[27] J. Hupfeld, T. Efferth, Drug resistance of human immunodeficiency virus and overcoming it by natural products, In Vivo 23 (1) (2009) 1–6.
[28] K. Andrae-Marobela, et al., Polyphenols: a diverse class of multi-target anti-HIV-1 agents, Curr. Drug Metabol. 14 (4) (2013) 392–413.
[29] G. Prinsloo, C.K. Marokane, R.A. Street, Anti-HIV activity of southern African plants: current developments, phytochemistry and future research, J. Ethnopharmacol. 210 (2018) 133–155.
[30] B. Salehi, et al., Medicinal plants used in the treatment of human immunodeficiency virus, Int. J. Mol. Sci. 19 (5) (2018) 1459.
[31] T.M. Ehrman, D.J. Barlow, P.J. Hylands, Virtual screening of Chinese herbs with random forest, J. Chem. Inform. Model. 47 (2) (2007) 264–278.
[32] O. Kadioglu, T. Efferth, Peptide aptamer identified by molecular docking targeting translationally controlled tumor protein in leukemia cells, Invest. N. Drugs 34 (4) (2016) 515–521.
[33] O. Kadioglu, et al., Interactions of human P-glycoprotein transport substrates and inhibitors at the drug binding domain: functional and molecular docking analyses, Biochem. Pharmacol. 104 (2016) 42–51.
[34] China trials anti-HIV drug on coronavirus patients. https://www.theguardian.com/world/2020/feb/07/china-trials-anti-hiv-drug-coronavirus-patients. (accessed February 13, 2020).

[35] Y. Saplakoglu, How Experts Plan to Treat the New Coronavirus, Live Science, 2020. https://www.livescience.com/possible-treatments-new-coronavirus.html. accessed on March 23, 2021.
[36] J. Cohen, Can an anti-HIV combination or other existing drugs outwit the new coronavirus? Science (2020). https://www.sciencemag.org/news/2020/01/can-anti-hiv-combination-or-other-existing-drugs-outwit-new-coronavirus. accessed on March 23, 2021.
[37] C. Offord, Flu and HIV Drugs Show Efficacy against Coronavirus, The Scientist, 2020. https://www.the-scientist.com/news-opinion/flu-and-anti-hiv-drugs-show-efficacy-against-coronavirus-67052. accessed on March 23, 2021.
[38] C. Devereux, et al., Extending the applicability of the ANI deep learning molecular potential to sulfur and halogens, J. Chem. Theor. Comput. 16 (7) (2020) 4192–4202.
[39] J.S. Smith, et al., Less is more: sampling chemical space with active learning, J. Chem. Phys. 148 (24) (2018) 241733.
[40] G. Martinez-Rosell, T. Giorgino, G. De Fabritiis, PlayMolecule ProteinPrepare: a web application for protein preparation for molecular dynamics simulations, J. Chem. Inform. Model. 57 (7) (2017) 1511–1516.
[41] S. Doerr, et al., HTMD: high-throughput molecular dynamics for molecular discovery, J. Chem. Theor. Comput. 12 (4) (2016) 1845–1852.
[42] P.S. Masters, Coronavirus genomic RNA packaging, Virology 537 (2019) 198–207.
[43] L. Du, et al., MERS-CoV spike protein: a key target for antivirals, Expert Opin. Ther. Targets 21 (2) (2017) 131–143.
[44] Y. Zhou, et al., Advances in MERS-CoV vaccines and therapeutics based on the receptor-binding domain, Viruses 11 (1) (2019) 60.
[45] R. McBride, M. van Zyl, B.C. Fielding, The coronavirus nucleocapsid is a multifunctional protein, Viruses 6 (8) (2014) 2991–3018.
[46] C.K. Chang, et al., Recent insights into the development of therapeutics against coronavirus diseases by targeting N protein, Drug Discov. Today 21 (4) (2016) 562–572.
[47] F. Almazan, et al., Construction of a severe acute respiratory syndrome coronavirus infectious cDNA clone and a replicon to study coronavirus RNA synthesis, J. Virol. 80 (21) (2006) 10900–10906.
[48] A. Izzi, et al., Editorial - sofosbuvir/Velpatasvir as a combination with strong potential activity against SARS-CoV2 (COVID-19) infection: how to use direct-acting antivirals as broad-spectrum antiviral agents, Eur. Rev. Med. Pharmacol. Sci. 24 (10) (2020) 5193–5194.
[49] A.A. Elfiky, S.M. Mahdy, W.M. Elshemey, Quantitative structure-activity relationship and molecular docking revealed a potency of anti-hepatitis C virus drugs against human corona viruses, J. Med. Virol. 89 (6) (2017) 1040–1047.
[50] R.J. Khan, et al., Targeting SARS-CoV-2: a systematic drug repurposing approach to identify promising inhibitors against 3C-like proteinase and 2'-O-ribose methyltransferase, J. Biomol. Struct. Dyn. (2020) 1–14.
[51] A. J, et al., Repurposing simeprevir, calpain inhibitor IV and a cathepsin F inhibitor against SARS-CoV-2 and insights into their interactions with M(pro), J. Biomol. Struct. Dyn. (2020) 1–12.
[52] N.A. Murugan, et al., Searching for target-specific and multi-targeting organics for Covid-19 in the Drugbank database with a double scoring approach, Sci. Rep. 10 (1) (2020) 19125.
[53] N. Maroli, et al., The potential role of procyanidin as a therapeutic agent against SARS-CoV-2: a text mining, molecular docking and molecular dynamics simulation approach, J. Biomol. Struct. Dyn. (2020) 1–16.
[54] T. Efferth, E. Koch, Complex interactions between phytochemicals. The multi-target therapeutic concept of phytotherapy, Curr. Drug Targets 12 (1) (2011) 122–132.
[55] M.E.M. Saeed, et al., Antischistosomal activity of artemisinin derivatives in vivo and in patients, Pharmacol. Res. 110 (2016) 216–226.
[56] T. Efferth, From ancient herb to modern drug: artemisia annua and artemisinin for cancer therapy, Semin. Canc. Biol. 46 (2017) 65–83.
[57] T. Efferth, Beyond malaria: the inhibition of viruses by artemisinin-type compounds, Biotechnol. Adv. 36 (6) (2018) 1730–1737.
[58] J. Nass, T. Efferth, The activity of Artemisia spp. and their constituents against Trypanosomiasis, Phytomedicine 47 (2018) 184–191.
[59] L. Senerovic, et al., Quinolines and quinolones as antibacterial, antifungal,anti-virulence, antiviral and anti-parasitic agents, Adv. Exp. Med. Biol. 1282 (2020) 37–69.
[60] Y.B. Ryu, et al., Bioflavonoids from Torreya nucifera displaying SARS-CoV 3CL (pro) inhibition, Bioorg. Med. Chem. 18 (22) (2010) 7940–7947.
[61] T.T. Nguyen, et al., Flavonoid-mediated inhibition of SARS coronavirus 3C-like protease expressed in Pichia pastoris, Biotechnol. Lett. 34 (5) (2012) 831–838.
[62] H. Lee, et al., Identification of novel drug scaffolds for inhibition of SARS-CoV 3-Chymotrypsin-like protease using virtual and high-throughput screenings, Bioorg. Med. Chem. 22 (1) (2014) 167–177.
[63] M. Berry, B.C. Fielding, J. Gamieldien, Potential broad spectrum inhibitors of the coronavirus 3CLpro: a virtual screening and structure-based drug design study, Viruses 7 (12) (2015) 6642–6660.
[64] A.A. Elfiky, S.M. Mahdy, W.M. Elshemey, Quantitative structure-activity relationship and molecular docking revealed a potency of anti-hepatitis C virus drugs against human corona viruses, J. Med. Virol. 89 (6) (2017) 1040–1047.
[65] N.H. Zaher, M.I. Mostafa, A.Y. Altaher, Design, synthesis and molecular docking of novel triazole derivatives as potential CoV helicase inhibitors, Acta Pharm. 70 (2) (2020) 145–159.