Rapid Identification of Potential Inhibitors of SARS-CoV-2 Main Protease by Deep Docking of 1.3 Billion Compounds

*Anh-Tien Ton‡, Francesco Gentile‡, Michael Hsing, Fuqiang Ban, Artem Cherkasov**

Vancouver Prostate Centre, University of British Columbia, 2660 Oak Street, Vancouver, BC V6H 3Z6, Canada

KEYWORDS: COVID-19, deep learning, virtual screening, protease inhibitors

ABSTRACT: The recently emerged 2019 Novel Coronavirus (SARS-CoV-2) and associated COVID-19 disease cause serious or even fatal respiratory tract infection and yet no FDA-approved therapeutics or effective treatment is currently available to effectively combat the outbreak. This urgent situation is pressing the world to respond with the development of novel vaccine or a small molecule therapeutics for SARS-CoV-2. Along these efforts, the structure of SARS-CoV-2 main protease (Mpro) has been rapidly resolved and made publicly available to facilitate global efforts to develop novel drug candidates.

In recent month, our group has developed a novel deep learning platform – Deep Docking (DD) which enables very fast docking of billions of molecular structures and provides up to 6,000X enrichment on the top-predicted ligands compared to conventional docking workflow (without notable loss of information on potential hits). In the current work we applied DD to entire 1.3 billion compounds from ZINC15 library to identify top 1,000 potential ligands for SARS-CoV-2 Mpro. The compounds are made publicly available for further characterization and development by scientific community.

INTRODUCTION

Coronaviruses (CoVs) are enveloped viruses containing a single positive-stranded RNA, and causing a wide array of respiratory, gastrointestinal, and neurological diseases in human hosts[1,2]. It has been established that strains of CoVs were at the source of the 2002 severe acute respiratory syndrome (SARS) and 2012 middle east respiratory syndrome (MERS) epidemics[3]. In late December 2019, a novel CoV of SARS-COV-2 was identified to be the cause of atypical pneumonia outbreak in Wuhan, China, named COVID-19[4]. The rapidly increasing number of infected patients worldwide prompted the World Health Organization to declare a state of global health emergency to coordinate scientific and medical efforts to rapidly develop a cure for patients[5]. While drug repurposing may be a short-term and non-specific solution to treat COVID-19 patients[6], development of more targeted inhibitors is highly desirable.

CoVs main target proteins

Previous research efforts to develop anti-viral agents against members of *Coronaviridae* family demonstrated that the Angiotensin-converting enzyme II (ACE2) entry receptor, the RNA-dependent RNA polymerase (RdRp) and the main protease (Mpro) proteins may represent suitable drug targets[7]. Although initially promising, inhibitors targeting ACE2 (hence aiming to block critical coronavirus-host interactions) did not advance clinically due to significant side effects[8]. Identified RdRp inhibitors appeared to be not very specific and demonstrated overall lower potency, that also translated into common side-effects in patients[1,9]. Concurrently, CoV infected patients administered with protease inhibitors have shown improved outcome[1,10], demonstrating the potential of the main protease (Mpro) as the most promising drug target in CoVs[11,12]. Hence, a recently published X-ray crystal structure of the SARS-COV-2 Mpro provides an excellent ground for structure-based drug discovery efforts[13].

Known CoV main protease inhibitors

Earlier efforts to target SARS-CoV resulted in identification of several covalent Mpro inhibitors targeting the catalytic dyad of the protein defined by His41 and Cys145[14] residues. However, covalent inhibitors are often marked by adverse drug responses, off-target side effects, toxicity and lower potency[15–19]. Therefore, noncovalent protease inhibitors may have advantages for the treatment of SARS-COV-2 infection. Still, the majority of approved drugs administered as anti-SARS were designed for other viral strains (Table S1). Notably, no CoV-protease specific inhibitor has yet successfully completed a clinical development program to date[16,20].

Deep Docking

The impact of current COVID-19 outbreak and the likelihood of future CoV epidemics strongly advocate for rapid development of new treatments and fast intervention protocols. Few research groups have already suggested potential repurposing strategies for clinically approved drugs[21–23] or proposed *de novo* agents[24] as therapeutic solutions for SARS-COV-2. However, previously reported docking (virtual screening) campaigns with Mpro targets were able to process only few millions or even thousands compounds[6,25–27] The main reason for that is that conventional docking is too computationally expensive and slow, while the libraries of available chemicals are growing exponentially[28].

To address this general challenge, we have recently developed a novel deep learning-based approach for accelerated screening of large chemical libraries, consisting of billions of entities. This Deep Docking (DD) platform utilizes quantitative structure-activity relationship (QSAR) models trained on docking scores of database subsets to approximate in an iterative manner the docking outcome of the remaining entries. More details can be found in our recent preprint[29].

Herein we have used DD for large-scale virtual screening against the SARS-COV-2 Mpro active site.

MATERIALS AND METHODS

We used DD to virtually screen all ZINC15 (1.36 billion compounds)[30] against the SARS-COV-2 Mpro. The model was initialized by randomly sampling 3 million molecules and dividing them evenly into training, validation and test set. Most probable tautomer and ionization states at pH 7.4 were calculated with OpenEye QUACPAC package[31] and starting 3D conformations were generated using Omega *pose* routine[32]. The structure PDB 6LU7 (resolution 2.16 Å)[33] of the SARS-COV-2 Mpro bound to a covalent N3 inhibitor was obtained from the Protein Data Bank[34], and prepared for docking using Protein Preparation Wizard[35]. Docking was performed using Glide SP module[36], and computed scores were used for DNN initialization. We then ran 3 iterations, adding each time 1 million of docked molecules sampled from previous predictions to the training set and setting the recall of top scoring compounds to 0.75. The top 1 million molecules predicted to have favorable scores were then docked to the protease site. The set of protease inhibitors (7,800 compounds) from the BindingDB repository was also docked to the same site[37]. Our computational setup consisted of 13 Intel(R) Xeon(R) Gold 6130 CPU @ 2.10GHz (a total of 390 cores) for docking, and 40 Nvidia Tesla V100 GPUs with 32GB memory for deep learning.

RESULTS AND DICUSSION

The use of DD platform enabled us to dock 1.3B compounds from ZINC15 database[30] into SARS-COV-2 main protease active site using standard Glide SP protocol[36] in a week.

The predicted interaction between the top four hits, selected by Glie SP docking score are presented on Figure 1. The data demonstrate that common ligand anchoring interactions correspond to

hydrogen bonds with Cys145 and Leu141 residues. Encouragingly, it appears that our top predicted inhibitor ZINC000541677852 shares a number of features with two known protease inhibitors, which are also likely to bind to the SARS-COV-2 Mpro (Also shown on Figure 1). One of them - Lopinavir is a clinically approved HIV protease inhibitor, which is being evaluated in combination with ritonavir in a randomized controlled trial for SARS-COV-2 infection in China, based on its activity in past CoV epidemics[38]. The drug is a large peptide-like molecule that docked well in the binding site having many contacts with the binding residues due to its size. The second drug molecule termed "Compound 80" is a non-peptide small molecule inhibitor of SARS Mpro, with a reported IC_{50} of 0.95 μM[20,39,40]. Compound 80 and ZINC000541677852 share one hydrogen bond with Cys145; additionally, the two phenyl rings of compound 80 share the two hydrophobic sites of the diazole and 2-ethyl-6-oxopiperidin-4-yl moieties of ZINC000541677852. Also, the trifluoromethyl-phenyl moiety of ZINC000541677852 overlays well with the 1,3-xylene moiety of lopinavir. Thus, our top identified molecule appears to have binding features with the site that are proper of protease inhibitors. Nevertheless, all our compounds featured on Figure 1 demonstrate significantly better docking scores than the two protease inhibitors.

We have also analyzed the origin of top 1,000 ZINC hits (selected by LE), and observed that 99% of them are not present in the commonly used ZINC15 in-stock library (~11 millions of molecules), commonly used in routine docking campaigns, demonstrating that the DD methodology can access complete and diverse chemical space beyond classical docking. The Glide SP scores of the top 1,000 candidates we selected were significantly better than top 1,000 molecules from a 1 million random sample of ZINC15 entries, and even better than top candidates from BindingDB protease inhibitor library, which were docked to the same site (Figure 2).

We also evaluated the chemical diversity of the newly identified set of inhibitors compared to the protease library. Calculation of Murcko frameworks[41] for hits from such library and DD hits revealed a similar number of frameworks present in the two sets (603 and 587 scaffolds, respectively). Encouragingly, we observed just two common frameworks, clearly indicating that screening 1.36 billion enables identification of new chemical classes that can potentially inhibit SARS-COV-2 Mpro. Thus, DD allowed us to rapidly narrow down ZINC15 to a smaller dataset enriched with high scoring compounds, which consists of novel molecules with highly favourable docking scores as well as significantly different than known protease inhibitors.

Collectively, our results strongly support the use of docking beyond libraries of few millions compounds. In a recent article, Lyu et al.[28] have showed that such strategy leads to identifying new scaffolds as well as chemicals of unprecedented potency, that cannot be retrieved from small chemical libraries (i.e. few millions of molecules). Likewise, our DD screening identified 585 new scaffolds for SARS-COV-2 which are not shared with known protease inhibitors, although they can establish all the critical interactions with the protease active site, thus providing a completely new set of chemicals for testing and optimization.

CONCLUSIONS

The use of DD methodology in conjunction with Glide allowed rapid docking of 1.3 chemical structures into an active site of novel SARS-COV-2 Mpro. The candidate inhibitors in the top-1,000 hit list are chemically diverse, exhibit superior docking scores compared to known protease inhibitors, and can be readily sourced from established vendors. The structures of the identified compounds are made publicly available and should facilitate international efforts in rapid development of suitable drug candidates against COVID-19.

ASSOCIATED CONTENT

Docking results for all molecules discussed in the text can be obtained from the authors. The following files are available free of charge.

SARS main protease inhibitors (SM_1.pdf)

Top 1,000 ZINC15 compounds identified as potential inhibitors of the active site of SARS-COV-2 main protease (SM_2.xlsx)

AUTHOR INFORMATION

CORRESPONDING AUTHOR

*E-mail: acherkasov@prostatecentre.com

AUTHORS CONTRIBUTIONS

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. ‡A.T.T. and F.G. contributed equally.

NOTES

The authors declare no competing financial interest.

ABBREVIATIONS

CoVs, coronaviruses; SARS, severe acute respiratory syndrome; MERS, middle east respiratory syndrome; ACE2, Angiotensin-converting enzyme II; RdRp, RNA-dependent RNA polymerase; Mpro, main protease; DD, Deep Docking; QSAR, quantitative structure-activity relationship; PDB, Protein Data Bank; LE, ligand efficiency

REFERENCES

(1) Zumla, A.; Chan, J. F. W.; Azhar, E. I.; Hui, D. S. C.; Yuen, K.-Y. Coronaviruses — Drug Discovery and Therapeutic Options. *Nat Rev Drug Discov* **2016**, *15* (5), 327–347. https://doi.org/10.1038/nrd.2015.37.

(2) de Wit, E.; van Doremalen, N.; Falzarano, D.; Munster, V. J. SARS and MERS: Recent Insights into Emerging Coronaviruses. *Nat Rev Microbiol* **2016**, *14* (8), 523–534. https://doi.org/10.1038/nrmicro.2016.81.

(3) Song, Z.; Xu, Y.; Bao, L.; Zhang, L.; Yu, P.; Qu, Y.; Zhu, H.; Zhao, W.; Han, Y.; Qin, C. From SARS to MERS, Thrusting Coronaviruses into the Spotlight. *Viruses* **2019**, *11* (1). https://doi.org/10.3390/v11010059.

(4) Hui, D. S.; I Azhar, E.; Madani, T. A.; Ntoumi, F.; Kock, R.; Dar, O.; Ippolito, G.; Mchugh, T. D.; Memish, Z. A.; Drosten, C.; et al. The Continuing 2019-NCoV Epidemic Threat of Novel Coronaviruses to Global Health — The Latest 2019 Novel Coronavirus Outbreak in Wuhan, China. *International Journal of Infectious Diseases* **2020**, *91*, 264–266. https://doi.org/10.1016/j.ijid.2020.01.009.

(5) Coronavirus Latest: Chinese Cases Spike after Changes to Diagnosis Method. *Nature* **2020**, d41586-020-00154-w. https://doi.org/10.1038/d41586-020-00154-w.

(6) Wang, M.; Cao, R.; Zhang, L.; Yang, X.; Liu, J.; Xu, M.; Shi, Z.; Hu, Z.; Zhong, W.; Xiao, G. Remdesivir and Chloroquine Effectively Inhibit the Recently Emerged Novel Coronavirus (2019-NCoV) in Vitro. *Cell Res* **2020**. https://doi.org/10.1038/s41422-020-0282-0.

(7) Li, G.; De Clercq, E. Therapeutic Options for the 2019 Novel Coronavirus (2019-NCoV). *Nat Rev Drug Discov* **2020**, d41573-020-00016–0. https://doi.org/10.1038/d41573-020-00016-0.

(8) Han, D. P.; Penn-Nicholson, A.; Cho, M. W. Identification of Critical Determinants on ACE2 for SARS-CoV Entry and Development of a Potent Entry Inhibitor. *Virology* **2006**, *350* (1), 15–25. https://doi.org/10.1016/j.virol.2006.01.029.

(9) Cameron, C. E.; Castro, C. The Mechanism of Action of Ribavirin: Lethal Mutagenesis of RNA Virus Genomes Mediated by the Viral RNA-Dependent RNA Polymerase: *Current Opinion in Infectious Diseases* **2001**, *14* (6), 757–764. https://doi.org/10.1097/00001432-200112000-00015.

(10) Chu, C. M. Role of Lopinavir/Ritonavir in the Treatment of SARS: Initial Virological and Clinical Findings. *Thorax* **2004**, *59* (3), 252–256. https://doi.org/10.1136/thorax.2003.012658.

(11) Lu, I.-L.; Mahindroo, N.; Liang, P.-H.; Peng, Y.-H.; Kuo, C.-J.; Tsai, K.-C.; Hsieh, H.-P.; Chao, Y.-S.; Wu, S.-Y. Structure-Based Drug Design and Structural Biology Study of Novel Nonpeptide Inhibitors of Severe Acute Respiratory Syndrome Coronavirus Main Protease. *J. Med. Chem.* **2006**, *49* (17), 5154–5161. https://doi.org/10.1021/jm060207o.

(12) Blanchard, J. E.; Elowe, N. H.; Huitema, C.; Fortin, P. D.; Cechetto, J. D.; Eltis, L. D.; Brown, E. D. High-Throughput Screening Identifies Inhibitors of the SARS Coronavirus Main Proteinase. *Chemistry & Biology* **2004**, *11* (10), 1445–1453. https://doi.org/10.1016/j.chembiol.2004.08.011.

(13) Liu, X.; Zhang, B.; Jin, Z.; Yang, H.; Rao, Z. The Crytal Structure of 2019-NCoV Main Protease in Complex with an Inhibitor N3. *PDB* **2020**. https://doi.org/10.2210/pdb6lu7/pdb.

(14) Paasche, A.; Zipper, A.; Schäfer, S.; Ziebuhr, J.; Schirmeister, T.; Engels, B. Evidence for Substrate Binding-Induced Zwitterion Formation in the Catalytic Cys-His Dyad of the SARS-CoV Main Protease. *Biochemistry* **2014**, *53* (37), 5930–5946. https://doi.org/10.1021/bi400604t.

(15) Lee, H.; Mittal, A.; Patel, K.; Gatuz, J. L.; Truong, L.; Torres, J.; Mulhearn, D. C.; Johnson, M. E. Identification of Novel Drug Scaffolds for Inhibition of SARS-CoV 3-Chymotrypsin-like Protease Using Virtual and High-Throughput Screenings. *Bioorg. Med. Chem.* **2014**, *22* (1), 167–177. https://doi.org/10.1016/j.bmc.2013.11.041.

(16) Ghosh, A. K.; Xi, K.; Johnson, M. E.; Baker, S. C.; Mesecar, A. D. Progress in Anti-SARS Coronavirus Chemistry, Biology and Chemotherapy. In *Annual Reports in Medicinal Chemistry*; Elsevier, 2006; Vol. 41, pp 183–196. https://doi.org/10.1016/S0065-7743(06)41011-3.

(17) Tuley, A.; Fast, W. The Taxonomy of Covalent Inhibitors. *Biochemistry* **2018**, *57* (24), 3326–3337. https://doi.org/10.1021/acs.biochem.8b00315.

(18) Turk, B. Targeting Proteases: Successes, Failures and Future Prospects. *Nat Rev Drug Discov* **2006**, *5* (9), 785–799. https://doi.org/10.1038/nrd2092.

(19) Ghosh, A. K.; Gong, G.; Grum-Tokars, V.; Mulhearn, D. C.; Baker, S. C.; Coughlin, M.; Prabhakar, B. S.; Sleeman, K.; Johnson, M. E.; Mesecar, A. D. Design, Synthesis and Antiviral Efficacy of a Series of Potent Chloropyridyl Ester-Derived SARS-CoV 3CLpro Inhibitors. *Bioorg. Med. Chem. Lett.* **2008**, *18* (20), 5684–5688. https://doi.org/10.1016/j.bmcl.2008.08.082.

(20) Pillaiyar, T.; Manickam, M.; Namasivayam, V.; Hayashi, Y.; Jung, S.-H. An Overview of Severe Acute Respiratory Syndrome–Coronavirus (SARS-CoV) 3CL Protease Inhibitors:

Peptidomimetics and Small Molecule Chemotherapy. *J. Med. Chem.* **2016**, *59* (14), 6595–6628. https://doi.org/10.1021/acs.jmedchem.5b01461.

(21) Li, Y.; Zhang, J.; Wang, N.; Li, H.; Shi, Y.; Guo, G.; Liu, K.; Zeng, H.; Zou, Q. Therapeutic Drugs Targeting 2019-NCoV Main Protease by High-Throughput Screening. *bioRxiv* **2020**, 2020.01.28.922922. https://doi.org/10.1101/2020.01.28.922922.

(22) Xu, Z.; Peng, C.; Shi, Y.; Zhu, Z.; Mu, K.; Wang, X.; Zhu, W. Nelfinavir Was Predicted to Be a Potential Inhibitor of 2019-NCov Main Protease by an Integrative Approach Combining Homology Modelling, Molecular Docking and Binding Free Energy Calculation. *bioRxiv* **2020**, 2020.01.27.921627. https://doi.org/10.1101/2020.01.27.921627.

(23) Liu, X.; Wang, X.-J. Potential Inhibitors for 2019-NCoV Coronavirus M Protease from Clinically Approved Medicines. *bioRxiv* **2020**, 2020.01.29.924100. https://doi.org/10.1101/2020.01.29.924100.

(24) Zhavoronkov, A.; Aladinskiy, V.; Zhebrak, A.; Zagribelnyy, B.; Terentiev, V.; Bezrukov, D. S.; Polykovskiy, D.; Shayakhmetov, R.; Filimonov, A.; Orekhov, P.; et al. Potential 2019-NCoV 3C-like Protease Inhibitors Designed Using Generative Deep Learning Approaches. **2020**. https://doi.org/10.26434/CHEMRXIV.11829102.V1.

(25) Zhang, H.; Saravanan, K. M.; Yang, Y.; Hossain, Md. T.; Li, J.; Ren, X.; Wei, Y. *Deep Learning Based Drug Screening for Novel Coronavirus 2019-NCov*; preprint; other, 2020. https://doi.org/10.20944/preprints202002.0061.v1.

(26) Abuhammad, A.; Al-Aqtash, R. A.; Anson, B. J.; Mesecar, A. D.; Taha, M. O. Computational Modeling of the Bat HKU4 Coronavirus 3CL ^{pro} Inhibitors as a Tool for the Development of Antivirals against the Emerging Middle East Respiratory Syndrome

(MERS) Coronavirus. *J Mol Recognit* **2017**, *30* (11), e2644. https://doi.org/10.1002/jmr.2644.

(27) Berry, M.; Fielding, B.; Gamieldien, J. Potential Broad Spectrum Inhibitors of the Coronavirus 3CLpro: A Virtual Screening and Structure-Based Drug Design Study. *Viruses* **2015**, *7* (12), 6642–6660. https://doi.org/10.3390/v7122963.

(28) Lyu, J.; Wang, S.; Balius, T. E.; Singh, I.; Levit, A.; Moroz, Y. S.; O'Meara, M. J.; Che, T.; Algaa, E.; Tolmachova, K.; et al. Ultra-Large Library Docking for Discovering New Chemotypes. *Nature* **2019**, *566* (7743), 224–229. https://doi.org/10.1038/s41586-019-0917-9.

(29) Gentile, F.; Agrawal, V.; Hsing, M.; Ban, F.; Norinder, U.; Gleave, M. E.; Cherkasov, A. Deep Docking - a Deep Learning Approach for Virtual Screening of Big Chemical Datasets. *bioRxiv* **2019**, 2019.12.15.877316. https://doi.org/10.1101/2019.12.15.877316.

(30) Sterling, T.; Irwin, J. J. ZINC 15 - Ligand Discovery for Everyone. *Journal of Chemical Information and Modeling* **2015**, *55* (11), 2324–2337. https://doi.org/10.1021/acs.jcim.5b00559.

(31) QUACPAC 2.0.2.2. OpenEye Scientific Software: Santa Fe, NM, USA 2019.

(32) Hawkins, P. C. D.; Skillman, A. G.; Warren, G. L.; Ellingson, B. A.; Stahl, M. T. Conformer Generation with OMEGA: Algorithm and Validation Using High Quality Structures from the Protein Databank and Cambridge Structural Database. *Journal of Chemical Information and Modeling* **2010**, *50* (4), 572–584. https://doi.org/10.1021/ci100031x.

(33) Liu, X.; Zhang, B.; Jin, Z.; Yang, H.; Rao, Z. The Crytal Structure of 2019-NCoV Main Protease in Complex with an Inhibitor N3. *RCSB Protein Data Bank* **2020**. https://doi.org/10.2210/PDB6LU7/PDB.

(34) Berman, H. M. The Protein Data Bank. *Nucleic Acids Research* **2000**, *28* (1), 235–242. https://doi.org/10.1093/nar/28.1.235.

(35) Schrödinger LLC. Small-Molecule Drug Discovery Suite 2019-1. New York, NY, USA 2019.

(36) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; et al. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *Journal of medicinal chemistry* **2004**, *47* (7), 1739–1749. https://doi.org/10.1021/jm0306430.

(37) Liu, T.; Lin, Y.; Wen, X.; Jorissen, R. N.; Gilson, M. K. BindingDB: A Web-Accessible Database of Experimentally Determined Protein-Ligand Binding Affinities. *Nucleic Acids Research* **2007**, *35* (Database), D198–D201. https://doi.org/10.1093/nar/gkl999.

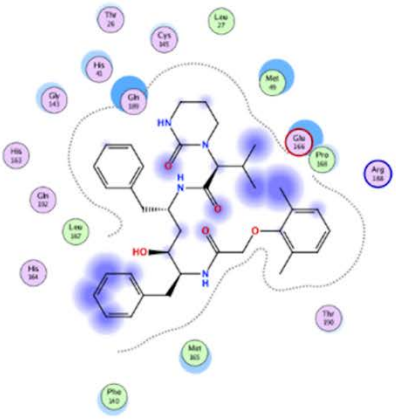
(38) Huang, C.; Wang, Y.; Li, X.; Ren, L.; Zhao, J.; Hu, Y.; Zhang, L.; Fan, G.; Xu, J.; Gu, X.; et al. Clinical Features of Patients Infected with 2019 Novel Coronavirus in Wuhan, China. *The Lancet* **2020**, *395* (10223), 497–506. https://doi.org/10.1016/S0140-6736(20)30183-5.

(39) Chen, L.-R.; Wang, Y.-C.; Lin, Y. W.; Chou, S.-Y.; Chen, S.-F.; Liu, L. T.; Wu, Y.-T.; Kuo, C.-J.; Chen, T. S.-S.; Juang, S.-H. Synthesis and Evaluation of Isatin Derivatives as Effective SARS Coronavirus 3CL Protease Inhibitors. *Bioorganic & Medicinal Chemistry Letters* **2005**, *15* (12), 3058–3062. https://doi.org/10.1016/j.bmcl.2005.04.027.

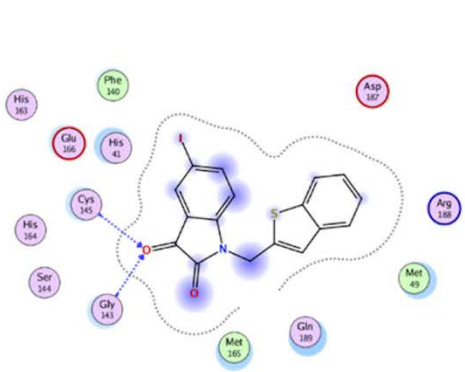
(40) Liu, W.; Zhu, H.-M.; Niu, G.-J.; Shi, E.-Z.; Chen, J.; Sun, B.; Chen, W.-Q.; Zhou, H.-G.; Yang, C. Synthesis, Modification and Docking Studies of 5-Sulfonyl Isatin Derivatives as SARS-CoV 3C-like Protease Inhibitors. *Bioorganic & Medicinal Chemistry* **2014**, *22* (1), 292–302. https://doi.org/10.1016/j.bmc.2013.11.028.

(41) Bemis, G. W.; Murcko, M. A. The Properties of Known Drugs. 1. Molecular Frameworks. *J. Med. Chem.* **1996**, *39* (15), 2887–2893. https://doi.org/10.1021/jm9602928.

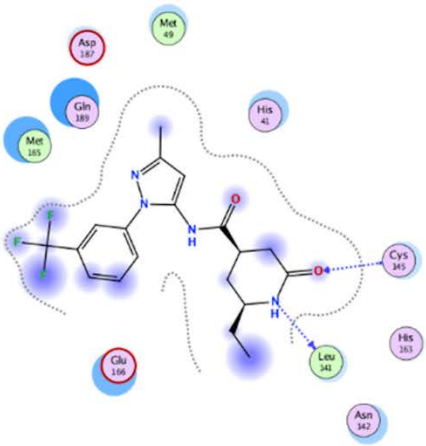
FIGURES AND TABLES



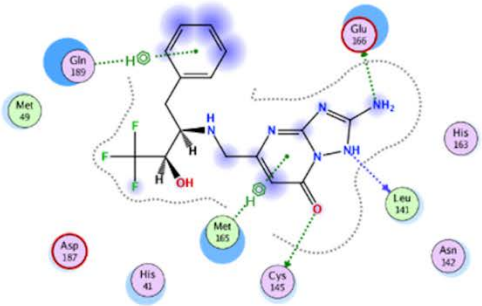
Lopinavir (GS=-7.12 kcal/mol)



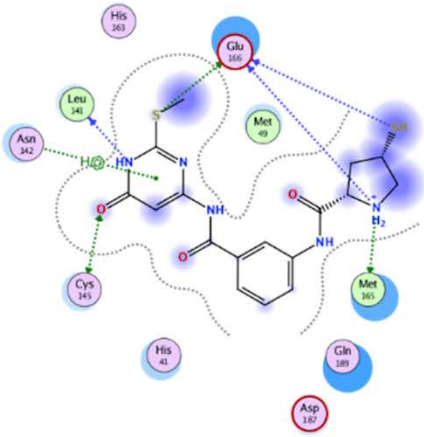
Compound 80 (GS=-6.01 kcal/mol)



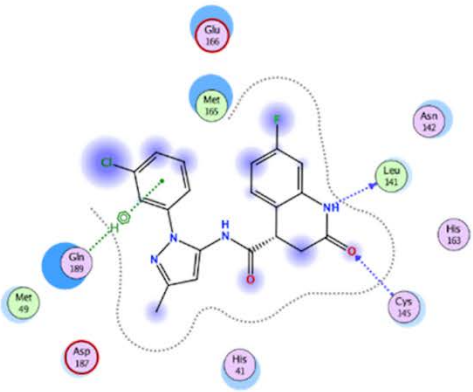
ZINC000541677852 (GS=-11.32 kcal/mol)



ZINC001655436520 (GS=-11.15 kcal/mol)



ZINC001218583693 (GS=-11.04 kcal/mol)



ZINC000636416501 (GS=-10.85 kcal/mol)

Figure 1. *Interaction diagrams of two protease inhibitors, lopinavir and compound 80, and the top four compounds identified by DD screening. Common interactions between these four compounds and the SARS-COV-2 Mpro binding site are the two hydrogen-bonding interactions with Cys145 and Leu141. These molecules also showed similar interaction patterns of two protease inhibitors docked at the same site, and significantly better Glide scores (GS).*

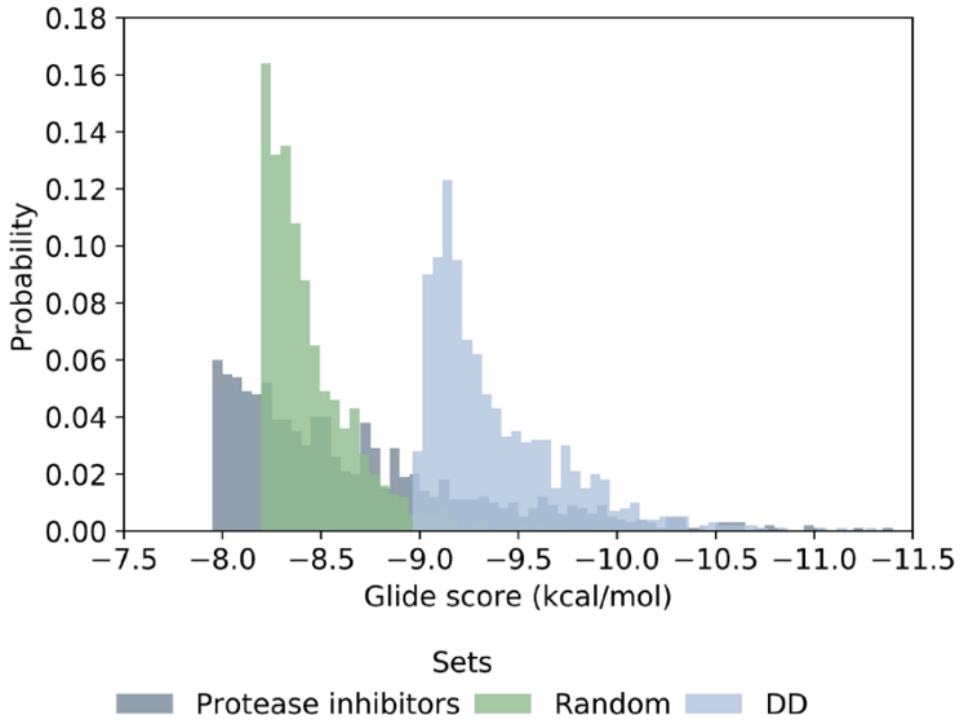


Figure 2. *Score probability of top 1,000 ranked compounds extracted from docking of a set of protease inhibitors (7,800 compounds), a random sample of ZINC15 (1 million molecules) and top 1 million molecules from DD.*