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

Onat Kadioglu1*, Mohamed Saeed1*, Henry Johannes Greten2,3, Thomas Efferth1§
*both authors contributed equally

1. Department of Pharmaceutical Biology, Institute of Pharmacy and Biochemistry, Johannes Gutenberg University, Mainz. Germany
2. Abel Salazar Biomedical Sciences Institute, University of Porto, Portugal
3. Heidelberg School of Chinese Medicine, Heidelberg, Germany

§ Corresponding address: Staudinger Weg 5, 55128 Mainz, Germany: Tel.: 49-6131-3925751, Fax: 49-6131-3923752, Email: efferth@uni-mainz.de

(Submitted: 20 March 2020 – Published online: 21 March 2020)

Running title: Compounds against SARS-CoV-2

Key words: Artificial intelligence; Chemotherapy; COVID-19; Infectious diseases; Natural products

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.

RECOMMENDED CITATION

Kadioglu O, Saeed M, Johannes Greten H & Efferth T. Identification of novel compounds against three targets of SARS CoV-2 coronavirus by combined virtual screening and supervised machine learning. [Preprint]. Bull World Health Organ. E-pub: 21 March 2020. doi: http://dx.doi.org/10.2471/BLT.20.255943

Abstract

Coronavirus disease 2019 (COVID-19) is a major threat worldwide due to its fast spreading. As yet, there are no established drugs or vaccines 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 II, 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, grazoprevir, 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.

Introduction

In the Chinese city of Wuhan, Hubei province, several cases of novel, SARS-like, severe pneumonia occurred for the first time in December 2019, as firstly reported by the physician Li Wenliang and officially 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 transmitted from these animals to human hosts.

Some coronaviruses (*e.g.* HCoV-229E, -NL63, -OC43, and -HKU1) usually cause respiratory infections and circulate worldwide in human populations (7). Other coronaviruses (*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, muscle pain and other symptoms (8). As of March 20, 2020, 260,000 people were infected and more than 10,000 deaths occurred. On the other hand, 90,000 patients recovered from COVID-19 (https://www.worldometers.info/coronavirus/).

As of yet, there are no drugs or vaccines available 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). Randomized, placebo-controlled, double-blind studies are still missing. The current clinical treatment is largely based on symptom-based therapies (11, 13). 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 (14). As it became apparent that drugs approved for one disease, may also exert activity for other indications, FDA-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 (15, 16). 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 development of new drugs (17, 18), overcoming multidrug resistance (19), and applications in precision medicine to select drugs for individualized therapies (20, 21).

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.

Materials and methods

Workflow

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

(1) 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 spike protein, nucleocapsid protein, and 2'-o-ribose methyltransferase were taken as templates to generate 3D homology models of the three SARS-CoV-2 proteins.

(2) Construction of compound databases: (A) 1,577 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-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.

Homology modeling

Three proteins from SARS-CoV-2 were selected as templates for homology modelling, *i.e.* spike protein (YP_009724390.1), nucleocapsid protein (YP_009724397.2), and 2'-o-ribose methyltransferase (YP_009725311.1). The genome sequence of SARS-CoV-2 (GenBank ID: NC_045512.2) was used to deduce the protein sequences and to generate homology models derived from YP_009725311.1, YP_009724397.2 and YP_009724390.1. The Phyre2 software (22) (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index) was used to build homology protein structures with the highest available homology and coverage percentages. The reported pharmacophores were the receptor binding domain of spike glycoprotein, the RNA binding domain of nucleocapsid protein, and the catalytic site of 2'-o-ribose methyltransferase) (23-25).

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). The protein sequences were deduced from the NCBI website. FDA-approved drugs (1,577 compounds), natural compounds from the ZINC database (39,442 compounds), and natural compounds mined from the literature with antiviral activity (115 compounds) (26-30). 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) the from DrugBank database (https://www.drugbank.ca/). The positive control drugs revealed binding energies of ≤-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 PyRx algorithm (blind docking mode) and the high-performance supercomputer MOGON II (Johannes Gutenberg University, Mainz).

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. Lamarckian algorithm was used to analyze the docking poses and binding energies with as described before (19, 31). The ligand moved around the rigid protein was rigid 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/).

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 (32-35), we used the above mentioned presumable positive control and negative control drugs. After random selection was applied, 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) (19). 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. Leave one out random sampling was used to build the models. To select the most suited algorithm, we applied 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 **Figure 2**.

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**). After 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). We first 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 ACE3-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) ma 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. **Figure 2** displays 3 out of 11 tested algorithms for illustration. With these prediction models, the test compounds were calculated, and excellent ROC 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, grazoprevir, and velpatasvir), indicating that these drugs may also be suitable 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, paritaprevir and teniposide bound to spike protein and 2'-o-ribose methyltransferase and dihydroergotamine and venetoclax to nucleocapsid protein and 2'-o-ribose methyltransferase. Among the natural products, amyrin, ZINC000027215482 and ZINC000252515584 bound to spike protein and 2'-o-ribose methyltransferase, while procyanidin bound to spike 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 simepravir, loniflavone and ZINC000027215482, (2) with the nucleocapsid protein for conivaptan, amyrin, and ZINC000027215482, and with 2'-o-ribose methyltransferase for ZINC000008635475. The protein-drug interactions are illustrated in **Figures 3-5**.

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 a transmission rate of 2-3.5 and 183 countries and territories affected (status March 20, 2020), 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 (36), it represents a premier target for the development of drugs and vaccines against coronaviruses (37, 38). The nucleocapsid protein forms complexes with genomic RNA of the virus and plays a crucial role in coronavirus transcription and assembly (39). It has recently been discussed as valuable target for the development of drugs against coronaviruses (40). 2'-O-ribose

methyltransferase is involved in the capping of coronaviral mRNAs and is essential for efficient coronavirus RNA synthesis and processing (41). 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 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, semeprevir, grazoprevir, and velpatasvir) deserve attention, since the hepatitis C virus is also an enveloped ssRNA virus. 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 the anti-HCV drug IDX-184 was also active against Middle East Respiratory Syndrome (MERS) coronavirus (42). 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 protein but also to another one (paritaprevir, teniposide, dihydroergotamine, venetoclax, amyrin, ZINC000027215482, ZINC000252515584, and procyanidin). 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 (43). 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 1,577 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 (44-47). Another example is the antimalarial chloroquine, which also inhibits cancer (48, 49). Recently, chloroquine has also been used to treat COVID-19 (50). Broad-spectrum activities have also been reported for other classes of pharmacological drugs (51), 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 (52-55), mRNA polymerases (56), and helicase (57). 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*.

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).

Conflicts of interest: The authors declare no conflict of interest.

References

1. World Health Organization Coronavirus disease (COVID-2019) situation reports 2020, accessed on February 23, 2020. [cited 2020 February 23]; Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports.
2. Wu F ZS, Yu B, Chen YM, Wang W, Song ZG, Hu Y, Tao ZW, Tian JH, Pei YY, Yuan ML, Zhang YL, Dai FH, Liu Y, Wang QM, Zheng JJ, Xu L, Holmes EC, Zhang YZ. A new coronavirus associated with human respiratory disease in China. Nature. 2020. Epub Epub ahead of print.
3. Zhou P YX-L, Wang X-G, Hu B, Zhang L, Zhang W, Si H-R, Zhu Y, Li B, Huang C-L, Chen H-D, Chen J, Luo Y, Guo H, Jiang R-D, Liu M-Q, Chen Y, Shen X-R, Wang X, Zheng X-S, Zhao K, Chen Q-J, Deng F, Liu L-L, Yan B, Zhan F-X, Wang Y-Y, Xiao G-F, Shi Z-L. Discovery of a novel coronavirus associated with the recent pneumonia outbreak in humans and its potential bat origin. bioRxiv. 2020. Epub Januar 23.
4. Chen Y, Guo Y, Pan Y, Zhao ZJ. Structure analysis of the receptor binding of 2019-nCoV. Biochemical and biophysical research communications. 2020.
5. Liu P, Chen W, Chen JP. Viral Metagenomics Revealed Sendai Virus and Coronavirus Infection of Malayan Pangolins (Manis javanica). Viruses. 2019;11(11).
6. Paraskevis D, Kostaki EG, Magiorkinis G, Panayiotakopoulos G, Sourvinos G, Tsiodras S. 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. 2020;79.
7. Corman VM, Muth D, Niemeyer D, Drosten C. Hosts and Sources of Endemic Human Coronaviruses. Adv Virus Res. 2018;100:163-88.
8. Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, et al. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. JAMA. 2020.
9. Wang Z, Chen X, Lu Y, Chen F, Zhang W. Clinical characteristics and therapeutic procedure for four cases with 2019 novel coronavirus pneumonia receiving combined Chinese and Western medicine treatment. Biosci Trends. 2020.
10. de Wit E, Feldmann F, Cronin J, Jordan R, Okumura A, Thomas T, 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. 2020.
11. Arabi YM, Fowler R, Hayden FG. Critical care management of adults with community-acquired severe respiratory viral infection. Intensive Care Med. 2020.
12. Lim J, Jeon S, Shin HY, Kim MJ, Seong YM, Lee WJ, 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. 2020;35(6):e79.
13. Wax RS, Christian MD. Practical recommendations for critical care and anesthesiology teams caring for novel coronavirus (2019-nCoV) patients. Can J Anaesth. 2020. Directives concretes a l'intention des equipes de soins intensifs et d'anesthesiologie prenant soin de patients atteints du coronavirus 2019-nCoV.
14. Ashburn TT, Thor KB. Drug repositioning: identifying and developing new uses for existing drugs. Nature reviews Drug discovery. 2004;3(8):673-83.
15. Newman DJ, Cragg GM. Natural Products as Sources of New Drugs from 1981 to 2014. J Nat Prod. 2016;79(3):629-61.
16. Newman DJ, Cragg GM. Natural Products As Sources of New Drugs over the 30 Years from 1981 to 2010. J Nat Prod. 2012;75(3):311-35.
17. Wang T, Wu MB, Lin JP, Yang LR. Quantitative structure-activity relationship: promising advances in drug discovery platforms. Expert Opin Drug Discov. 2015;10(12):1283-300.

18. Yang X, Wang Y, Byrne R, Schneider G, Yang S. Concepts of Artificial Intelligence for Computer-Assisted Drug Discovery. Chem Rev. 2019;119(18):10520-94.
19. Kadioglu O, Efferth T. A Machine Learning-Based Prediction Platform for P-Glycoprotein Modulators and Its Validation by Molecular Docking. Cells. 2019;8(10).
20. Robinson MC, Glen RC, Lee AA. Validating the validation: reanalyzing a large-scale comparison of deep learning and machine learning models for bioactivity prediction. Journal of computer-aided molecular design. 2020.
21. Chang Y, Park H, Yang HJ, Lee S, Lee KY, Kim TS, et al. Cancer Drug Response Profile scan (CDRscan): A Deep Learning Model That Predicts Drug Effectiveness from Cancer Genomic Signature. Scientific reports. 2018;8.
22. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. Nature Protocols. 2015;10(6):845-58.
23. Chen Y, Su CY, Ke M, Jin X, Xu LR, Zhang Z, et al. Biochemical and Structural Insights into the Mechanisms of SARS Coronavirus RNA Ribose 2 '-O-Methylation by nsp16/nsp10 Protein Complex. Plos Pathog. 2011;7(10).
24. Galloux M, Tarus B, Blazevic I, Fix J, Duquerroy S, Eleouet JF. Characterization of a Viral Phosphoprotein Binding Site on the Surface of the Respiratory Syncytial Nucleoprotein. Journal of virology. 2012;86(16):8375-87.
25. Yuan Y, Cao DF, Zhang YF, Ma J, Qi JX, Wang QH, et al. Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. Nature Communications. 2017;8.
26. Hupfeld J, Efferth T. Drug Resistance of Human Immunodeficiency Virus and Overcoming it by Natural Products. In Vivo. 2009;23(1):1-6.
27. Andrae-Marobela K, Ghislain FW, Okatch H, Majinda RRT. Polyphenols: A Diverse Class of Multi-Target Anti-HIV-1 Agents. Curr Drug Metab. 2013;14(4):392-413.
28. Prinsloo G, Marokane CK, Street RA. Anti-HIV activity of southern African plants: Current developments, phytochemistry and future research. Journal of ethnopharmacology. 2018;210:133-55.
29. Salehi B, Kumar NVA, Sener B, Sharifi-Rad M, Kilic M, Mahady GB, et al. Medicinal Plants Used in the Treatment of Human Immunodeficiency Virus. International journal of molecular sciences. 2018;19(5).
30. Ehrman TM, Barlow DJ, Hylands PJ. Virtual screening of Chinese herbs with Random Forest. Journal of chemical information and modeling. 2007;47(2):264-78.
31. Kadioglu O, Saeed MEM, Valoti M, Frosini M, Sgaragli G, Efferth T. Interactions of human P-glycoprotein transport substrates and inhibitors at the drug binding domain: Functional and molecular docking analyses. Biochem Pharmacol. 2016;104:42-51.
32. China trials anti-HIV drug on coronavirus patients, accessed on 13 February 2020. The Guardian; 2020; Available from: https://www.theguardian.com/world/2020/feb/07/china-trials-anti-hiv-drug-coronavirus-patients.
33. How experts plan to treat the new coronavirus, accessed on 13 February 2020

2020; Available from: https://www.livescience.com/possible-treatments-new-coronavirus.html.
34. Can an anti-HIV combination or other existing drugs outwit the new coronavirus?, accessed on 13 February 2020. 2020; Available from: https://www.sciencemag.org/news/2020/01/can-anti-hiv-combination-or-other-existing-drugs-outwit-new-coronavirus.
35. Flu and HIV Drugs Show Efficacy Against Coronavirus, accessed on 13 February 2020. 2020; Available from: https://www.the-scientist.com/news-opinion/flu-and-anti-hiv-drugs-show-efficacy-against-coronavirus-67052.
36. Masters PS. Coronavirus genomic RNA packaging. Virology. 2019;537:198-207.

37. Du L, Yang Y, Zhou Y, Lu L, Li F, Jiang S. MERS-CoV spike protein: a key target for antivirals. Expert opinion on therapeutic targets. 2017;21(2):131-43.
38. Zhou Y, Yang Y, Huang J, Jiang S, Du L. Advances in MERS-CoV Vaccines and Therapeutics Based on the Receptor-Binding Domain. Viruses. 2019;11(1).
39. McBride R, van Zyl M, Fielding BC. The Coronavirus Nucleocapsid Is a Multifunctional Protein. Viruses-Basel. 2014;6(8):2991-3018.
40. Chang CK, Lo SC, Wang YS, Hou MH. Recent insights into the development of therapeutics against coronavirus diseases by targeting N protein. Drug Discov Today. 2016;21(4):562-72.
41. Almazan F, DeDiego ML, Galan C, Escors D, Alvarez E, Ortego J, et al. Construction of a severe acute respiratory syndrome coronavirus infectious cDNA clone and a replicon to study coronavirus RNA synthesis. Journal of virology. 2006;80(21):10900-6.
42. Elfiky AA, Mahdy SM, Elshemey WM. Quantitative structure-activity relationship and molecular docking revealed a potency of anti-hepatitis C virus drugs against human corona viruses. J Med Virol. 2017;89(6):1040-7.
43. Efferth T, Koch E. Complex interactions between phytochemicals. The multi-target therapeutic concept of phytotherapy. Current drug targets. 2011;12(1):122-32. Epub 2010/08/26.
44. Saeed MEM, Krishna S, Greten HJ, Kremsner PG, Efferth T. Antischistosomal activity of artemisinin derivatives in vivo and in patients. Pharmacological Research. 2016;110:216-26.
45. Efferth T. From ancient herb to modern drug: Artemisia annua and artemisinin for cancer therapy. Semin Cancer Biol. 2017;46:65-83.
46. Efferth T. Beyond malaria: The inhibition of viruses by artemisinin-type compounds. Biotechnol Adv. 2018;36(6):1730-7.
47. Nass J, Efferth T. The activity of Artemisia spp. and their constituents against Trypanosomiasis. Phytomedicine : international journal of phytotherapy and phytopharmacology. 2018;47:184-91. Epub 2018/09/01.
48. Xu R, Ji ZY, Xu C, Zhu J. The clinical value of using chloroquine or hydroxychloroquine as autophagy inhibitors in the treatment of cancers A systematic review and meta-analysis. Medicine. 2018;97(46).
49. Varisli L, Cen O, Vlahopoulos S. Dissecting pharmacological effects of chloroquine in cancer treatment: interference with inflammatory signaling pathways. Immunology. 2020;159(3):257-78.
50. Gao J, Tian Z, Yang X. Breakthrough: Chloroquine phosphate has shown apparent efficacy in treatment of COVID-19 associated pneumonia in clinical studies. Bioscience trends. 2020.
51. Senerovic L, Opsenica D, Moric I, Aleksic I, Spasić M, Vasiljevic B. Quinolines and Quinolones as Antibacterial, Antifungal, Anti-virulence, Antiviral and Anti-parasitic Agents. 2019.
52. Ryu YB, Jeong HJ, Kim JH, Kim YM, Park JY, Kim D, et al. Biflavonoids from Torreya nucifera displaying SARS-CoV 3CL(pro) inhibition. Bioorganic & medicinal chemistry. 2010;18(22):7940-7.
53. Nguyen TT, Woo HJ, Kang HK, Nguyen VD, Kim YM, Kim DW, et al. Flavonoid-mediated inhibition of SARS coronavirus 3C-like protease expressed in Pichia pastoris. Biotechnology letters. 2012;34(5):831-8.
54. Lee H, Mittal A, Patel K, Gatuz JL, Truong L, Torres J, et al. Identification of novel drug scaffolds for inhibition of SARS-CoV 3-Chymotrypsin-like protease using virtual and high-throughput screenings. Bioorganic & medicinal chemistry. 2014;22(1):167-77.
55. Berry M, Fielding BC, Gamieldien J. Potential Broad Spectrum Inhibitors of the Coronavirus 3CLpro: A Virtual Screening and Structure-Based Drug Design Study. Viruses. 2015;7(12):6642-60.
56. Elfiky AA, Mahdy SM, Elshemey WM. Quantitative structure-activity relationship and molecular docking revealed a potency of anti-hepatitis C virus drugs against human corona viruses. Journal of medical virology. 2017;89(6):1040-7.
57. Zaher NH, Mostafa MI, Altaher AY. Design, synthesis and molecular docking of novel triazole derivatives as potential CoV helicase inhibitors. Acta pharmaceutica. 2020;70(2):145-59.

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)

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

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 all three docking programs revealed binding energies < -7 kcal/mol Amino acid residues forming hydrogen bonds are labeled in bold.

Dataset	ROC probability	VINA blind (LBE)	VINA defined (LBE)	Autodock defined (LBE)	Interacting amino acid residues
FDA-approved drugs:					
Simeprevir	0.993	-8.73±0.85	-8.00±0.46	-9.81±0.31	**Arg346**, Tyr351, Ser494, Leu492, Phe490, Asn450, Phe347, Ala348
Paritaprevir	0.997	-9.47±1.16	-8.43±0.64	-8.52±0.16	**Arg346, Ser349, Tyr351**, Asn450, Ser494, Tyr449, Ile468, Ala348
Grazoprevir	0.995	-8.97±0.38	-7.80±0.78	-8.36±0.34	**Ser494, Gly496, Gln498,** Leu492, Tyr505, Asn501, Gly496, Tyr449, Tyr495
Teniposide	0.997	-9.13±0.25	-7.17±0.12	-7.93±0.24	Val357, Asn437, Trp436, Phe374, Cys336, **Asn343**, **Ser373**
Velpatasvir	0.999	-9.47±0.15	-8.90±1.13	-7.77±0.18	**Ser494**, Phe490, Leu492, Tyr449, Tyr452, Tyr451, Ala348, Arg346
Rifabutin	0.992	-8.83±0.61	-6.40±0.10	-7.50±0.58	**Asn450**, Asn448, Tyr449, Ser494, Leu492, Phe490, Leu452
Ledipasvir	0.999	-9.07±0.25	-8.63±0.76	-6.72±0.54	Leu441, Asn440, Asn437, Trp436, Ser375, Ala372, Ser373, Phe374, Tyr369
Ivermectin	0.999	-9.00±1.18	-8.13±1.02	-6.48±0.39	**Asn440**, Ser438, Asn437, Asn439, Ala372, Ser373, Trp436, Leu368, Phe342
Everolimus	0.996	-9.57±0.06	-8.77±1.36	-6.16±0.61	Arg346, Asn343, Thr345, Trp436, Ser373, Leu368, Gly339, Phe342
Nystatin	0.994	-8.33±1.18	-6.93±0.59	-6.00±0.40	**Glu340**, Asn343, Phe342, Gly339, Phe338, Ser373
Natural compounds from literature:					
Loniflavone	0.996	-9.97±0.15	-9.63±0.57	-8.18±0.06	**Arg509**, Ala348, Tyr451, Ser349, Ala352, **Asn354,** Thr345
Amyrin	0.996	-9.00±0.69	-8.70±1.05	-7.42±0.00	Ser375, Ala372, Phe374, Asn370, Tyr369, Phe377, Lys378
Procyanidin	0.999	-8.60±0.00	-8.13±0.45	-7.29±0.15	**Asn440**, **Ser373**, **Ala372**, Phe374, Asn343, Phe342, Thr345
Phillyrin	0.999	-8.00±0.10	-7.97±0.75	-6.72±0.26	**Arg346**, Lys444, Ile468, Asn450, Phe347, Ala352, Arg509, **Asp442,** Asn448
Proanthocyanidin	0.999	-7.84±0.63	-6.73±0.32	-6.05±0.77	**Arg346**, **Asn354**, **Lys356**, Ser399, **Ala397**, Asn343, **Phe347**, Thr345
Sericoside	0.999	-8.43±0.61	-7.93±1.15	-5.13±0.09	**Asn343, Ser373**, Asn437, **Phe342**, Leu368, Trp436
Punicalagin	1.000	-9.70±0.70	-9.10±0.82	-4.74±0.04	**Arg346, Lys356**, Asn450, Thr345, Phe347, Ala344, **Lys356, Glu340**
Strictinin	0.998	-9.03±0.29	-8.13±0.92	-4.63±0.35	**Asn440, Ser373, Ser375**, **Asn437**, Leu441, Trp436, Phe374, **Asn343**
Rutin	0.999	-8.17±0.45	-7.10±0.72	-4.61±0.78	**Asn343, Ser373, Asn437**, Phe374, Ala372, Ser371, Val367
Tirucallina	1.000	-10.07±0.49	-7.90±1.65	-3.05±0.70	**Ala372, Thr345**, Ser371, Asn440, Asn437, Leu441, Trp436, Arg509, Asn343, Phe342
Natural compounds from ZINC database:					

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^{10},^{15}]hexadeca-10,12,14-trien-9-yl]-2,5,16-triazatetracyclo[7.7.0.0^{2},^{7}.0^{10},^{15}]hexadeca-10(15),11,13-triene-3,6-dione	0.997	-10.37±0.23	-9.97±0.61	-8.89±0.04	**Val357, Ser373**, Leu441, Leu368, Phe342, Phe338, Asn343
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	-10.47±0.83	-9.83±0.35	-8.32±0.35	**Thr470, Ser494, Ser349, Arg346,** Phe490, Leu492, Ala352, Leu452
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^{10},^{15}]hexadeca-10,12,14-trien-9-yl]-2,5,16-triazatetracyclo[7.7.0.0^{2},^{7}.0^{10},^{15}]hexadeca-10(15),11,13-triene-3,6-dione	0.997	-10.90±0.10	-9.77±1.07	-7.91±0.13	Trp436, Phe374, Ser373, Ala372, Asn343, Gly339
ZINC000253532663; 3-(10-[(4-methoxyphenyl)methyl]-4,9,13,15,29-pentamethyl-2,5,8,11,14,30-hexaoxo-24-([3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy)-22-oxa-3,6,9,12,15,29-hexaazatetracyclo[14.12.2.2^{18},^{21}.1^{23},^{27}]tritriaconta-18,20,23(31),24,26,32-hexaen-7-yl)propanoic acid	0.999	-9.70±1.11	-8.67±0.85	-7.50±0.41	**Tyr351, Asn450, Tyr451**, Leu452, Asp442, Arg346, Asn354, Trp353, Ala352, Ser349
ZINC000257466563; Saikosaponin E	0.999	-9.67±1.50	-8.67±0.98	-7.38±0.05	**Asn440**, Ser371, Leu368, **Ser373**, Trp436, Asn343, Phe342, Val367
ZINC000253389151; (2S,3R,4S,5R,6R)-6-([(3S,6aR,6bS,8aS,12aS,14bR)-4,4,6a,6b,11,11,14b-heptamethyl-8a-(([(2S,3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy)carbonyl)-1,2,3,4,4a,5,6,6a,6b,7,8,8a,9,10,11,12,12a,14,14a,14b-icosahydropicen-3-yl]oxy)-3,4,5-trihydroxyoxane-2-carboxylic acid	0.998	-10.33±0.35	-8.83±1.27	-6.93±0.47	**Lys444**, Asn450, Tyr451, Leu441, Arg509, Thr345, **Phe342**, Asn343
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	-10.57±0.06	-9.47±0.99	-5.97±0.32	**Arg346**, Thr345, Ala352, Ala348, Asn450, **Asn354**, Asn448, Asp442, Leu441
ZINC000253389415; (2S,3S,4S,5R,6R)-6-([(3S,6aR,6bS,8aS,12aS,14bR)-8a-(([(2S,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)carbonyl)-4,4,6a,6b,11,11,14b-heptamethyl-1,2,3,4,4a,5,6,6a,6b,7,8,8a,9,10,11,12,12a,14,14a,14b-icosahydropicen-3-yl]oxy)-3,4,5-trihydroxyoxane-2-carboxylic acid	0.999	-10.87±0.15	-9.63±2.02	-5.36±0.26	**Arg335**, Arg346, Ile468, Ala348, Trp353, Asn354, Ser349, Leu452
ZINC000253387436; (2S,3S,4S,5R,6R)-6-)[(3S,4R,6aR,6bS,8aS,14bR)-4-(hydroxymethyl)-4,6a,6b,11,11,14b-hexamethyl-8a-()[(2S,3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy)carbonyl)-1,2,3,4,4a,5,6,6a,6b,7,8,8a,9,10,11,12,12a,14,14a,14b-icosahydropicen-3-yl]oxy)-3,4,5-trihydroxyoxane-2-carboxylic acid	0.999	-9.70±1.14	-8.33±1.62	-5.31±0.66	**Asn437, Asn439**, Leu441, Arg509, Trp436, Ser373, Asn343, Phe342
ZINC000252504401; (2S,3R,4S,5S,6R)-6-()[(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)methyl)-3,4,5-trihydroxyoxan-2-yl (4aR,6aR,6bR,9R,10R,11R,12aR)-10,11-dihydroxy-9-(hydroxymethyl)-1,2,6a,6b,9,12a-hexamethyl-1,4,4a,5,6,6a,6b,7,8,8a,9,10,11,12,12a,12b,13,14,14a,14b-icosahydropicene-4a-carboxylate	0.999	-9.90±0.56	-8.30±0.87	-4.41±0.40	**Arg466, Ile468**, Thr470, Ser469, **Phe464,** Asp467, Tyr351

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 blind (LBE)	VINA defined (LBE)	Autodock defined (LBE)	Interacting amino acid residues
FDA-approved drugs:					
Conivaptan	0.995	-9.13±0.58	-8.77±0.06	-9.93±0.38	Thr57, Arg107, Tyr109, Ala156, Ile157, Val158, Glu174, Gly175, Arg177
Ergotamine	0.993	-9.43±0.40	-9.20±0.00	-9.49±0.91	**Gly69**, Gln70, Ile74, Asn75, Thr76, **Gln83**, Thr135, Glu136, Gln160
Venetoclax	0.995	-9.37±0.35	-9.00±0.20	-8.89±0.73	Ile74, Asn154, Ala155, Ala156, Ile157, Val158, Leu159, Gln160, Leu161, Pro162, Thr166, Leu167, Glu174, Ser176
Eribulin	0.999	-8.90±0.50	-8.37±0.55	-7.46±0.08	Leu161, Pro162, Gly164, Thr166, Leu167, Tyr172, Ser176
Rifapentine	0.999	-9.17±0.21	-9.07±0.06	-6.88±0.03	Val158, Gln160, Leu161, Pro162, Thr166, Leu167, Tyr172, Ala173, **Glu174**, Ser176, Arg177
Dihydroergotamine	0.999	-8.67±0.72	-8.30±0.00	-9.27±0.41	Trp52, Thr54, Arg107, Tyr109, Asn154, Ala155, Ala156, **Ile157**, Val158, Arg177
Rifabutin	0.999	-9.20±0.50	-8.43±0.40	-7.74±0.01	Arg149, Ala155, **Ala156**, Ile157, Val158, Gln160, Ser176, **Arg177**
Natamycin	0.999	-9.00±0.52	-8.70±0.00	-5.77±0.09	Gln70, Ile74, **Pro80, Asp81**, Gln83, Thr135, **Glu136**
Nystatin	0.999	-8.53±1.18	-8.13±0.49	-5.16±0.14	Arg107, Ala155, Ala156, Ile157, Val158, Tyr172, Ser176, Arg177, Gly178, Gly179, **Ser180**
Valrubicin	0.999	-8.83±0.51	-8.87±0.21	-4.37±0.26	Val158, Leu159, Gln160, Leu161, **Pro162**, Ala173, Gly174, Ser176
Natural compounds from literature:					
Amyrin	0.998	-8.90±0.17	-8.80±0.00	-8.80±0.01	Thr49, Ala50, Ser51, Arg88, Ala90, Arg92, Tyr109, Phe110, Tyr111, Gly147
Euphol	1.000	-8.40±0.56	-8.30±0.44	-8.73±0.68	Ser51, Trp52, Phe53, Arg88, Ala90, Arg92, Tyr109, Tyr111
Strictinin	0.999	-8.87±0.72	-8.47±0.06	-5.31±0.40	Ala50, Ser51, Phe53, Arg92, Tyr109, Phe110, Tyr111, **Tyr112**, **Glu118**
Procyanidin	0.999	-9.10±0.00	-9.07±0.06	-5.16±0.28	**Gln70**, Val72, Ile74, **Thr76**, Gln83, **Thr135**, Glu136, Pro162
Sericoside	0.999	-8.03±0.06	-8.00±0.00	-4.76±0.24	Glu160, **Leu161**, Pro162, Thr166, Leu167, Tyr172, Ala173, Glu174, Ser176
Punicalagin	1.000	-9.33±0.23	-8.93±0.46	-4.67±0.04	**Leu161**, Pro162, Tyr172, Ala173, Glu174, **Ser176**, Arg177, Ser180
Ilexsaponinb2	1.000	-8.23±0.23	-8.27±0.15	-4.52±0.59	Ile157, Val158, Gln160, Leu161, Pro162, Thr166, Leu167, Ala173, Ser176, Ser180
Ilexsaponinb3	1.000	-8.57±0.31	-8.37±0.12	-3.37±0.44	Ala156, Ile157, Val158, Gln160, Leu161, Tyr172, Glu174, Ser176, Ser180
Forsythiaside	0.999	-8.10±0.53	-7.53±0.06	-2.47±0.17	Ala50, Thr54, Arg92, Tyr109, Arg149, Ala155, Ala156, Val158, Tyr172, Gly175, Ser176
Tirucallina	1.000	-9.13±0.55	-9.33±0.21	-1.65±0.48	**Ile157**, Val158, Gln160, Leu161, **Pro162**, Ser176, **Arg177**, Gly178
Natural compounds from ZINC database:					
ZINC000027215482; (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^{10},^{15}]hexadeca-10,12,14-trien-9-yl]-2,5,16-	0.999	-10.13±0.23	-10.00±0.00	-8.18±0.05	**Ile74**, Thr76, Pro80, Asp81, **Gln83**, Thr135, Glu136, Gly137

triazatetracyclo[7.7.0.0^{2},^{7}.0^{10},^{15}]hexadeca-10(15),11,13-triene-3,6-dione					
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	-10.30±0.00	-10.33±0.06	-7.92±0.70	Trp52, **Asn154**, Ala155, Ala156, Ile157, Val158, Gln160, Gly164, Thr166, Leu167, **Phe171**, Ser176
ZINC000103216961; Fumiquinazoline D	0.999	-10.30±0.00	-10.30±0.00	-7.87±0.01	**Thr49**, Ser51, Tyr109, Tyr111, Tyr112, Gly147
ZINC000253504772; 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	-10.33±0.06	-10.07±0.31	-7.61±0.27	Thr54, **Ala55**, Arg107, Val158, Leu159, Gln160, Leu161, Thr166, Leu167, **Glu174**, Ser176
ZINC000253504760; 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	-10.30±0.00	-10.30±0.00	-7.56±0.59	**Thr49**, Ala50, Ser51, Trp52, Phe53, Thr54, Arg92, Arg107, Tyr109, Phe110, Tyr111, **Tyr112**, Gly147, Arg149, Gly175, Arg177
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	-11.13±0.06	-11.17±0.06	-7.41±0.38	Ala155, Ala156, Ile157, Val158, Gln160, Leu161, Thr166, Leu167, Ala173, Ser176, Ser180
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	0.999	-10.30±0.00	-10.30±0.00	-7.18±0.68	Thr54, Ala55, Arg107, Tyr109, Ala155, Ala156, Val158, Leu159, Gln160, Leu161, Pro162, Gly164, Glu174, Ser176, **Arg177**
ZINC000253394134; (2E,4Z,8R,9S,10S,11R,13R,18R,22Z)-27,28-dihydroxy-9,15-dimethyl-7,12,20,26,29-pentaoxaspiro[hexacyclo[21.5.2.1^{8},^{11}.0^{1},^{25}.0^{9},^{18}.0^{13},^{18}]hentriacontane-10,2'-oxirane]-2,4,14,22-tetraene-6,21-dione	0.999	-10.20±0.00	-10.20±0.00	-6.78±0.01	Gln160, Leu161, Pro162, Gln163, Gly164, Leu167, Ala173, Glu174
ZINC000226650999; [5-(6-benzamidopurin-9-yl)-3,4-bis[(2,4-dichlorobenzoyl)oxy]oxolan-2-yl]methyl 2,4-dichlorobenzoate	0.999	-10.07±0.25	-8.83±0.84	-6.49±0.20	Val158, Leu159, Gln160, Leu161, Pro162, Gly164, Thr166, Leu167, Tyr172, Ala173, **Gly175**, Ser176
ZINC000253500795; 3-hydroxy-6-[(9-hydroxy-4,4,6a,6b,8a,11,11,14b-octamethyl-1,2,3,4,4a,5,6,6a,6b,7,8,8a,9,10,11,12,14a,14b-octadecahydropicen-3-yl)oxy]-4-[(3,4,5-trihydroxy-6-methyloxan-2-yl)oxy]-5-[(3,4,5-trihydroxyoxan-2-yl)oxy]oxane-2-carboxylic acid	1.000	-10.30±0.00	-10.30±0.00	-5.71±0.34	Ala155, Ala156, Ile157, Val158, Gln160, Leu161, Pro162, Gly164, Ala173, Glu174, Ser176

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 blind (LBE)	VINA defined (LBE)	Autodock defined (LBE)	Interacting amino acid residues
FDA approved drugs:					
Dihydroergotamine	1.000	-10.40±1.80	-11.90±0.00	-12.74±0.29	**Asp75**, Met131, Tyr132, Asp130, Asn43, Tyr47, Ser74, Leu100, Asp99, Cys115, Gly73, Ala72
Nilotinib	1.000	-9.80±0.36	-9.70±0.17	-11.75±0.13	**Asn101**, Gly71, Asp130, Met131, Asp133, Ser74, Tyr47, Asn43, Gly81, Pro80, Asp99, His69, Gly73, Ala72
Telithromycin	1.000	-10.00±0.17	-8.80±0.17	-10.99±0.18	**Asn29, Leu100, Lys170, Asn198**, Glu147, Lys148, Phe150, Phe149, Asp133, Tyr132, Pro134, Ser201, Asn29, Asp99, Asp114
Posaconazole	1.000	-9.67±0.64	-9.43±0.06	-8.64±0.46	**Asn138**, Cys25, Lys137, Glu173, Leu27, **Ser202**, Lys135, Thr136, Thr172, **Pro132**, Ser201, Asn29, Glu203, Tyr132, Lys170, Lsy46
Ergotamine	1.000	-10.10±1.56	-9.10±0.46	-12.26±0.67	**Asp114**, Leu100, Met131, Tyr132, Pro134, Lys170, Asp130, Tyr47, Asn43, Pro80, Gly81, Ala72, Gly73
Lumacaftor	0.999	-9.97±0.29	-9.80±0.00	-9.95±0.18	**Cys115**, Phe149, Met131, Tyr132, Lys170, Gly71, Asp114, Leu100, Asp99, Gly73, Asp130
Venetoclax	1.000	-10.63±0.12	-9.50±0.35	-11.20±0.71	**Lys170**, Thr172, Met132, Pro134, Tyr132, Asp130, Lys46, Gly71, Leu100, Asp99, His69, Asn101, Ser74, Asp75, Gly81, Pro80, Ala79, Thr82
Paritaprevir	1.000	-10.47±0.64	-9.67±0.75	-10.89±0.64	**Lys170, Asn198**, Leu27, Ser202, Ser201, Gln28, Ser200, Asn29, Tyr30, Lys170, Asp130, Lys46, Asn198, Leu239, Ser33, Met42, Asn43
Tenoposide	1.000	-10.07±0.50	-7.63±0.49	-9.64±0.10	**Asn29, Lys46, Lys170**, Thr172, Ser202, Ser201, Asn29, Tyr132, Lys170, Asp130, Glu203, Asn198, Asp130, Gly70, Gly73, Asn43, Gly81, Pro80
Ivermectin	1.000	-8.97±1.76	-6.30±1.39	-8.06±0.30	**Gln28, Asn29, Lys170,** Leu27, Ser202, Tyr132, Asp130, Gly73, Ser74, Tyr30
Natural compounds from literature:					
Procyanidin	1.000	-9.00±0.00	-9.20±0.72	-10.64±0.81	**Gly71, Asn101**, Lys170, Asp130, Cys115, Asp114, Gly113, Asn101, Ser98, Ser74, Asp75, Asp99
Loniflavone	1.000	-11.27±0.06	-11.23±0.06	-10.16±0.26	**Asn29, Leu100, Cys115, Tyr132, Asn198,** Asn138, Ser201, Asn198, Gly71, Asp99, Cys115
Tingeninb	0.999	-9.27±0.12	-8.87±0.29	-9.04±0.00	**Leu100, Tyr132**, Lys170, Gly73, Gly71, Asp99, Asn43, Lys46, Asp130, Met131, Lys76
3,5-Dicaffeoylquinic Acid	1.000	-9.53±0.40	-9.30±0.10	-7.28±0.23	**Cys115**, **Lys146**, Glu147, Gly148, Phe149, Asp133, Tyr132, Met131, Asp114, Cys115
3,4,Dicaffeoylquinicacid	1.000	-8.87±0.12	-8.97±0.06	-7.23±0.27	**Ser33**, Tyr30, Ser201, Lys170, Asp130, Gly71, His69, Ala72, Gly81, Met42
4,5,Dicaffeylquinicacid	1.000	-8.80±0.17	-9.00±0.00	-7.21±0.15	**Lys46**, **Asn198**, Glu173, Thr172, Tyr132, Lys170, Glu203, Leu27, Gly73, Gly71
Strictinin	1.000	-9.97±0.21	-9.83±0.06	-6.54±0.29	**Gln28, His174, Asn198**, Ser33, Gln28, Thr174, Ser202, Asp32, Asn29
TirucallinA	1.000	-10.60±0.44	-10.37±0.84	-6.27±1.09	**Asn138, Asn198, Lys46**, Val139, Lys137, His174, Tyr132, Lys45, Ser201, Pro134, Asp75, Asn43

Punicalagin	1.000	-9.17±0.06	-5.47±0.31	-6.16±0.04	**lys46, Lys170**, Lys137, Asn138, Thr136, Pro134, Tyr132, Lys170, Met131, Asp130, Asp75, Asp32
Rutin	1.000	-8.83±0.29	-8.73±0.46	-5.86±0.41	**Leu100, Tyr132, Lys135**, Gly148, Ph149, Cys115, Asp114, Asp99, Gly71
Natural compounds from ZINC database:					
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	-11.50±0.17	-11.40±0.17	-11.65±0.29	**Gly81**, Phe149, Met131, Tyr132, Pr134, Met131, Cys115, Asp130, Gly71, Leu100, Asp99, His69, Ala72, Gly73, Thr82, Asp75, Pro80
ZINC000008299969; 3-[(3S,3aR,6S,6aR)-6-[(4-)[1,1'-biphenyl]-4-yl)pyrimidin-2-yl)amino]-hexahydrofuro[3,2-b]furan-3-yl]-1-[3-(trifluoromethyl)phenyl]urea	1.000	-11.03±0.58	-10.70±0.00	-11.03±0.26	**Gly71, Asn43**, Phe149, Met131, Tyr132, Cys115, Leu1100, Asp99, Ser74, Asp75, Pro80, Lys76
ZINC000253504772; 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	1.000	-11.77±0.42	-11.57±0.46	-11.02±0.10	**Asn29, Leu100, Lys170, Asn198,** Gly148, Phe150, Phe149, Met131, Lys170, Asn198, Ser201, Asn29, Asp199
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	-11.47±0.38	-11.27±0.06	-10.59±0.47	**Leu100, Glu147, Lys170, Asn198,** Gly148, Phe149, Cys115, Asp114, Asp130, Lys46, Gly71, Met42
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	-11.10±0.87	-10.10±0.00	-10.35±0.16	**Ser33, Tyr30, Ser201, Lys170, Asp130, Gly71, His69, Ala72, Gly81, Met42**
ZINC000003841299; N-[(6aS,8S)-2-(4-chlorophenyl)-6,12-dioxo-5,6a,7,8,9,10-hexahydropyrido[2,1-c][1,4]benzodiazepin-8-yl]cyclopentanecarboxamid	1.000	-11.50±0.00	-11.50±0.00	-10.11±0.02	**Tyr132**, Thr172, Met131, Asp130, Lys170, Gly71, Leu100, Asp99
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	1.000	-11.47±0.29	-11.30±0.00	-10.01±1.52	**Asn29, GLN28, Asn198**, Ser201, Glu203, Tyr132, Lys170, Asp75, Met42, Ser33, Val197
ZINC000004222225; N-[(6As,8S)-6,12-dioxo-2-[3-(trifluoromethyl)phenyl]-5,6a,7,8,9,10-hexahydropyrido[2,1-c][1,4]benzodiazepin-8-yl]pyrazine-2-carboxamide	1.000	-12.10±0.00	-12.10±0.00	-9.95±0.04	**Gly71**, Gly73, Ser74, Phe149, Asp133, Tyr132, Met131, Cys115, Asp114, Leu100, Asp99, Glu147, Gly148, Lys146, Asp133
ZINC000253532087; 3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl 9-(hydroxymethyl)-2,2,6a,6b,9,12a-hexamethyl-10-[(3,4,5-trihydroxy-6-)[(3,4,5-trihydroxy-6-methyloxan-2-yl)oxy]methyl)oxan-2-yl)oxy]-1,2,3,4,4a,5,6,6a,6b,7,8,8a,9,10,11,12,12a,12b,13,14b-icosahydropicene-4a-carboxylate	1.000	-10.17±1.67	-9.20±0.00	-8.54±0.58	**Gln28, Asn29**, His174, Val139, Glu173, Thr172, Ser202, Gln28, Leu27, Gly71, Asn101, Asp99, Ser74, Gly73, Asp75, Asp130
ZINC000253532091; 3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl 9-(hydroxymethyl)-2,2,6a,6b,9,12a-hexamethyl-10-[(3,4,5-trihydroxy-6-)[(3,4,5-trihydroxy-6-methyloxan-2-yl)oxy]methyl)oxan-2-yl)oxy]-1,2,3,4,4a,5,6,6a,6b,7,8,8a,9,10,11,12,12a,12b,13,14b-icosahydropicene-4a-carboxylate	1.000	-10.20±1.56	-9.47±0.29	-8.25±1.14	**Lys170, Asp130, Gly73, Lys46**, **Asn43**, Asp75, Ser74, Glu203, Glu173, Leu27, Gln28, Asn29, **Ser202**

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

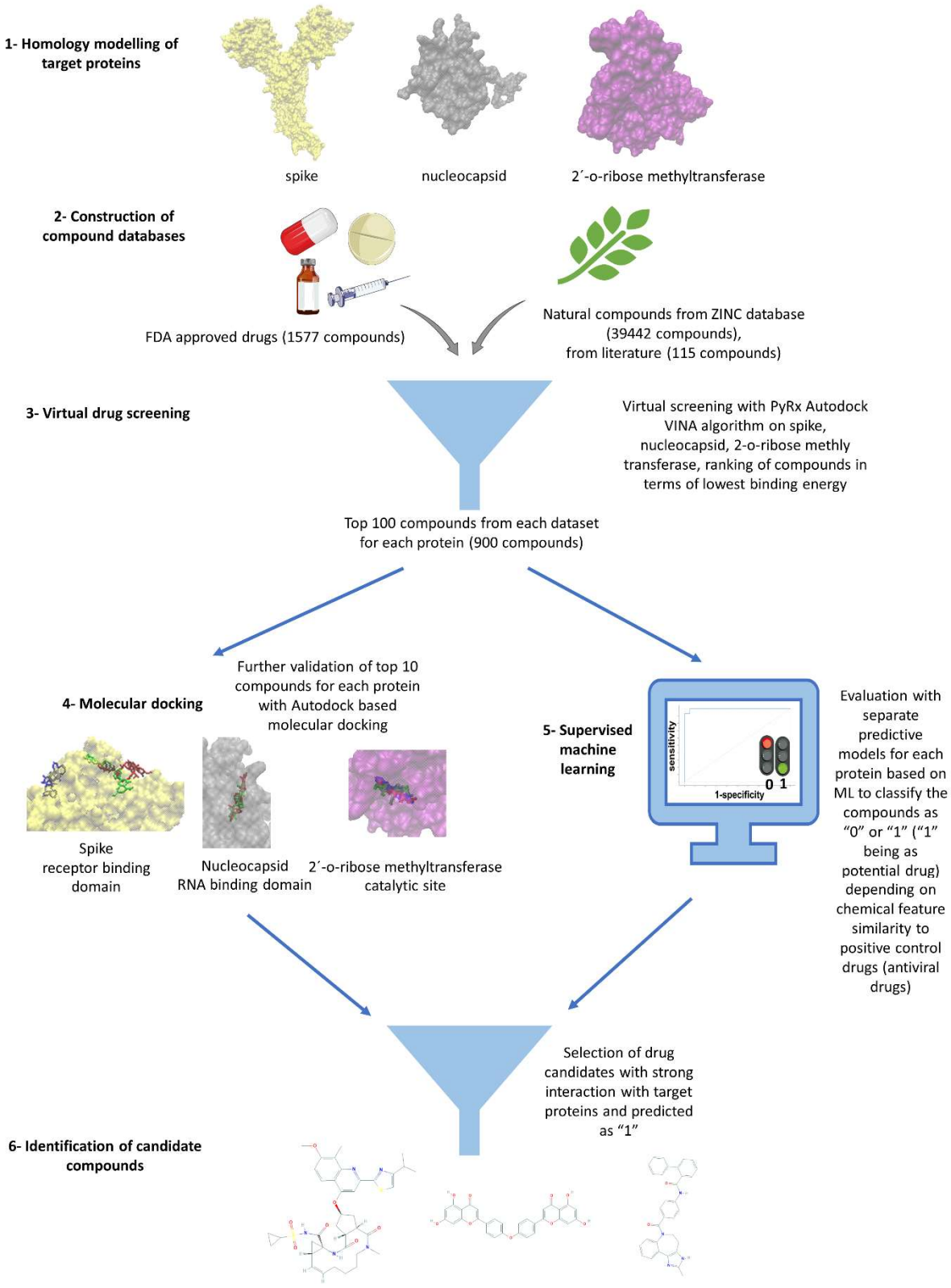

Figure 2: Receiver operating characteristic (ROC) curves for spike protein (A), nucleocapsid protein (B), 2'-o-ribose-methyltransferase (C).

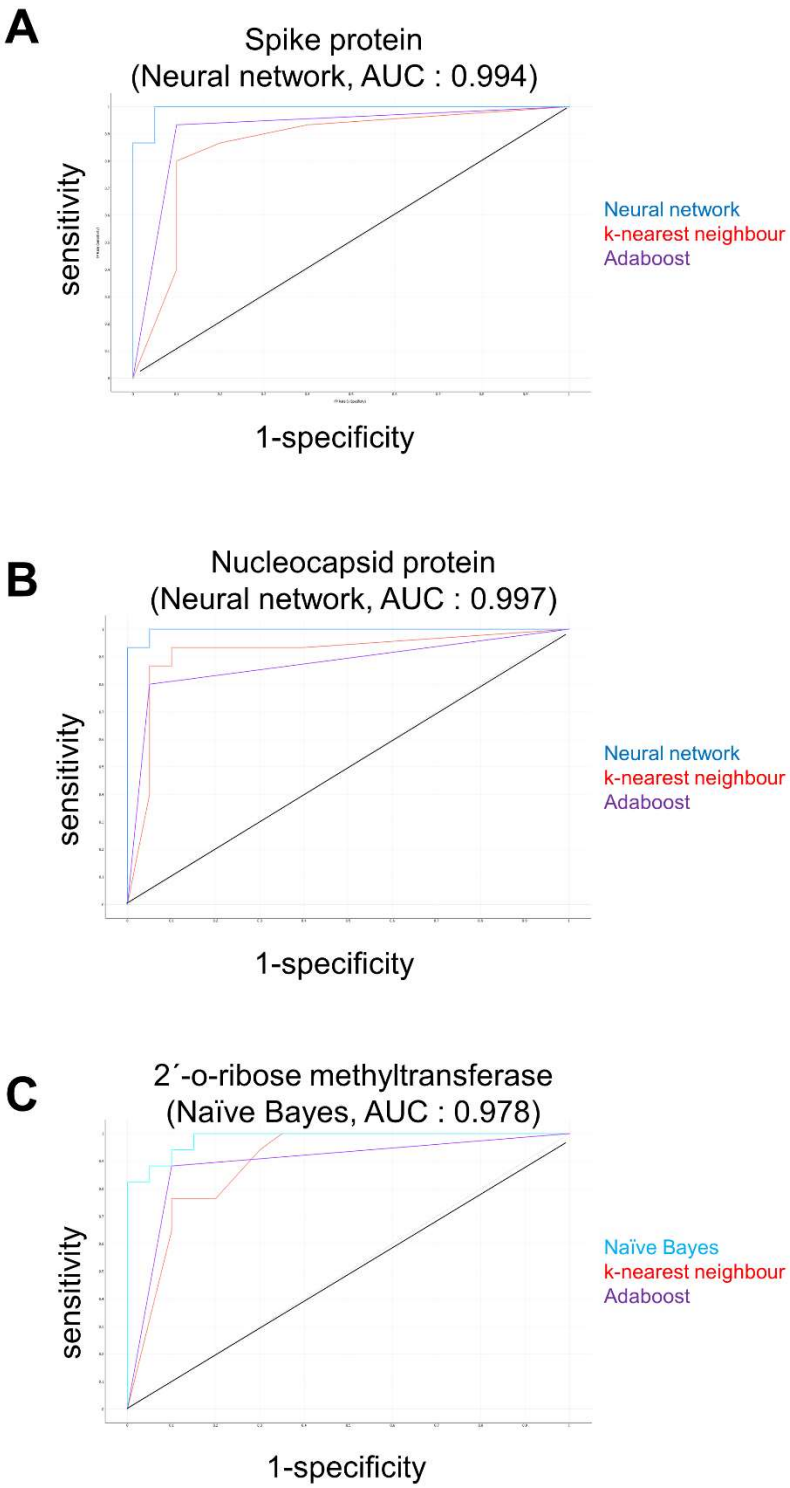

Figure 3: Docking poses of simeprevir (red), loniflavone (green) and ZINC27215482 (blue) on spike protein (yellow). Residues forming hydrogen bonds are labelled bold.

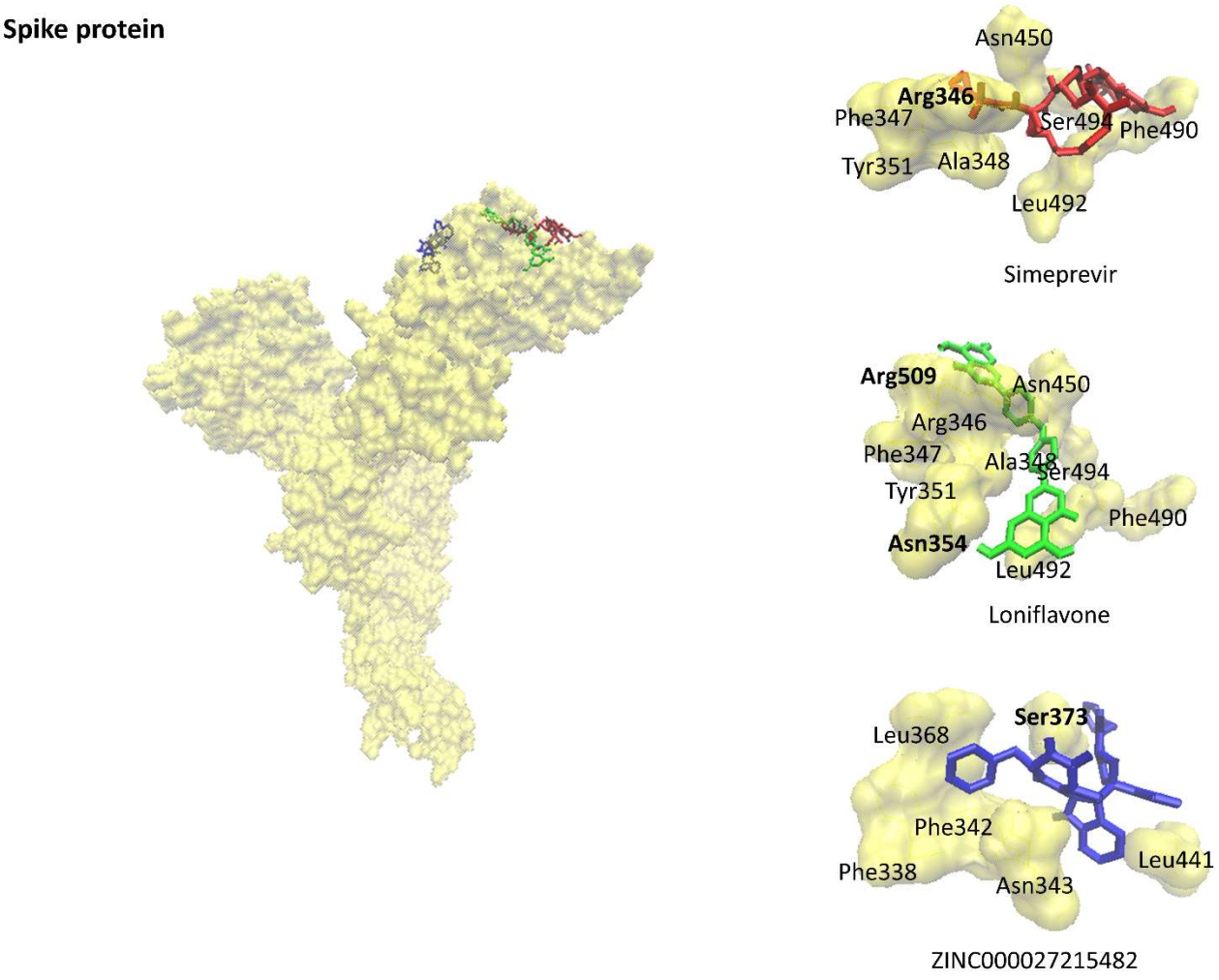

Figure 4: Docking poses of conivaptan (red), amyrin (green) and ZINC27215482 (blue) on nucleocapsid protein (gray). Residues forming hydrogen bonds are labelled bold.

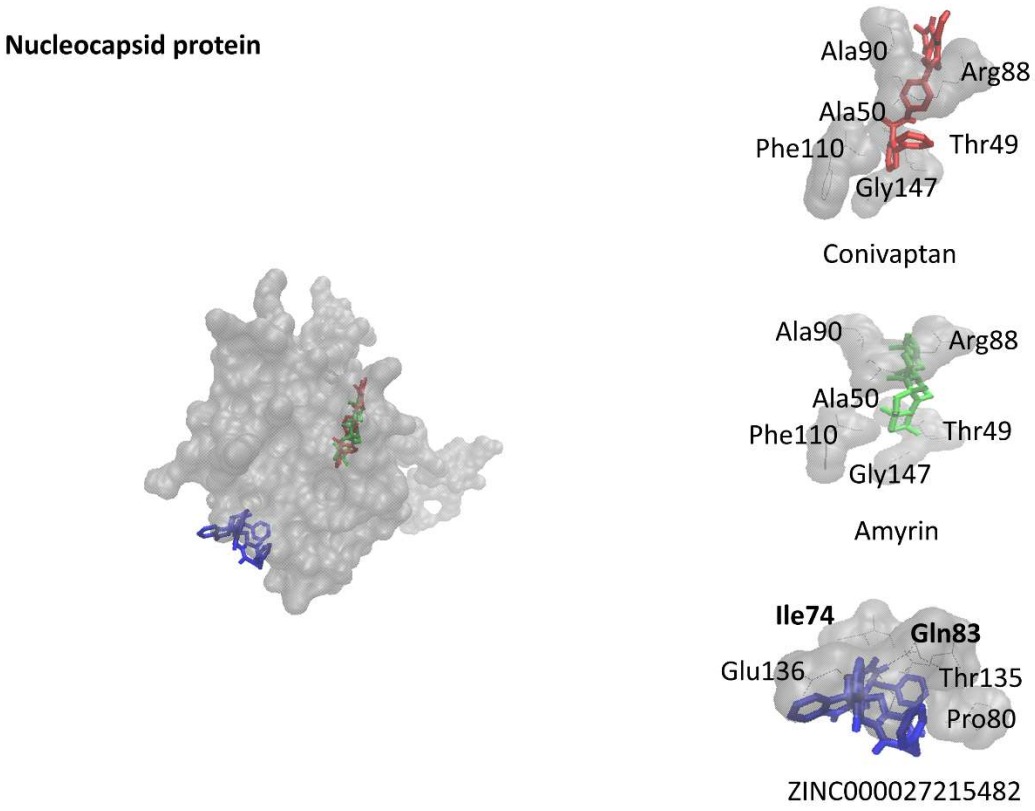

Figure 5: Docking poses of dihydroergotamine (red), procyanidin (green) and ZINC8635475 (blue) on 2'-o-ribose-methyltransferase (purple). Residues forming hydrogen bonds are labeled bold.