Ro5 Bioactivity Lab: Identification of Drug Candidates for COVID-19

Zeyu Yang [1,*] , **Orestis Bastas** [1,*] , **Mikhail Demtchenko** [1] , **Aurimas Pabrinkis** [1] , **Cooper Stergis Jamieson**[2,3] , **Danius Bačkis**[1] , **Charles Dazler Knuff** [3] , **Žygimantas Jočys**[1,3] and **Roy Tal**[1,3,†]

Abstract

The public health emergency known as the coronavirus disease 2019 (COVID-19), caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has led to a large number of deaths worldwide and major socioeconomic disruption. To date, no broadly effective antiviral treatment or vaccine has been developed for COVID-19. In response to this dire situation, Ro5 deployed its AI Lab to accelerate the search for potential treatments. This report focuses on our use of the Ro5 Bioactivity model, which has been designed to predict the inhibitory activity of small molecules against protein targets. The model screened a vast range of compounds *in silico* to uncover potential inhibitors of the SARS-CoV-2 3CL protease. We hereby present the most propitious candidates from this screen. The highest-ranking molecules include Nelfinavir, Saquinavir, Itacitinib, Kynostatin-272, BOG-INS-6c2-1, and BEN-VAN-d2b-11. Subsequent docking simulations corroborate their plausibility as 3CLpro inhibitors. Nelfinavir and Itacitinib hold the most promise for drug repurposing, among all the molecules proposed herein, due to their high predicted inhibition and affinity against the 3CL protease, favourable pharmacokinetics, and encouraging experimental data for treating viral replication and hyperinflammation, respectively.

1 Introduction

The coronavirus disease 2019 (COVID-19) has caused a global health crisis [WHO, 2020b]. As of May 7 2020, more than 3.5 million confirmed cases and almost 250,000 deaths have been associated with the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [WHO, 2020a]. Coronaviruses are enveloped, positive-sense RNA viruses, able to infect a range of animals species, including humans. Several strains of human coronaviruses are known to cause acute and chronic diseases [Weiss and Leibowitz, 2011]. SARS-CoV-2 is believed to be of zoonotic origin and closely related to bat coronaviruses [Zhu *et al.*, 2020].

The main 3-chymotrypsin-like protease (3CLpro) and papain-like protease (PLpro) are the enzymes responsible for the post-translation cleavage of the replicase polyproteins PP1a and PP1ab. As an integral part of the viral reproduction process, 3CLpro cleaves the polyprotein at 11 sites into individual functional proteins using the thiol group of the cysteine residue, located in its active site [Muramatsu *et al.*, 2016]. 3CLpro is a proven drug discovery target for past respiratory syndrome coronaviruses and is currently under concerted study as an inhibitory target to disrupt the viral replication process of SARS-CoV-2 [ul Qamar *et al.*, 2020; Sisay, 2020].
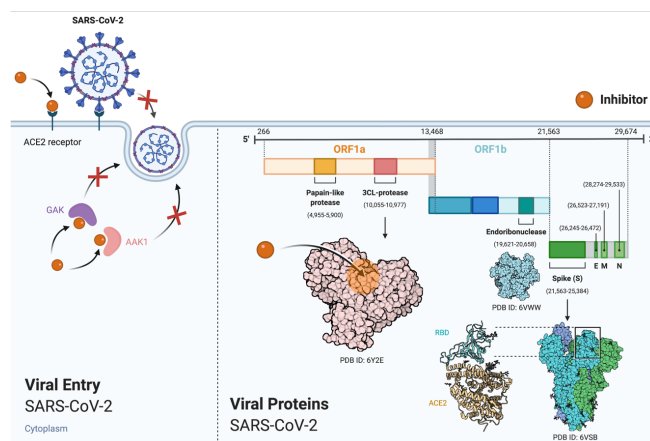

Figure 1: Potential inhibitor targeting pathways. The left panel illustrates the entry of SARS-CoV-2 into a host cell by binding to the ACE2 receptor through its spike protein. This leads to clathrin-mediated endocytosis, for which the GAK and AAK1 enzymes play a key role. The right panel highlights the genomic organisation of SARS-CoV-2 and a few of its important proteins. Figure created with Biorender.com

[1] Ro5, East Side, Platform 1, King's Cross Station Euston Road, London N1C 4AX, United Kingdom

[*] Equal contribution

[2] Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095, United States

[3] Ro5, 2801 Gateway Drive, Irving, Texas 75063, United States

[†] **Corresponding author: rtal@ro5.ai**

Figure 1 displays potential targets for the inhibition of SARS-CoV-2, including important viral proteins necessary for its replication process and those associated with its infection of host cells. Hoffmann et. al. demonstrated that SARS-CoV-2 uses the angiotensin-converting enzyme 2 (ACE2) as the entry receptor. ACE2 is the same receptor that SARS-CoV, the virus that caused the 2003 SARS outbreak, used for host cell entry [Hoffmann *et al.*, 2020]. Cyclin G-associated kinase (GAK) and adaptor-associated protein kinase 1 (AAK1) are two important kinases involved in intracellular membrane trafficking; they were previously studied as broad-spectrum antiviral drug targets [Schor and Einav, 2018]. By taking the approach of inhibiting the kinases fundamental to the endocytosis process, BenevolentAI used their artificial intelligence (AI)-derived knowledge graph to identify a group of approved drugs that could potentially inhibit the infection of cells by SARS-CoV-2 [Richardson *et al.*, 2020; Stebbing *et al.*, 2020]. A recent review based on clinical evidence postulated that Janus kinase (JAK) inhibitors could be used as immunosuppressants to mitigate the hyperinflammation characteristic of COVID-19, as systemic inflammation is associated with adverse outcomes [Ritchie and Singanayagam, 2020; Russell *et al.*, 2020].

Following the outbreak of the COVID-19 pandemic in January 2020, the global scientific community, from all disciplines, has been actively researching disease treatment and prevention. Both experimental and computational high-throughput screenings are underway at unprecedented speed. Diamond Light Source solved the structure of the main protease at high resolution and conducted a crystal-based fragment screen with 3CLpro. Of the 66 active site fragments identified, 44 were covalently bound [Diamond, 2020]. This set of data allowed PostEra to initiate the Moonshot project that crowdsources fragment-based rational designs of molecules from chemists. Molecules selected based on synthesisability and toxicity modelling will be synthesised and subjected to experimental testing [PostEra, 2020]. Meanwhile, many academic and industrial initiatives have screened *in silico* against a variety of chemical databases with the hope of accelerating drug development for COVID-19 [Wu *et al.*, 2020; Contini, 2020; Teams *et al.*, 2020; Zhavoronkov *et al.*, 2020]. As part of this effort, we decided to deploy the Ro5 AI Lab, which contains machine learning models for Bioactivity and ADMET predictions, augmenting clinical trials and *de novo* generation of new drug candidates. In this report, we would like to offer our insight from applying Ro5's Bioactivity model towards identifying appropriate 3CLpro inhibitors.

2 Methodology

2.1 Data Collection

Protein Targets

Testing targets included the 3CLpro, spike protein, RNA-binding domain of the nucleocapsid phosphoprotein, PL2pro, and a range of non-structural proteins (NSP3, 10, 12, 15). In addition, we selected a range of proteins involved in the cellular uptake of the virus and the hyperinflammation associated with COVID-19. These included the ACE2 spike receptor-binding domain, AAK1 and GAK enzymes, and the JH1 kinase domain of the JAK enzyme. Some of the larger proteins were truncated to sequences of approximately 200 amino acids around the active site. This allowed our model to focus on the well-studied inhibition region and ensured that the tested target lengths were more similar to the training set distribution. The analyses presented herein focus on 3CLpro, a proven target for developing SARS-CoV-2 inhibitors; the other targets screened will be subject to further research.

Compounds

In order to locate the relevant molecules and related annotations in our internally curated database, we used a comprehensive search routine to discover information both directly and indirectly tagged to the virus entry and replication pathways. Furthermore, we prioritised compounds possessing antiviral or anti-inflammatory activity. This was incorporated into our screening dataset, which also consisted of the drug repurposing database from the Broad Institute, the DrugBank database, the crowdsourced compound designs submitted to the Moonshot project, potential therapeutics proposed by BenevolentAI and Insilico Medicine, as well as several promising molecules suggested by the scientific community. Preprocessing of compounds included stripping counterions and removing duplicates. A total of 18,943 unique compounds were present in the final dataset.

2.2 Machine Learning

We used Ro5's Bioactivity machine learning model, which was trained to predict the inhibitory activity level between a small molecule and a protein target. The model was designed to learn from graph representations of molecules by generating a vector to capture their chemical structure and reactivity. Combining this with a protein representation derived from sequence information alone, the model arrives at a bioactivity prediction for the pair.

A crucial feature of this machine learning model is the attention mechanism, which creates a graph-based attention map highlighting the bonds and atoms in the molecule that the model deems important. During the model development process, this enables a comparison between model predictions and the qualitative evaluation of chemists regarding the most important areas of the molecule.

For the presented case study, the model was used to evaluate all combinations of compounds and targets. The model predictions were filtered to a set of compounds exhibiting the desired activity level across a set of targets. Furthermore, the molecules were ranked by their probability of inhibiting 3CLpro; the five highest-scoring compounds were subsequently assessed according to their chemical and pharmacological relevance. Exploratory analysis of the results allowed for the identification of additional interesting drug candidates.

2.3 Docking

Structures of interest were covalently and non-covalently docked into the 3CLpro using AutoDock4 and AutoDock Vina [Morris *et al.*, 2009; Trott and Olson, 2010]. First, the structures of two identified ligands, kynostatin-272 (PDBe ID: KNI) and saquinavir (PDBe ID: ROC), were obtained from the European Bioinformatics Institute. The geometries of the remaining ligands were optimised using Grimme's program xtb [Bannwarth *et al.*, 2019]. Covalent docking parameters were generated using MGLTools supplied with AutoDock4 and AutoDockTools. Covalent AutoDock4 simulations used 10 Lamarckian genetic algorithm docking runs centered on a grid of (65,61,61) cartesian points spaced at 0.375 Å centered on the nucleophilic cysteine 145 (CYS145). AutoDock Vina simulations were exhaustively run in a 30 Å grid box centered on the CYS145.

3 Results

By predicting the probability of inhibitory activity between molecules and the target, our model identified several compounds highly likely to be 3CLpro inhibitors. The top five ranked molecules overall are presented in Table A1.

One of the molecules, DB07964, is an inhibitor of the endogenous disintegrin and metalloproteinase 17 (ADAM17). Evidence suggests that ADAM17 cleaves and thereby mediates shedding of the ACE2 ectodomain, leading to a reduction in cellular ACE2 [Grobe *et al.*, 2015]. Since SARS-CoV-2 binds to the ACE2 receptor during infection [Hoffmann *et al.*, 2020], ADAM17 might play a natural role in decreasing coronavirus-ACE2 complexes and consequently hindering viral entry. Cellular experiments suggest that ADAM17 and transmembrane protease serine 2 (TMPRSS2) compete for ACE2 cleavage and only proteolysis by TMPRSS2 promotes spike protein-driven entry [Levin *et al.*, 2006]. Therefore, antagonism of ADAM17 would be unfavorable for a COVID-19 treatment. As a consequence, DB07964 and its analogue, Apratastat, which possesses the same ADAM17-binding moiety, hydroxamate, were excluded from further analyses.

This report thus focuses on three candidates from the top five overall ranked compounds. Itacitinib, one of these molecules, is an anti-inflammatory drug that recently fell short of meeting its primary endpoint for the treatment of acute graft-versus-host disease (GVHD) in phase 3 clinical trials. The two remaining candidates are novel compounds that were proposed through generative modelling and fragment-based lead discovery. In addition, we report three, among several, noteworthy compounds from the top 5% of the dataset, by examining their structure-activity relationship and comparing them to the Diamond's ligand screening results. These include Nelfinavir and Saquinavir, two HIV drugs that have been approved for market, and Kynostatin-272, an HIV protease inhibitor that was found to be inadequately efficacious as an HIV therapeutic after phase 1 clinical trials.

These six molecules were then subjected to docking against the main protease of SARS-CoV-2. Four of the proposed compounds contain amide linkages and are either a polypeptide or a peptidomimetic. Other common fragments and protein-ligand interactions will provide insight to further optimise the lead compounds. Herein, we report the potential binding poses of these compounds and their available pharmacological data.

3.1 Highest-scoring compounds
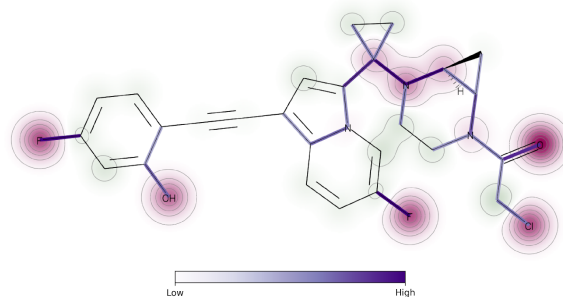
BOG-INS-6c2-1

Figure 2: BOG-INS-6c2-1 molecule with an overlaid attention map. The Ro5 Bioactivity model focused on the oxygen and chlorine of the alpha-chloroketone, the fluorines, and the hydroxyl group.

Insilico Medicine used their fragment-based generative model to generate and optimise plausible lead candidates [Zhavoronkov *et al.*, 2020]. One of their proposed molecules ranked highly according to our model. This compound (Figure 2) contains an alpha-chloroketone moiety, which is a fragment known to undergo nucleophilic substitution by sulphur, making it possible to form a covalent bond with CYS145 at the active site of 3CLpro. Non-covalent docking revealed that the inolizine inserts deeply into the binding pocket with a moderate binding affinity of -7.8 kcal/mol (Figure A6). The fluorine may also form a strong halogen-pi interaction with His41. As in the non-covalent docking simulations, the covalent docking simulations, where the alpha-chloroketone reacts to form a thioate linkage (Figure A7), show a weak hydrogen bond between the phenol and the Glu166 backbone. Autodock4 predicts this binding pose to have a poor binding affinity of +11.2 kcal/mol.

We applied the in-house Ro5 ADMET prediction tools to analyse BOG-INS-6c2-1, since no experimental data is available for this compound. Our model predicted that the compound should have good cell permeability based on Caco-2 cell line predictions and that it is not a P-gp substrate, in line with prediction results from SwissADME [Daina *et al.*, 2017].

Itacitinib

Itacitinib (Figure 3), a Janus-associated kinase inhibitor, selective towards JAK-1, has been tried against melanoma, endometrial cancer, B-cell malignancies, and more recently against GVHD [Im *et al.*, 2019]. The JAK-STAT signalling
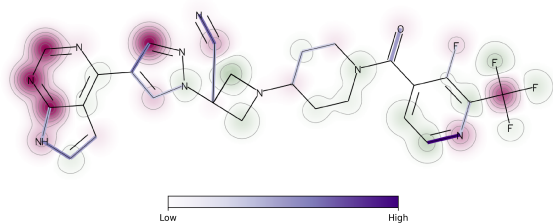
Figure 3: Itacitinib molecule with an overlaid attention map. The attention map shows that the Ro5 Bioactivity model focused on parts of the heteroaromatic rings and the carbon of the trifluoromethyl group.
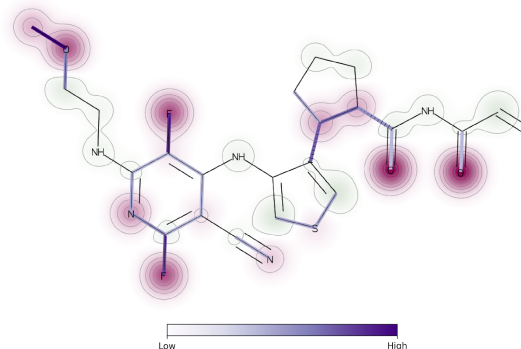

Figure 4: BEN-VAN-d2b-11 molecule with an overlaid attention map. The Ro5 Bioactivity model focused on the oxygens of the acetylacetamide group, the fluorines, and the ether linkage.

pathway plays an important role in the host inflammatory response. Our model predicts that Itacitinib could be a potent 3CLpro inhibitor. By targeting both 3CLpro and JAK, Itacitinib could potentially disrupt the viral replication process of SARS-CoV-2 and also reduce the damaging inflammatory response that is observed in the later, more severe stages of the disease, where patients suffer from the consequences of hyperinflammation [Richardson *et al.*, 2020]. Interestingly, our model ranked Itacitinib to more actively inhibit 3CLpro than would Baricitinib, a marketed JAK inhibitor recently proposed as a potential COVID-19 treatment by Benevolent AI's knowledge graph [Stebbing *et al.*, 2020]. Akin to Baricitinib, the Ro5 model also predicts that Itacitinib is likely to have inhibitory activity against GAK and AAK1, suggesting that it could potentially be a virus uptake inhibitor as well. Detailed docking studies and *in vitro* testing are needed to confirm its mechanism of action. Furthermore, it should be noted that the JAK-STAT pathway is also involved in the body's natural antiviral response; thus, caution must be taken when modulating the JAK protein. Results from ongoing clinical studies will help ascertain whether the JAK-STAT pathway is indeed a promising target for treating COVID-19.

Based on studies of nitrile-containing inhibitors, Itacitinib could potentially act as a covalent inhibitor against 3CLpro [Furber *et al.*, 2014]. Non-covalent docking (Figure A3) shows a favourable binding affinity of -8.3 kcal/mol. It should be noted that the electrophilic nitrile is positioned next to the nucleophilic cysteine, which could lead to covalent binding. The CF$_3$ forms electrostatic interactions with Thr25 and Ser46. The covalent docking simulations, where the nitrile and cysteine react to form a thioimidate adduct (Figure A4), predict a binding affinity of –5.6 kcal/mol. In the covalently docked structure, the piperidine and pyridine extend from the binding pocket, whereas the indole nitrogen forms a hydrogen bond with Tyr54. Remarkably, Itacitinib was independently proposed by another drug repurposing screening study due to the its potential as a nitrile covalent inhibitor [Liu *et al.*, 2020]; however, the molecule was not tested *in vitro*.

BEN-VAN-d2b-11

One of the compounds (Figure 4) submitted to the Moonshot project ranked highly according to the Bioactivity model. Brown *et al.* generated linkers to combine fragments dis-

covered by Diamond's crystal-based screening [Brown and Meiler, 2020]. The compounds were optimised, scored, and filtered based on interaction poses and predicted physicochemical properties [Brown and Meiler, 2020]. Docking results showed a -6.9 kcal/mol binding affinity, where the acetylacetamide forms a hydrogen bond with the Glu166 backbone and the ether chain inserts deeply into the active site pocket (Figure A8). This compound has potential to be a covalent inhibitor due to the presence of a Michael acceptor moiety and an aromatic nitrile functional group. Covalent docking simulations, where the acrylamides react to form a sulfide linkage (Figure A9), show a weak hydrogen bond between the pyridine and the Glu166 backbone. The binding affinity of this pose is predicted to be -2.2 kcal/mol. Our ADMET predictions suggest that the compound is unlikely to be a P-gp substrate and that it has low cell permeability. SwissADME predicts the molecule to be a P-gp substrate with low cell permeability. Furthermore, our microsomal clearance model predicted the compound to have a high hepatic clearance, which would render it unfeasible for clinical development, if indeed verified by *in vitro* testing.

3.2 Noteworthy compounds

Considerable effort has gone towards evaluating the suitability of repurposing existing antiviral drugs towards COVID-19 treatment. Moreover, another widely studied enveloped, positive-sense RNA virus is the Human Immunodeficiency Virus (HIV). If left untreated, HIV can lead to acquired immunodeficiency syndrome (AIDS). Several potent HIV drugs are on the market, including Atazanavir, Ritonavir, and Lopinavir/Ritonavir. Use of Lopinavir/Ritonavir for treatment during the outbreak in China sparked wide interest globally. A later randomised controlled trial, however, concluded that no benefit was observed in hospitalised adult patients with severe COVID-19 [Cao *et al.*, 2020]. Preliminary results from a retrospective, single-center case study conducted in Hangzhou, China has shown that the combination of Lopinavir/Ritonavir, interferon, and arbidol could be an effective antiviral therapy, especially for common cases of the disease in adults [Yu *et al.*, 2020]. Therefore, HIV drugs are

of particular interest to us and three HIV drugs from the top 5% of the screened dataset were selected for further analysis.
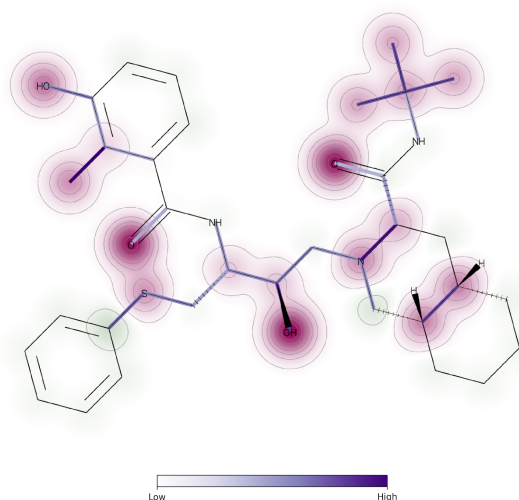
Nelfinavir

Figure 5: Nelfinavir molecule with an overlaid attention map. The attention map shows that the Ro5 Bioactivity model focused on the hydroxyl groups, oxygens of the amide linkages, and the *tert*-butyl group.

Nelfinavir (Figure 5) is a marketed antiretroviral drug with activity against HIV Type 1 (HIV-1). It binds to the HIV-1 protease active site and thereby inhibits the activity of the enzyme. By blocking the cleavage of the gag-pol polyprotein, the virion maturation process is hindered, resulting in immature non-infectious viral particles [Bardsley-Elliot and Plosker, 2000]. Subsequent docking analyses revealed a -7.1 kcal/mol binding affinity. The favourable interactions (Figure A1) include a hydrogen bond between the nitrogen of *tert*-butylacetamide and the His164 backbone amide oxygen, a hydrogen bond between the phenol group of Nelfinavir and the backbone of Thr24, as well as a weak pi–pi stacking with the neighboring phenyl sulfane of Nelfinavir. The hydrophobic *tert*-butyl moiety inserts deeply into the binding pocket, further stabilising the binding pose.

Recent studies subjected Nelfinavir to experimental testing against SARS-CoV-2 in the VeroE6/TMPRSS2 cell line. Data from several independent groups identified Nelfinavir as a potent replication inhibitor of SARS-CoV-2 and showed that Nelfinavir protected cell viability twentyfold compared to the untreated control group [Xu *et al.*, 2020; Ohashi *et al.*, 2020]. This is consistent with the predictions of the Ro5 Bioactivity model, as well as the docking results. Nelfinavir showed promising *in vitro* data with EC50 and EC90 values of 1.13 uM and 1.76 uM, respectively, which was the highest efficacy observed out of all nine HIV protease inhibitors tested, including lopinavir [Yamamoto *et al.*, 2020]. These results, coupled with a selective index (SI) of 12.99 and a suitable serum concentration (Cmax/EC50 = 6.23, Ctrough/EC50

= 3.43) make Nelfinavir a potential drug candidate to treat COVID-19.
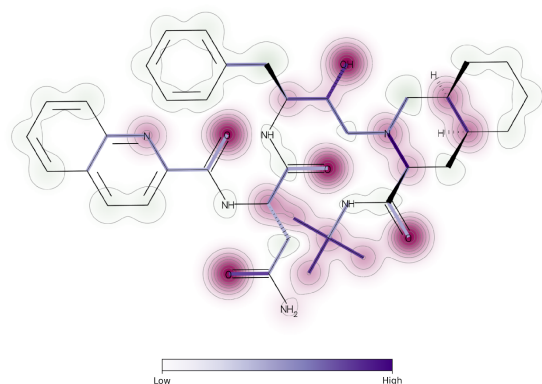
Saquinavir

Figure 6: Saquinavir molecule with an overlaid attention map. The attention maps shows that the Ro5 Bioactivity model focused on the oxygens of the amide linkages, the hydroxyl group, and the *tert*-butyl group.

Saquinavir (Figure 6) is an HIV-1 and HIV-2 protease inhibitor that ranked highly in our inhibitory activity model. Docking shows a binding affinity of -9.3 kcal/mol, the strongest binding among the six compounds. The binding site (Figure A2) is highly hydrophilic leading to many electrostatic interactions. Decahydroisoquinoline points into a hydrophobic pocket. Quinoline extends outwards from the protein cleft and is exposed to the solvent environment, which reduces binding stability. The plausible cation-pi interaction between toluene and HIS41 contributes to stabilise the compound in the binding site. Similar to Nelfinavir, the toluene moiety resides deep within the the pocket.

Saquinavir was one of the drugs tested by Yamamoto *et al.*, among eight other HIV protease inhibitors [Yamamoto *et al.*, 2020]. Despite its relatively high efficacy against SARS-CoV-2 *in vitro* (EC50 = 8.83 uM) and its high predicted binding affinity to 3CLpro, the low serum concentration *in vitro* (Cmax/EC50 = 0.73) suggests that Saquinavir is a less suitable drug candidate than Nelfinavir.

Kynostatin-272

Kyostatin-272 (KNI-272, Figure 7) is a transition-state mimetic tripeptide inhibitor of the HIV-1 protease. Despite demonstrating promising inhibitory activity *in vitro* and a suitable serum concentration, KNI-272 showed insufficient clinical activity in phase I clinical trials [Humphrey *et al.*, 1999]. The low *in vivo* activity was attributed to the unfavorable pharmacokinetics of KNI-272, since it is likely to be 98-99% protein-bound in the circulating blood. Retrospectively, our microsomal clearance ADMET model predicted a relatively high clearance value for KNI-272, coinciding with the fact that KNI-272 is metabolized by hepatic P450 enzymes. Docking predicted a -7.1 kcal/mol binding affinity, for which
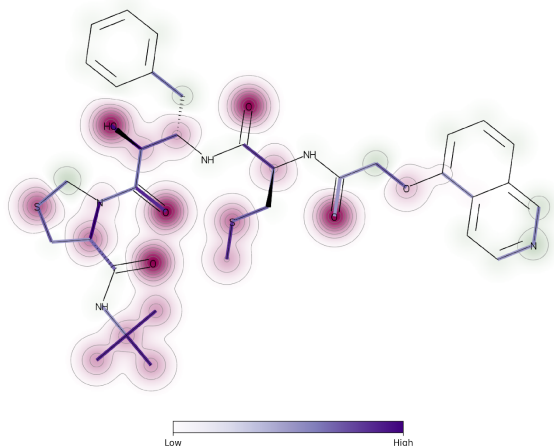
Figure 7: KNI-272 molecule with an overlaid attention map. The attention map shows that the Ro5 Bioactivity model focused on the oxygens of the amide linkages, the hydroxyl group, the *tert*-butyl group, and the sulphurs.

the main interactions include a hydrogen bond and hydrophobic interactions (Figure A5).

4 Discussion

The Bioactivity machine learning model of the Ro5 AI Lab uses protein sequence information alone to derive a vector representation of the target. This allowed us to test novel targets without the explicit input of structural information, eliminating the need for computationally expensive homology modelling or labour intensive crystallography. This enables a fast response towards drug repurposing as well as novel drug discovery. Repurposing unsuccessful drug candidates that have gone through several stages of clinical trials is a valuable approach; these candidates will typically have documented pharmacokinetic data and safety information. Testing them against a novel disease may accelerate the drug development process and lower costs.

Saquinavir and Kynostatin-272, two molecules that efficaciously target and inhibit HIV proteases, were shown to be suitable inhibitors against 3CLpro of SARS-CoV-2, according to our Bioactivity model and subsequent docking analyses. Notwithstanding, their pharmacokinetic disadvantages make them less viable drug candidates compared to our highest-scoring molecules.

BOG-INS-6c2-1 and BEN-VAN-d2b-11, the two novel compounds identified by the Bioactivity model, were also demonstrated to be plausible inhibitors of 3CLpro. This supports the notion that our model can reliably predict the efficacy and affinity of unseen molecules against an unseen target protein. However, since they are completely novel compounds, it is unrealistic to develop them within the short window of time necessary to tackle the current COVID-19 public health emergency.

In the current crisis, the most appropriate drug candidates are those that allow for speed and ease of both production and patient prescription. Nelfinavir is a clinically approved drug, and Itacitinib has passed the first two phases of clinical trials, giving the two the capacity to be repurposed in a relatively short period of time. Furthermore, they exhibit high predicted inhibition and affinity against the 3CL protease, favourable pharmacokinetics and encouraging experimental data for preventing viral replication and hyperinflammation, respectively, in COVID-19 patients.

5 Conclusion

Although the pandemic appears to be stabilising in some countries, several areas are experiencing a resurgence in COVID-19 cases [WHO, 2020a]. This volatility, combined with a lack of evidence for immunity after infection by the virus [GM, 2020] and the long wait until a vaccine can be made available, makes finding a treatment for this disease the urgent priority.

Our Bioactivity model enabled us to independently identify clinically viable drug repurposing candidates, corroborated by other computational and experimental publications. Furthermore, two novel compounds are predicted to be 3CLpro inhibitors, and we presented their predicted pharmacology. In the urgent quest to identify drug candidates that can improve outcomes for COVID-19 patients and prevent the further spread of the disease, we hope these findings bring insight and speed.

References

[Bannwarth *et al.*, 2019] Christoph Bannwarth, Sebastian Ehlert, and Stefan Grimme. GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *Journal of Chemical Theory and Computation*, 15(3):1652–1671, March 2019.

[Bardsley-Elliot and Plosker, 2000] Anne Bardsley-Elliot and Greg L. Plosker. Nelfinavir. *Drugs*, 59(3):581–620, March 2000.

[Brown and Meiler, 2020] Benjamin P. Brown and Jens Meiler. Submission to the Moonshot project: BEN-VAN-d2b, August 2020.

[Cao *et al.*, 2020] Bin Cao, Yeming Wang, Danning Wen, Wen Liu, Jingli Wang, Guohui Fan, Lianguo Ruan, Bin Song, Yanping Cai, Ming Wei, Xingwang Li, Jiaan Xia, Nanshan Chen, Jie Xiang, Ting Yu, Tao Bai, Xuelei Xie, Li Zhang, Caihong Li, Ye Yuan, Hua Chen, Huadong Li, Hanping Huang, Shengjing Tu, Fengyun Gong, Ying Liu, Yuan Wei, Chongya Dong, Fei Zhou, Xiaoying Gu, Jiuyang Xu, Zhibo Liu, Yi Zhang, Hui Li, Lianhan Shang, Ke Wang, Kunxia Li, Xia Zhou, Xuan Dong, Zhaohui Qu, Sixia Lu, Xujuan Hu, Shunan Ruan, Shanshan Luo, Jing Wu, Lu Peng, Fang Cheng, Lihong Pan, Jun Zou, Chunmin Jia, Juan Wang, Xia Liu, Shuzhen Wang, Xudong

Wu, Qin Ge, Jing He, Haiyan Zhan, Fang Qiu, Li Guo, Chaolin Huang, Thomas Jaki, Frederick G. Hayden, Peter W. Horby, Dingyu Zhang, and Chen Wang. A Trial of Lopinavir–Ritonavir in Adults Hospitalized with Severe Covid-19. *New England Journal of Medicine*, 0(0):null, March 2020. Publisher: Massachusetts Medical Society _eprint: https://doi.org/10.1056/NEJMoa2001282.

[Contini, 2020] Alessandro Contini. Virtual Screening of an FDA Approved Drugs Database on Two COVID-19 Coronavirus Proteins. *ChemRxiv*, February 2020. Publisher: ChemRxiv.

[Daina *et al.*, 2017] Antoine Daina, Olivier Michielin, and Vincent Zoete. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific Reports*, 7, March 2017.

[Diamond, 2020] Diamond. Diamond Light Source - XChem fragement screen, March 2020.

[Furber *et al.*, 2014] Mark Furber, Anna-Karin Tiden, Philip Gardiner, Antonio Mete, Rhonan Ford, Ian Millichip, Linda Stein, Andrew Mather, Elizabeth Kinchin, Christopher Luckhurst, Simon Barber, Peter Cage, Hitesh Sanganee, Rupert Austin, Kamaldeep Chohan, Raj Beri, Bob Thong, Alan Wallace, Victor Oreffo, Ray Hutchinson, Steve Harper, Judit Debreczeni, Jason Breed, Lisa Wissler, and Karl Edman. Cathepsin C Inhibitors: Property Optimization and Identification of a Clinical Candidate. *Journal of Medicinal Chemistry*, 57(6):2357–2367, March 2014. Publisher: American Chemical Society.

[GM, 2020] Flodgren GM. Immunity after sars-cov-2 infection. April 2020.

[Grobe *et al.*, 2015] Nadja Grobe, Mauricio Di Fulvio, Nada Kashkari, Harshita Chodavarapu, Hari K Somineni, Richa Singh, and Khalid M Elased. Functional and molecular evidence for expression of the renin angiotensin system and adam17-mediated ace2 shedding in cos7 cells. *American Journal of Physiology-Cell Physiology*, 308(9):C767–C777, 2015.

[Hoffmann *et al.*, 2020] Markus Hoffmann, Hannah Kleine-Weber, Simon Schroeder, Nadine Krüger, Tanja Herrler, Sandra Erichsen, Tobias S. Schiergens, Georg Herrler, Nai-Huei Wu, Andreas Nitsche, Marcel A. Müller, Christian Drosten, and Stefan Pöhlmann. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell*, 181(2):271–280.e8, April 2020.

[Humphrey *et al.*, 1999] Rachel W. Humphrey, Kathleen M. Wyvill, Bach-Yen Nguyen, Laura E. Shay, David R. Kohler, Seth M. Steinberg, Takamasa Ueno, Tominaga Fukasawa, Makoto Shintani, Hideya Hayashi, Hiroaki Mitsuya, and Robert Yarchoan. A phase I trial of the pharmacokinetics, toxicity, and activity of KNI-272, an inhibitor of HIV-1 protease, in patients with AIDS or symptomatic HIV infection. *Antiviral Research*, 41(1):21–33, February 1999.

[Im *et al.*, 2019] Annie Im, Rodica Morariu-Zamfir, Maureen Bleam, Ying Yan, and Steven Z. Pavletic. Trial in Progress: A Phase 3 Study of Itacitinib or Placebo in Combination with Corticosteroids As Initial Treatment for Chronic Graft-Versus-Host Disease (GRAVITAS-309). *Blood*, 134(Supplement_1):3277–3277, November 2019. Publisher: American Society of Hematology.

[Levin *et al.*, 2006] JI Levin, JM Chen, LM Laakso, M Du, J Schmid, W Xu, T Cummons, J Xu, G Jin, D Barone, et al. Acetylenic tace inhibitors. part 3: Thiomorpholine sulfonamide hydroxamates. *Bioorganic & medicinal chemistry letters*, 16(6):1605–1609, 2006.

[Liu *et al.*, 2020] Sen Liu, Qiang Zheng, and Zhiying Wang. Potential covalent drugs targeting the main protease of the SARS-CoV-2 coronavirus. *Bioinformatics*, January 2020.

[Morris *et al.*, 2009] Garrett M. Morris, Ruth Huey, William Lindstrom, Michel F. Sanner, Richard K. Belew, David S. Goodsell, and Arthur J. Olson. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry*, 30(16):2785–2791, December 2009.

[Muramatsu *et al.*, 2016] Tomonari Muramatsu, Chie Takemoto, Yong-Tae Kim, Hongfei Wang, Wataru Nishii, Takaho Terada, Mikako Shirouzu, and Shigeyuki Yokoyama. SARS-CoV 3CL protease cleaves its C-terminal autoprocessing site by novel subsite cooperativity. *Proceedings of the National Academy of Sciences*, 113(46):12997–13002, November 2016. Publisher: National Academy of Sciences Section: Biological Sciences.

[Ohashi *et al.*, 2020] Hirofumi Ohashi, Koichi Watashi, Wakana Saso, Kaho Shionoya, Shoya Iwanami, Takatsugu Hirokawa, Tsuyoshi Shirai, Shigehiko Kanaya, Yusuke Ito, Kwang Su Kim, Kazane Nishioka, Shuji Ando, Keisuke Ejima, Yoshiki Koizumi, Tomohiro Tanaka, Shin Aoki, Kouji Kuramochi, Tadaki Suzuki, Katsumi Maenaka, Tetsuro Matano, Masamichi Muramatsu, Masayuki Saijo, Kazuyuki Aihara, Shingo Iwami, Makoto Takeda, Jane A. McKeating, and Takaji Wakita. Multidrug treatment with nelfinavir and cepharanthine against COVID-19. *bioRxiv*, page 2020.04.14.039925, April 2020. Publisher: Cold Spring Harbor Laboratory Section: New Results.

[PostEra, 2020] PostEra. PostEra | COVID-19, April 2020.

[Richardson *et al.*, 2020] Peter J. Richardson, Mario Corbellino, and Justin Stebbing. Baricitinib for COVID-19: a suitable treatment? – Authors' reply. *The Lancet Infectious Diseases*, 0(0), April 2020. Publisher: Elsevier.

[Ritchie and Singanayagam, 2020] Andrew I. Ritchie and Aran Singanayagam. Immunosuppression for hyperinflammation in COVID-19: a double-edged sword? *The Lancet*, 395(10230):1111, April 2020. Publisher: Elsevier.

[Russell *et al.*, 2020] Beth Russell, Charlotte Moss, Gincy George, Aida Santaolalla, Andrew Cope, Sophie Papa, and Mieke Van Hemelrijck. Associations between immune-suppressive and stimulating drugs and novel COVID-

19—a systematic review of current evidence. *ecancermedicalscience*, 14, March 2020.

[Schor and Einav, 2018] Stanford Schor and Shirit Einav. Repurposing of Kinase Inhibitors as Broad-Spectrum Antiviral Drugs. *DNA and Cell Biology*, 37(2):63–69, February 2018.

[Sisay, 2020] Mekonnen Sisay. 3CLpro inhibitors as a potential therapeutic option for COVID-19: Available evidence and ongoing clinical trials. *Pharmacological Research*, 156:104779, June 2020.

[Stebbing *et al.*, 2020] Justin Stebbing, Anne Phelan, Ivan Griffin, Catherine Tucker, Olly Oechsle, Dan Smith, and Peter Richardson. COVID-19: combining antiviral and anti-inflammatory treatments. *The Lancet Infectious Diseases*, 20(4):400–402, April 2020. Publisher: Elsevier.

[Teams *et al.*, 2020] Galaxy and HyPhy developments Teams, Anton Nekrutenko, and Sergei L. Kosakovsky Pond. No more business as usual: agile and effective responses to emerging pathogen threats require open data and open analytics. *bioRxiv*, page 2020.02.21.959973, February 2020. Publisher: Cold Spring Harbor Laboratory Section: New Results.

[Trott and Olson, 2010] Oleg Trott and Arthur J. Olson. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *Journal of computational chemistry*, 31(2):455–461, January 2010.

[ul Qamar *et al.*, 2020] Muhammad Tahir ul Qamar, Safar M. Alqahtani, Mubarak A. Alamri, and Ling-Ling Chen. Structural basis of SARS-CoV-2 3CLpro and anti-COVID-19 drug discovery from medicinal plants†. *Journal of Pharmaceutical Analysis*, March 2020.

[Weiss and Leibowitz, 2011] Susan R. Weiss and Julian L. Leibowitz. Chapter 4 - Coronavirus Pathogenesis. In Karl Maramorosch, Aaron J. Shatkin, and Frederick A. Murphy, editors, *Advances in Virus Research*, volume 81, pages 85–164. Academic Press, January 2011.

[WHO, 2020a] WHO. Coronavirus disease (COVID-2019) situation reports. Situation report—96. Technical report, World Health Organisation, April 2020.

[WHO, 2020b] WHO. WHO declare global pandemic. Technical report, World Health Organisation, November 2020.

[Wu *et al.*, 2020] Canrong Wu, Yang Liu, Yueying Yang, Peng Zhang, Wu Zhong, Yali Wang, Qiqi Wang, Yang Xu, Mingxue Li, Xingzhou Li, Mengzhu Zheng, Lixia Chen, and Hua Li. Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. *Acta Pharmaceutica Sinica B*, February 2020.

[Xu *et al.*, 2020] Zhijian Xu, Hangping Yao, Jingshan Shen, Nanping Wu, Yechun Xu, Xiangyun Lu, Weiliang Zhu, and Lan-Juan Li. Nelfinavir Is Active Against SARS-CoV-2 in Vero E6 Cells. *ChemRxiv*, March 2020. Publisher: ChemRxiv.

[Yamamoto *et al.*, 2020] Norio Yamamoto, Shutoku Matsuyama, Tyuji Hoshino, and Naoki Yamamoto. Nelfinavir inhibits replication of severe acute respiratory syndrome coronavirus 2 in vitro. preprint, Microbiology, April 2020.

[Yu *et al.*, 2020] Jianhua Yu, Meihua Wang, Kexing Zhou, Siquan Zhang, Bo Xiang, Yan Zhang, Zuhua Chen, Jinsong Huang, Zhaobin Cai, Shourong Liu, Dingyan Yan, and Zongxing Yang. Clinical characteristics of Coronavirus Disease 2019 in Hangzhou, China: the combination of lopinavir/ritonavir, interferon, and arbidol may be a well choice for antiviral therapy in common cases. preprint, In Review, April 2020.

[Zhavoronkov *et al.*, 2020] Alex Zhavoronkov, Vladimir Aladinskiy, Alexander Zhebrak, Bogdan Zagribelnyy, Victor Terentiev, Dmitry S. Bezrukov, Daniil Polykovskiy, Rim Shayakhmetov, Andrey Filimonov, Philipp Orekhov, Yilin Yan, Olga Popova, Quentin Vanhaelen, Alex Aliper, and Yan Ivanenkov. Potential COVID-2019 3C-like Protease Inhibitors Designed Using Generative Deep Learning Approaches. *ChemRxiv*, February 2020. Publisher: ChemRxiv.

[Zhu *et al.*, 2020] Na Zhu, Dingyu Zhang, Wenling Wang, Xingwang Li, Bo Yang, Jingdong Song, Xiang Zhao, Baoying Huang, Weifeng Shi, Roujian Lu, Peihua Niu, Faxian Zhan, Xuejun Ma, Dayan Wang, Wenbo Xu, Guizhen Wu, George F. Gao, and Wenjie Tan. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New England Journal of Medicine*, 382(8):727–733, February 2020. Publisher: Massachusetts Medical Society _eprint: https://doi.org/10.1056/NEJMoa2001017.
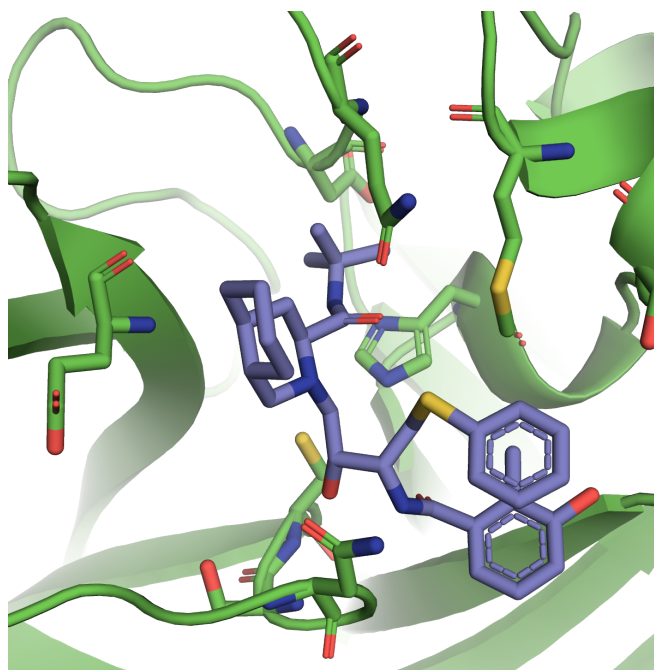
6 Appendix

Figure A1: Docking pose of Nelfinavir in 3CL protease active site. Non-covalent docking predicted the binding affinity to be -7.1 kcal/mol.
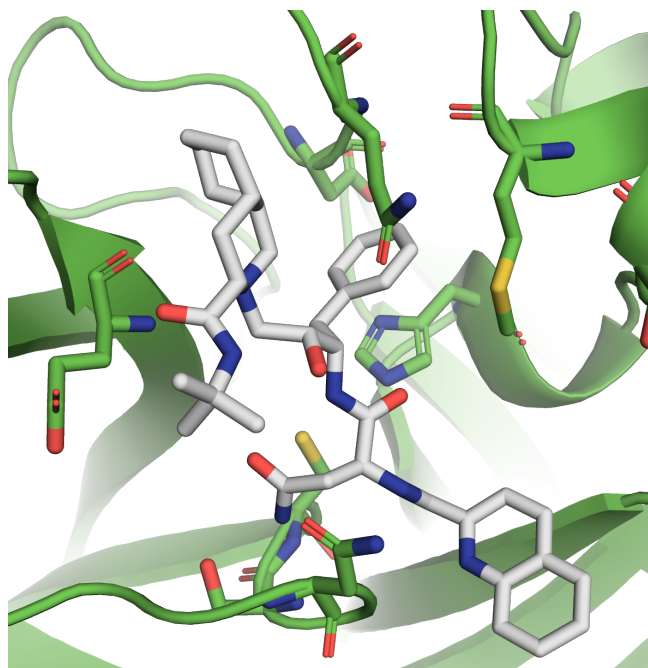

Figure A2: Docking pose of Saquinavir in 3CL protease active site. Non-covalent docking predicted the binding affinity to be -9.3 kcal/mol.
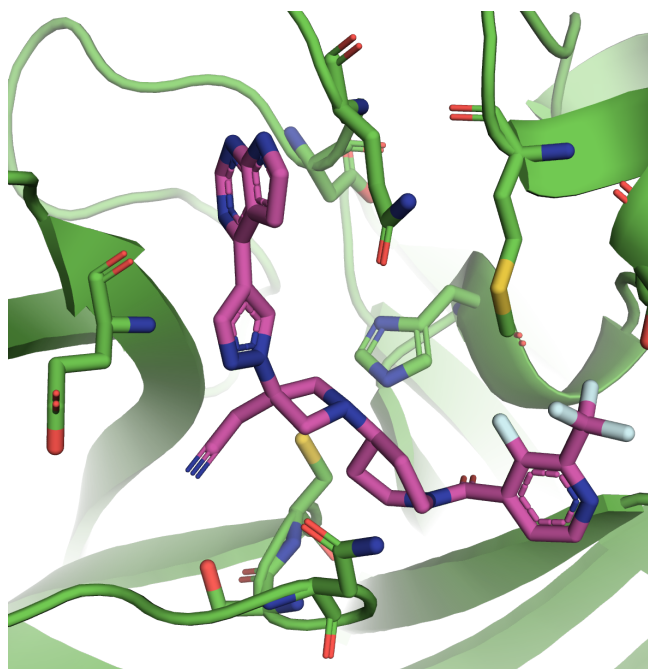

Figure A3: Docking pose of Itacitinib in 3CL protease active site. Non-covalent docking predicted the binding affinity to be -8.3 kcal/mol.
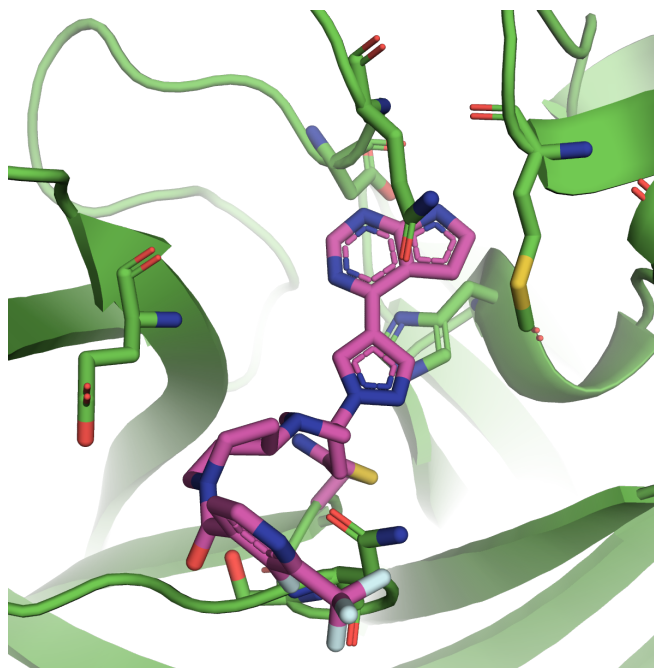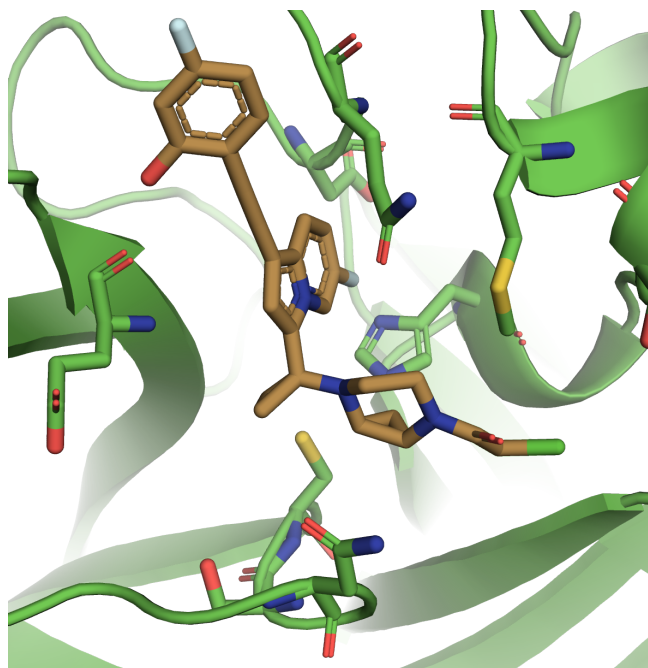
Figure A4: Covalent docking pose of Itacitinib in 3CL protease active site. Covalent docking predicted the binding affinity to be -5.6 kcal/mol.

Figure A6: Docking pose of BOG-INS-6c2-1 in 3CL protease active site. Non-covalent docking predicted the binding affinity to be -7.8 kcal/mol.
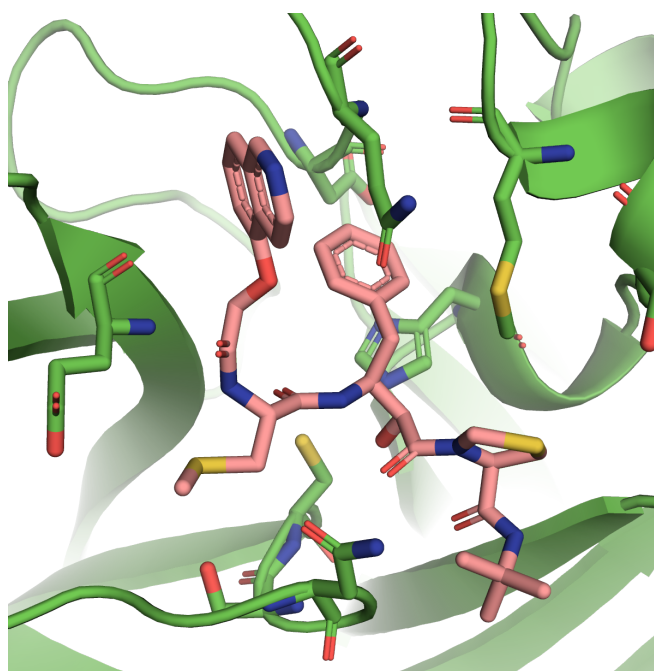

Figure A5: Docking pose of KNI-272 in 3CL protease active site. Non-covalent docking predicted the binding affinity to be -7.1 kcal/mol.
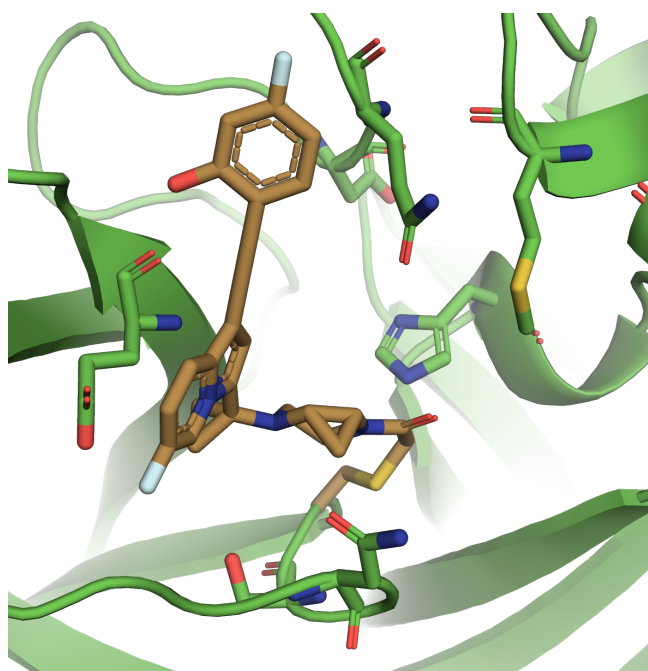

Figure A7: Covalent ocking pose of BOG-INS-6c2-1 in 3CL protease active site. Covalent docking predicted the binding affinity to be +11.2 kcal/mol.
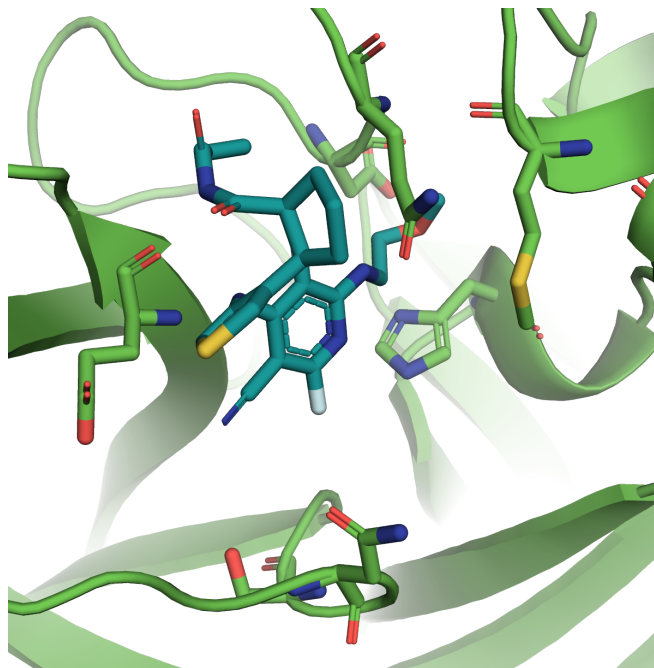
Figure A8: Docking pose of BEN-VAN-d2b-11 in 3CL protease active site. Non-covalent docking predicted the binding affinity to be -6.9 kcal/mol.
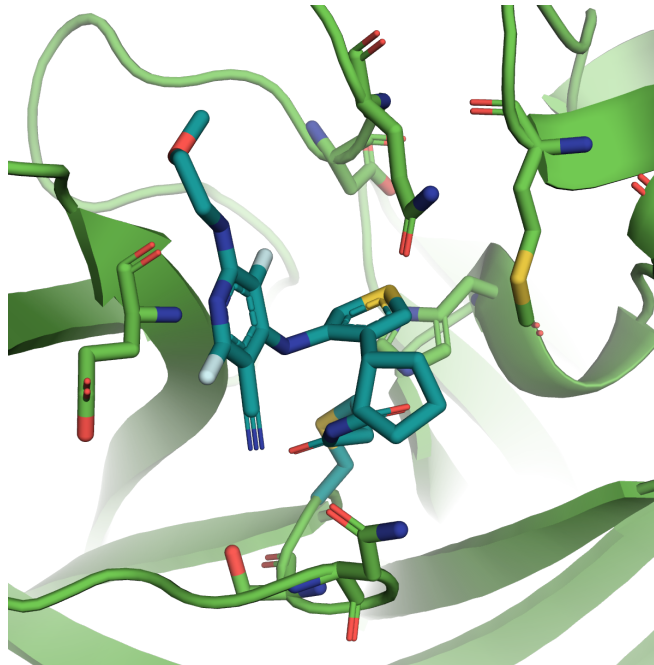

Figure A9: Covalent docking pose of BEN-VAN-d2b-11 in 3CL protease active site. Covalent docking predicted the binding affinity to be -2.2 kcal/mol.

Table A1: The five molecules with the highest probability of being active inhibitors of 3CLpro according to the Bioactivity model. The common name of the drug is included where available.

Ranking	Accession	Name	SMILES
1	BOG-INS-6c2-1		O=C(CCl)N1CCN(C2(c3cc(C#Cc4ccc(F)cc4O)c4ccc(F)cn34)CC2)[C@H]2CC21
2	CHEMBL3622820	Itacitinib	N#CCC1(n2cc(-c3ncnc4[nH]ccc34)cn2)CN(C2CCN(C(=O)c3ccnc(C(F)(F)F)c3F)CC2)C1
3	DB13020	Apratastat	CC1(C)SCCN(S(=O)(=O)c2ccc(OCC#CCO)cc2)[C@H]1C(=O)NO
4	BEN-VAN-d2b-11		C=CC(=O)NC(=O)[C@H]1CCC[C@H]1c1c1cscc1Nc1c(F)c(NCCOC)nc(F)c1C#N
5	DB07964		CC#CCOc1ccc(S(=O)(=O)N2CCSC(C)(C)[C@@H]2C(=O)NO)cc1