Structural insights into the binding modes of viral RNA-dependent RNA polymerases using a function-site interaction fingerprint method for RNA virus drug discovery

Zheng Zhao and Philip E. Bourne*

School of Data Science and Department of Biomedical Engineering, University of Virginia,

Charlottesville, Virginia 22904, United States of America

*Corresponding author

Email: peb6a@virginia.edu

Abstract

The COVID-19 pandemic speaks to the need for drugs that are not only effective but also remain so given the mutation rate of COVID-19. To this end, we describe a strategy to design potential drugs that target RNA-dependent RNA polymerase (RDRP), a common conserved component of RNA viruses. We combine an RDRP structure dataset and all RDRP-ligand interaction fingerprints into an RDRP-targeted drug discovery procedure. In so doing we reveal the ligand-binding modes and RDRP structural characteristics. Specifically, four types of binding modes with corresponding binding pockets were determined, suggesting two major potential sub-pockets available for drug discovery. We screened a drug dataset of approximately 8,000 compounds against these binding pockets and presented the top ten small molecules as a starting point in further exploring the prevention of virus replication. In summary, the binding characteristics determined here help rationalize RDRP targeted drug discovery and provide insights into the specific binding mechanisms.

1. Introduction

As of June 1, 2020, the COVID-19 pandemic has caused a severe threat to global public health and infected over 6.3 million people. Consequently, researchers have focused on developing new testing techniques, vaccines, and drug design and repurposing to mitigate coronavirus SARS-CoV-2[1-2]. However, to date, there are no effective COVID-19-specific therapeutic agents being prescribed. Laboratory testing techniques have made breakthroughs - the FDA has recently granted an emergency use authorization for the Sofia 2 SARS Antigen FIA Covid-19 test[3], which can detect the virus within minutes. Progress in detection is important but it does not speak to a treatment, hence the research community is paying more attention to finding a SARS-CoV-2 vaccine and/or drug to mitigate the pandemic[4].

Drug discovery is thwarted by the multiple mutations found in the COVID-19 family[5]. Three distinct "variants" from the COVID-19 genomes sampled[5] between December 24th 2019 and March 4th 2020 have been reported. Thus, it is challenging to design novel COVID-19 medications, which are not only effective, but remain so given the mutation rate.

Scientists have established COVID-19 as an RNA virus containing a single-stranded positive-sense RNA genome[1]. RNA viruses have been the main cause of epidemics in the last two decades - SARS[6] in 2003, MERS[7] in 2012, Ebola[8] in 2014, and Zika[9-10] in 2015 and now COVID-19. RNA viruses are divided into 4 classes[2]: single positive-strand RNA ((+)ssRNA) such as SARS, MERS, and COVID-19; single negative-strand RNA ((−)ssRNA) such as Ebola; double-strand RNA (dsRNA); and retroviruses such as HIV[11-12]. These viruses replicate their genetic material within host cells[13], hence one way to limit infection is to inhibit virus replication. Apart from retroviruses, the other classes all contain a common component, RNA-dependent RNA polymerase (RDRP)[14] which catalyzes the replication of viral RNA and hence is a prime drug target. Multiple high-

resolution RDRP 3D structures have been solved, each with a similar core architecture. Namely, a "cupped right hand" with 7 motifs (A-G) comprising "palm," "fingers," and "thumb" (Figure 1; Table S1)[15-16]. Here we propose a drug discovery scheme targeting the conserved RDRP.

On April 29th 2020, NIH indicated the repurposed drug Remdesivir, targeting RDRP, shortened patients' time to recovery by 4 days, or 31%. However, Remdesivir did not show significant efficacy in reducing mortality[17] but it is a start in the quest for RDRP targeted drugs. Here we advance anti-COVID-19 drug R&D by revealing new features of the binding characteristics to RDRP using a computational pharmacology approach.
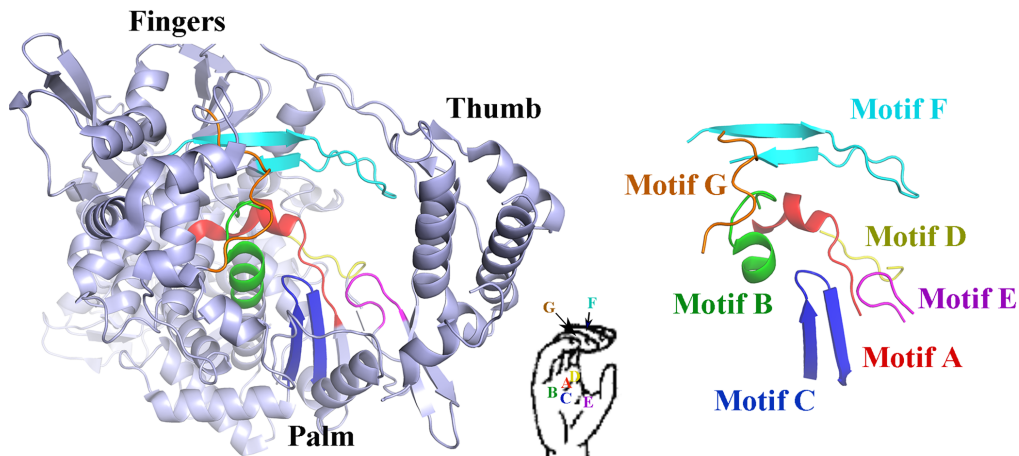

Figure 1. Conserved motifs in the RDRP binding site (PDB id: 7BV2).

We collected 375 PDB structures of RDRP catalytic domains and their complexes (as of Feb. 26th 2020) from 46 RNA viruses including coronavirus as our RDRP dataset. Then, using computational pharmacology methods, including protein-ligand interaction fingerprints, we characterized the RDRP-ligand interactions to provide new insights into antiviral drug design and discovery. Finally, combining a virtual docking process with an antiviral compound library from

Drugbank (www.drugbank.ca), we determined specific potential inhibitors as drug repurposing opportunities, as well as gaining new insights into possible modes of inhibition.

2. Results

2.1. RDRP global structure and COVID-19 - RDRP binding site similarity

The 375 released RDRP PDB structures from 46 different RNA viruses belong to 3 classes: (-)ssRNA (8), dsRNA (9), and (+)ssRNA viruses (29) as shown on the RDRP dendrogram (Figure 2a, and Table S2). Of the 375, 141 have ligands bound. The global structural similarity between all RDRP catalytic domains and COVID-19 RDRP was calculated (see Methods) (Figure 2b). The lowest similarity (see Methods), 0.33, is from bacteria, Escherichia coli (UniProt P0A6P1). The top three viruses with RDRP similarity above 0.65 are Poliovirus type 1 (UniProt P03300), Hepatitis C virus genotype 2a (UniProt Q99IB8), and Hepatitis C virus genotype 1b (UniProt P26663). In sum, structurally, the RDRP catalytic domains have high global structural similarity to COVID-19. Not surprisingly, with such high global structure similarity, the binding sites are highly similar (Figure 2c). Thus, the keys differences lie in the subtle details.
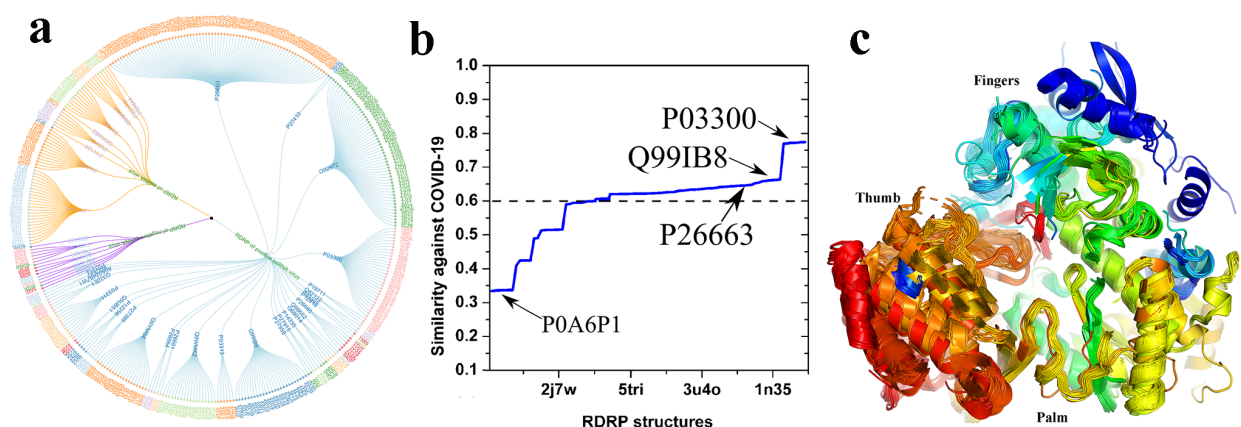

5

Figure 2. (a) RDRP structure dataset. The tree branches are color-coded by RNA virus categories. Each leaf indicates a PDB structure, and the PDB names are color-coded by different viruses. (b) Similarity of catalytic domains between all RDRPs and COVID-19 RDRP. (c) Alignment of all binding sites within the RDRP dataset.

2.2. RDRP ligand-binding characteristics

Using the COVID-19 RDRP binding site as a template, 141 binding sites were aligned (see Methods). Removing the columns of amino acids without encoded interaction information, each aligned binding site consists of 123 columns of amino acids – the interaction fingerprint (Table S3). According to the similarity of interaction fingerprints among all complexes it was possible to divide the binding modes into four classes, where each class contains multiple PDB structures from different kinds of viruses (Table 1). Each class possesses distinct binding characteristics (Figure 3).

Table 1. Clustered binding modes and the corresponding structural ids and UniProt entries.

| Classes	PDB ids	UniProt entries
ClassI	5F3T, 5F3Z, 5F41, 5HMX, 5HMY, 5HMZ, 5HN0, 5I3P, 5I3Q, 5JJR, 5JJS	Q6YMS4
ClassII	1GX6, 1HI0, 1HI1, 1N1H, 1N35, 1N38, 1NB6, 1RA7, 1S49, 1UON, 1UVN, 2ILY, 2ILZ, 2IM0, 2IM1, 2IM2, 2IM3, 2J7U, 2J7W, 2JLG, 2R7X, 2XI3, 3AVT, 3AVW, 3AVX, 3AVY, 3VNU, 3VNV, 3VWS, 4A8F, 4A8K, 4FWT, 4HDG, 4HDH, 4RY5, 4WTA, 4WTC, 4WTD, 4WTE, 4WTF, 4WTG, 4WTI, 4WTJ, 4WTK, 4WTL, 4WTM, 5IQ6, 5UJ2, 7BV2	O37061, O92972, P03300, P0A6P1, P0CK31, P11124, P19711, P26663, P27395, P27915, P27958, Q6YMS4, Q99IB8, P0DTD1
ClassIII	3FQK, 3FQL, 4KAI, 4KB7, 4KBI, 4KE5, 5PZK, 5PZM, 5PZN, 5PZO, 5PZP, 5QJ0, 5QJ1, 5TWM, 5TWN, 5W2E, 6MVP	P26663, Q99IB8, Q9WMX2
ClassIV	2AX1, 2GIQ, 2JC0, 2JC1, 2YOJ, 3BR9, 3BSA, 3BSC, 3CDE, 3CO9, 3CSO, 3CVK, 3CWJ, 3D28, 3D5M, 3E51, 3G86, 3GNV, 3GNW, 3GOL, 3GYN, 3H2L, 3H59, 3H5S, 3H5U, 3H98, 3HKW, 3HKY, 3IGV, 3LKH, 3QGF, 3QGG, 3QGH, 3QGI, 3VQS, 3SKA, 3SKE, 3SKH, 3TYQ, 3TYV, 3U4O, 3U4R, 3UPH, 3UPI, 4EAW, 4IH5, 4IH6, 4IH7, 4IZ0, 4JY0, 4MIA, 4MIB, 4MK7, 4MK8, 4MK9, 4MKA, 4MKB, 4MZ4, 5PZL, 5TRH, 5TRI, 5TRJ, 5TRK	O92972, P26663, P26664, Q9WMX2
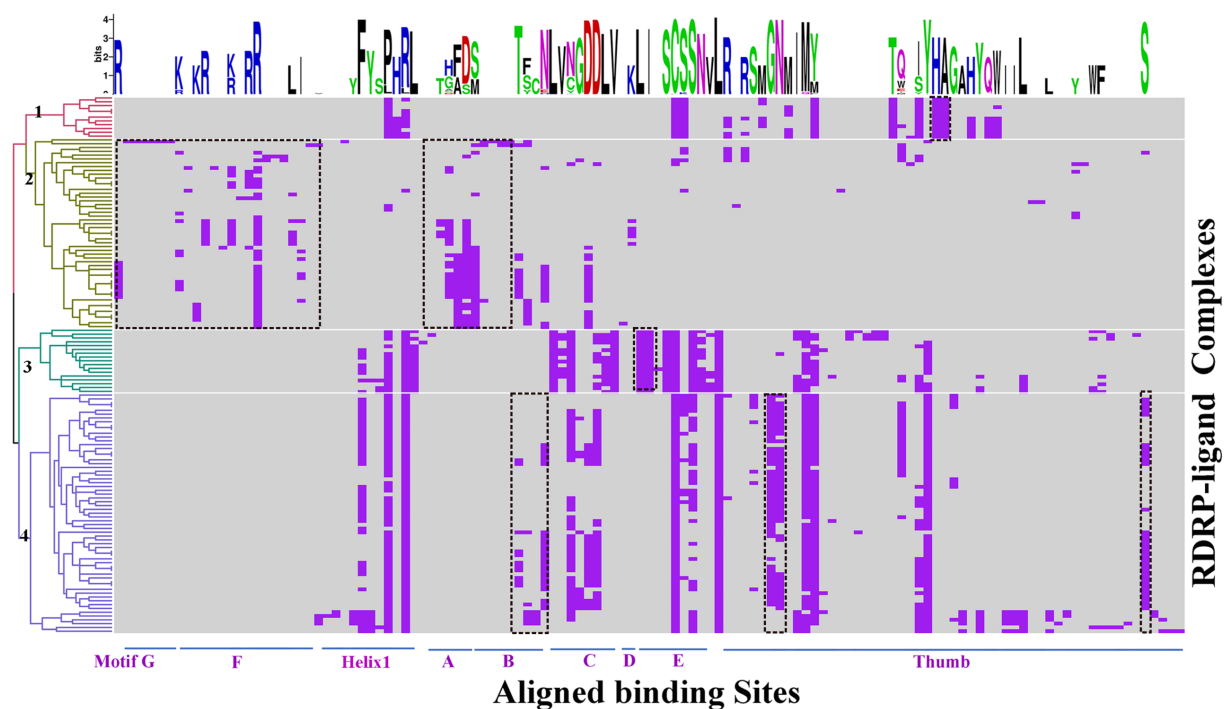
Figure 3. Aligned binding-site-ligand matrix with the clustering of interaction fingerprints and the sequence conservation logo of interaction-involved amino acids. Each row represents the interaction fingerprint of one complex. Each column represents the interaction fingerprint contributed by the amino acid in the same spatial position across all binding sites. The purple area indicates that an interaction exists with the bound ligand in the corresponding RDRP complex and the corresponding column, and grey indicates no interaction. The dashed rectangles approximately delineate the unique characteristics of each class.

Class I. There are 11 aligned PDB structures belonging to the same Dengue virus, (Uniprot entry: Q6YMS4) in this class (Table 1). The aligned binding sites have almost the same binding patterns (Figure 3). Within Motifs A-D and Motifs F-G there are no interaction fingerprints, however, conserved interaction fingerprints exist in Motif E, Helix1, and the Thumb domain, which implies the ligand binding site is located at the palm region and between Motif E and the Thumb domain

(Figure 4a). The residues L511, H512, and L514 of Helix1; C709 and S710 of Motif E provide the conserved interactions (Figure 4a). In the Thumb domain the interaction fingerprints of all complexes in the class are similar, especially in the columns marked with the dashed rectangle, (Figure 3). The role of this Class I binding pocket has been discussed by other groups previously[18]. Noble et al. inhibited enzyme activity through fragment screening[18] identifying this binding pocket. As part of their study, by changing a phenyl to a thiophene, higher binding affinity was obtained, highlighting the role of this pocket in subsequent drug design.

Class II. There are 50 PDB structures from 15 viruses (Table 1) in this class (Figure 3). Class II interaction fingerprints exist mainly in the region of Motifs A-D and Motifs F-G implying the ligand is located at the regions of "palm" and "fingers" (Figure 4b). Remdesivir is reported to bind in this subpocket[16] where K551 and R553 located within Motif F, D623 located within Motif A, S682 located within Motif B, and D760 located within Motif C are the major contributors to ligand binding (PDB id 7bv2). While these amino acids are conserved, Remdesivir only provides moderate improvement in the recovery time of patients with severe symptoms of COVID-19[19]. Further exploring this binding site with compounds of higher binding affinity would seem warranted.

Class III. There are 17 PDB structures belonging to three RNA viruses (Table 1) in this class. The interaction fingerprints are distributed in the regions Helix1, Motif C, Motif E, and the Thumb (Figure 3) which form a binding pocket to accommodate the ligand (Figure 4c). Specifically, in Helix1 the three residues, P197, R200, and L204, provide the primary interactions with the ligand and are conserved in the class (Figure 3). Motif C is a beta-hairpin folding (Figure 1) and on each strand there are 3 conserved amino acids (residue 314-316 and 319-321) contributing to forming the binding pocket (Figure 3 and 4c). Within Motif E, L360, I363, S365, C366, and S368 provide

the main binding interactions (Figure 4c). Two conserved residues (L360 and I363) define a unique fingerprint for the class. Compared with the Class I binding pocket, both are composed of Helix1, Motif E, and the Thumb domain. However, the difference is that in Class III, Motif C is involved as well, thus the binding pockets partially overlap each other. In a previous report[20], Mayland et al. discovered an inhibitor GSK-5852, which is just to target the Class III of pocket in HCV RDRP, to treat HCV infection[20].

Class IV. This is the largest class with 63 PDB structures belonging to 4 viruses (Table 1) and has interaction fingerprints most similar to Class I and Class III. Specifically, using an HCV complex (PDB id 3cwj) as the representative (Figure 4d), in the region of Helix1, F193, P197, and R200 interact with the ligand. Residues D318 and D319 from Motif C and residue C366 from Motif E are also conserved as was found in Class III. Distinct from Classes I and III, residues from Motif B participate in the binding interactions, notably N291. Another difference occurs in Motif E, only residue C366 from the hairpin loop interacts with the ligand, different from Class III which involves additional residues. Interestingly, within Motif E, the column where C366 is located is high conserved to provide the interaction as shown in the conserved logo (Figure 3). In the Thumb domain there are interactions not found in the other classes. Thus, in class IV the pocket is composed of the Thumb domain, Motifs B,C and E (Figure 4d).

In summary, according to our clustering, there are four distinct binding modes in the conserved core architecture of RDRP, each with different sub-pockets to accommodate diverse inhibitors. Classes I, III, and IV have Helix1 and Motif E always participating in ligand-binding interactions hence their binding pockets have a common overlap. Class II has a different sub-pocket which has been exploited as a primary target[9] to fight COVID-19[9].
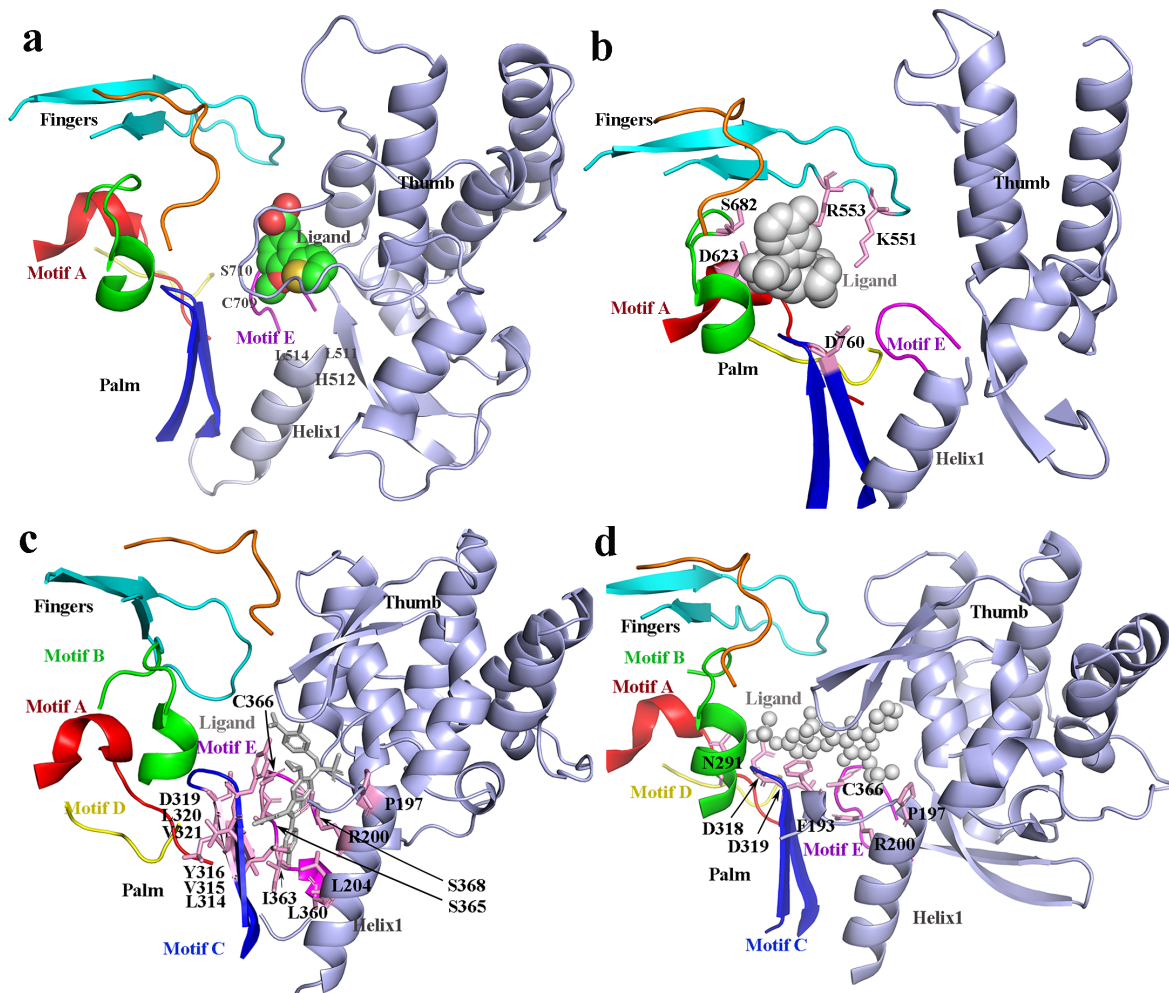
Figure 4. (a-d) classed binding patterns for the corresponding ClassI-IV (PDB ids: 5f3z, 7bv2, 4ke5, and 3cwj, respectively). Note: Helix1 refers to the helix located upstream of Motif A in all RDRP structures.

2.3. COVID-2019 RDRP-targeted drug screening

With the above-mentioned binding classes in mind, we screened 7894 FDA-approved small-molecule drugs targeting the RDRP catalytic domain. Two different sub-pockets were chosen as the binding pockets. one sub-pocket (sub-pocket 1) is located within Class II and the other (sub-pocket 2) is located in the area centered on the common region of Class I, III, and IV (See the two

sub-pockets highlighted with spheres[21] in Figure S1). Through virtual screening (see Methods), the top five highest scoring compounds against each sub-pocket are listed in Table 2. For sub-pocket 1, there is an inhibitor of Factor Xa (Darexaban)[22] which prevents venous thromboembolism by acting as an anticoagulant and antithrombotic after surgery, two inhibitors of histone deacetylase (4SC-202 and CUDC-907)[23-24], an inhibitor of dipeptidyl peptidase 4 (DB07779)[25], and an inhibitor of EGFR (Osimertinib)[26]. These sub-pocket 1 inhibitors interact with Motif A-D and F-G (Figure 5a). For comparison, Remdesivir (accession number: DB14761) is included in our compound lib and its docking score is 6.0 (Figure 4b), considerably less than our top scoring inhibitors.

Table 2. The top five putative inhibitors for sub-pockets 1 and 2 with the docking score and their corresponding primary targets.

Pockets	Name	Docking Score	Primary Target
Sub-pocket 1	Darexaban	9.8	Factor Xa
	4SC-202	9.4	Histone deacetylases
	DB07779	9.1	Dipeptidyl peptidase 4
	Osimertinib	9.0	Epidermal growth factor receptor (EGFR)
	CUDC-907	8.9	Phosphoinositide 3-kinase and Histone deacetylases
Sub-pocket 2	DB07005	8.4	Thrombospondin receptor
	LY-517717	8.2	Coagulation factor X
	Pentamidine	8.0	DNA and tRNA (cytosine(38)-C(5)) methyltransferase
	DB07074	8.0	Coagulation factor XI
	Nafamostat	7.9	Serine protease

Screening of sub-pocket 2 revealed five inhibitors with binding affinity > 7.9 (Figure 5b). It is noteworthy that two of the inhibitors (LY-517717 and DB07074)[27] target coagulation factors X and XI, respectively. As blood thinners these drugs might have the added value of reducing blood

clotting which has been reported in COVID-19 patients[28]. Pentamidine is an agent to treat pneumocystis pneumonia in HIV-infected patients[29]. Nafamostat is a short-acting anticoagulant which acts as a serine protease inhibitor and reported to have antiviral properties[30]. As such it is undergoing a clinical trial in Japan.

To summarize, we characterized RDRP binding pockets suggesting four classes of binding modes (Class I-IV), in silico screened against two completely different pockets (sub-pocket 1 and 2), and obtained a series of putative inhibitors with high binding-affinity. Further experimental validation is necessary.
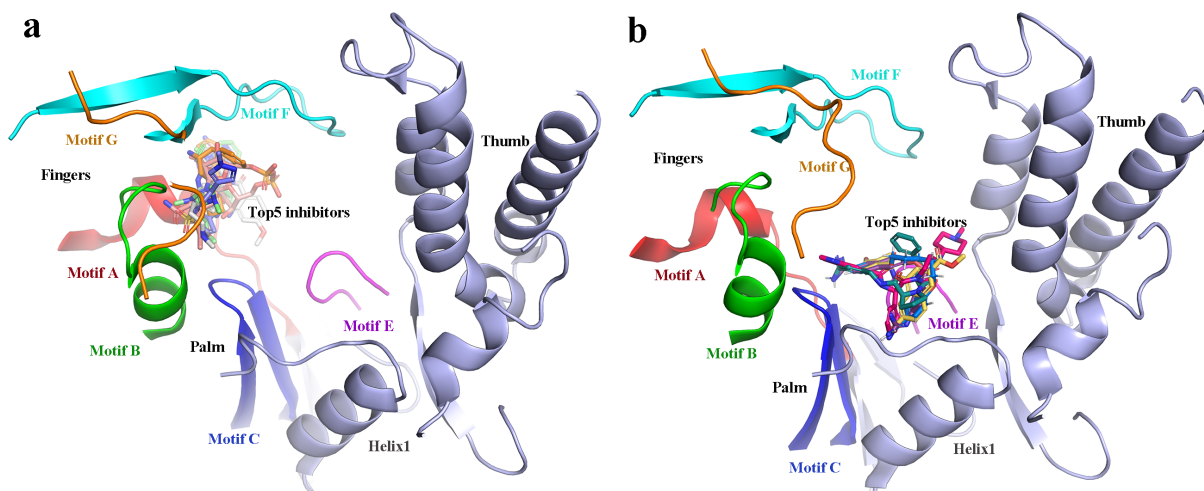

Figure 5. The putative inhibitors obtained via virtual screening (PDB id: 7bv2): (a) Targeting sub-pocket 1; (b) Targeting sub-pocket 2.

3. Conclusion

In the paper we explored the structural characteristics of the RDRP catalytic domain using a computational pharmacology method. More specifically, we focused on the ligand-binding characteristics of the RDRP binding site using a receptor-ligand interaction fingerprint strategy. We collected all RDRP structures and analyzed the conserved core structure. Across the entire

dataset a "cupped right hand" folding pattern and 7 conserved motifs characterize a high-similarity RDRP architecture. By analyzing these protein-ligand complexes with an overall shared architecture four different classes of binding modes were validated. Class II is based on the pocket consisting of Motifs A-D and F-G, whereas Classes I, III, and IV have distinct yet some overlapping characteristics, for example Helix1 and Motif E always participate in ligand binding. In terms of distinct characteristics, Class I has unique a unique binding mode in the Thumb domain, Class III in Motif E, and Class IV in Motifs B and E. Based on these RDRP-ligand binding features, multiple FDA drugs were screened to determine possible repurposing opportunities. The top 10 inhibitors against the two most distinct sub-pockets are discussed. One is already part of a clinical trial as a potential COVID-19 drug. Results here provide further potential repurposing opportunities that need experimental analysis.

4. Method

4.1. Structural RDRP dataset

We first counted the PDB ids of all RDRP structures by accessing the ProRule accession number of RDRP (PRU00539) using the PROSITE[31] database and then downloaded all corresponding PDB structures[25]. As of Feb. 26 2020, 375 RDRP (EC 2.7.7.48) catalytic domain structures were deposited in the PDB (Table S2). Among them were 141 ligand-bound complexes used here.

4.2. Encoding function-site interaction fingerprints (Fs-IFPs)

Fs-IFP is to describe protein-ligand interaction characteristics at the functional site and to do so on a proteome-wide scale as detailed in previous applications[32-35]. Generally, the Fs-IFP method follows three steps. First step is to align all of binding sites. Here, the secondary structures of all RDRP catalytic domains were aligned against the COVID-19 RDRP structure template using the

sequence-independent structural alignment program TM-align with the default scoring function. TM-align results have a value between 0 and 1. >0.3 implies similar fold and > 0.5 implies the same fold[36]. The alignment of the binding sites was performed using SMAP with default parameters[37-38]. The COVID-19 RDRP-remdesivir complex (PDB id: 7bv2) was used as the template and residues within 15 Angstroms of the ligand defined the binding site[39]. Second step is to encode the Fs-IFPs. Here, the interaction fingerprints are encoded using a previously described interaction fingerprint method[40]. Third, along with the aligned binding sites, the comparable interaction fingerprints of each complex were obtained and clustered using the k-means method in the R package.

4.3. High-throughput screening

7894 DrugBank-annotated drug molecules were downloaded from Drugbank as our compound library[41]. These drugs were docked to the RDRP catalytic domain using the docking software Surflex[21] 4.103. Using the distilled binding characteristics, multiple proto molecules were first obtained by predicting the binding pocket and protomol processing[21]. Then multiple proto molecules that bind different sub-pockets were chosen to screen the potential compounds using default parameters. All docked small molecules with different binding conformations were sorted based on the binding affinity score. The top five highest scoring molecules from different sub-pockets were further analyzed.

5. Supporting information

Conserved motifs A-G and their corresponding amino acid sequences (Table S1); The complete RDRP dataset (Table S2); The aligned ligand binding sites (Table S3).

6. Acknowledgement

7. Reference

1.   Zhou, P.; Yang, X. 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., A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **2020,** *579* (7798), 270-273.

2.   Liu, C.; Zhou, Q.; Li, Y.; Garner, L. V.; Watkins, S. P.; Carter, L. J.; Smoot, J.; Gregg, A. C.; Daniels, A. D.; Jervey, S.; Albaiu, D., Research and Development on Therapeutic Agents and Vaccines for COVID-19 and Related Human Coronavirus Diseases. *ACS Cent Sci* **2020,** *6* (3), 315-331.

3.   Sofia 2 SARS Antigen FIA. https://www.fda.gov/media/137886/download.

4.   Amanat, F.; Krammer, F., SARS-CoV-2 Vaccines: Status Report. *Immunity* **2020,** *52* (4), 583-589.

5.   Forster, P.; Forster, L.; Renfrew, C.; Forster, M., Phylogenetic network analysis of SARS-CoV-2 genomes. *Proc Natl Acad Sci U S A* **2020,** *117* (17), 9241-9243.

6.   Tsang, K.; Zhong, N. S., SARS: pharmacotherapy. *Respirology* **2003,** *8 Suppl*, S25-30.

7.   Chafekar, A.; Fielding, B. C., MERS-CoV: Understanding the Latest Human Coronavirus Threat. *Viruses* **2018,** *10* (2).

8.   Subissi, L.; Keita, M.; Mesfin, S.; Rezza, G.; Diallo, B.; Van Gucht, S.; Musa, E. O.; Yoti, Z.; Keita, S.; Djingarey, M. H.; Diallo, A. B.; Fall, I. S., Ebola Virus Transmission Caused by Persistently Infected Survivors of the 2014-2016 Outbreak in West Africa. *J Infect Dis* **2018,** *218* (suppl_5), S287-S291.

9.   Baud, D.; Gubler, D. J.; Schaub, B.; Lanteri, M. C.; Musso, D., An update on Zika virus infection. *The Lancet* **2017,** *390* (10107), 2099-2109.

10.   Zhao, Z.; Martin, C.; Fan, R.; Bourne, P. E.; Xie, L., Drug repurposing to target Ebola virus replication and virulence using structural systems pharmacology. *BMC bioinformatics* **2016,** *17*, 90.

11.   Koonin, E. V., The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. *J Gen Virol* **1991,** *72 (Pt 9)*, 2197-206.

12.   Li, C. X.; Shi, M.; Tian, J. H.; Lin, X. D.; Kang, Y. J.; Chen, L. J.; Qin, X. C.; Xu, J.; Holmes, E. C.; Zhang, Y. Z., Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. *Elife* **2015,** *4*.

13.   Plotch, S. J.; O'Hara, B.; Morin, J.; Palant, O.; LaRocque, J.; Bloom, J. D.; Lang, S. A.; DiGrandi, M. J.; Bradley, M.; Nilakantan, R.; Gluzman, Y., Inhibition of Influenza A Virus Replication by Compounds Interfering with the Fusogenic Function of the Viral Hemagglutinin. *Journal of Virology* **1999,** *73* (1), 140-151.

14.   Shu, B.; Gong, P., The uncoupling of catalysis and translocation in the viral RNA-dependent RNA polymerase. *RNA Biol* **2017,** *14* (10), 1314-1319.

15.   Venkataraman, S.; Prasad, B.; Selvarajan, R., RNA Dependent RNA Polymerases: Insights from Structure, Function and Evolution. *Viruses* **2018,** *10* (2).

16.   Yin, W.; Mao, C.; Luan, X.; Shen, D. D.; Shen, Q.; Su, H.; Wang, X.; Zhou, F.; Zhao, W.; Gao, M.; Chang, S.; Xie, Y. C.; Tian, G.; Jiang, H. W.; Tao, S. C.; Shen, J.; Jiang, Y.; Jiang, H.; Xu, Y.; Zhang, S.; Zhang, Y.; Xu, H. E., Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. *Science* **2020**.

17.   NIH clinical trial shows Remdesivir accelerates recovery from advanced COVID-19. https://www.nih.gov/news-events/news-releases/nih-clinical-trial-shows-remdesivir-accelerates-recovery-advanced-covid-19.

18.   Noble, C. G.; Lim, S. P.; Arora, R.; Yokokawa, F.; Nilar, S.; Seh, C. C.; Wright, S. K.; Benson, T. E.; Smith, P. W.; Shi, P. Y., A Conserved Pocket in the Dengue Virus Polymerase Identified through Fragment-based Screening. *J Biol Chem* **2016**, *291* (16), 8541-8.

19.   Beigel, J. H.; Tomashek, K. M.; Dodd, L. E.; Mehta, A. K.; Zingman, B. S.; Kalil, A. C.; Hohmann, E.; Chu, H. Y.; Luetkemeyer, A.; Kline, S.; Lopez de Castilla, D.; Finberg, R. W.; Dierberg, K.; Tapson, V.; Hsieh, L.; Patterson, T. F.; Paredes, R.; Sweeney, D. A.; Short, W. R.; Touloumi, G.; Lye, D. C.; Ohmagari, N.; Oh, M. D.; Ruiz-Palacios, G. M.; Benfield, T.; Fatkenheuer, G.; Kortepeter, M. G.; Atmar, R. L.; Creech, C. B.; Lundgren, J.; Babiker, A. G.; Pett, S.; Neaton, J. D.; Burgess, T. H.; Bonnett, T.; Green, M.; Makowski, M.; Osinusi, A.; Nayak, S.; Lane, H. C.; Members, A.-S. G., Remdesivir for the Treatment of Covid-19 - Preliminary Report. *N Engl J Med* **2020**.

20.   Maynard, A.; Crosby, R. M.; Ellis, B.; Hamatake, R.; Hong, Z.; Johns, B. A.; Kahler, K. M.; Koble, C.; Leivers, A.; Leivers, M. R.; Mathis, A.; Peat, A. J.; Pouliot, J. J.; Roberts, C. D.; Samano, V.; Schmidt, R. M.; Smith, G. K.; Spaltenstein, A.; Stewart, E. L.; Thommes, P.; Turner, E. M.; Voitenleitner, C.; Walker, J. T.; Waitt, G.; Weatherhead, J.; Weaver, K.; Williams, S.; Wright, L.; Xiong, Z. Z.; Haigh, D.; Shotwell, J. B., Discovery of a potent boronic

acid derived inhibitor of the HCV RNA-dependent RNA polymerase. *J Med Chem* **2014,** *57* (5), 1902-13.

21.   Jain, A. N., Surflex-Dock 2.1: robust performance from ligand energetic modeling, ring flexibility, and knowledge-based search. *J Comput Aided Mol Des* **2007,** *21* (5), 281-306.

22.   Eriksson, B. I.; Turpie, A. G.; Lassen, M. R.; Prins, M. H.; Agnelli, G.; Kalebo, P.; Gaillard, M. L.; Meems, L.; group, O. s., A dose escalation study of YM150, an oral direct factor Xa inhibitor, in the prevention of venous thromboembolism in elective primary hip replacement surgery. *J Thromb Haemost* **2007,** *5* (8), 1660-5.

23.   Wobser, M.; Weber, A.; Glunz, A.; Tauch, S.; Seitz, K.; Butelmann, T.; Hesbacher, S.; Goebeler, M.; Bartz, R.; Kohlhof, H.; Schrama, D.; Houben, R., Elucidating the mechanism of action of domatinostat (4SC-202) in cutaneous T cell lymphoma cells. *J Hematol Oncol* **2019,** *12* (1), 30.

24.   Oki, Y.; Kelly, K. R.; Flinn, I.; Patel, M. R.; Gharavi, R.; Ma, A.; Parker, J.; Hafeez, A.; Tuck, D.; Younes, A., CUDC-907 in relapsed/refractory diffuse large B-cell lymphoma, including patients with MYC-alterations: results from an expanded phase I trial. *Haematologica* **2017,** *102* (11), 1923-1930.

25.   Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E., The Protein Data Bank. *Nucleic Acids Res* **2000,** *28* (1), 235-42.

26.   Tan, C.-S.; Gilligan, D.; Pacey, S., Treatment approaches for EGFR-inhibitor-resistant patients with non-small-cell lung cancer. *The Lancet Oncology* **2015,** *16* (9), e447-e459.

27.   Agnelli, G.; Haas, S.; Ginsberg, J. S.; Krueger, K. A.; Dmitrienko, A.; Brandt, J. T., A phase II study of the oral factor Xa inhibitor LY517717 for the prevention of venous thromboembolism after hip or knee replacement. *J Thromb Haemost* **2007,** *5* (4), 746-53.

28.   Paranjpe, I.; Fuster, V.; Lala, A.; Russak, A.; Glicksberg, B. S.; Levin, M. A.; Charney, A. W.; Narula, J.; Fayad, Z. A.; Bagiella, E.; Zhao, S.; Nadkarni, G. N., Association of Treatment Dose Anticoagulation with In-Hospital Survival Among Hospitalized Patients with COVID-19. *J Am Coll Cardiol* **2020**.

29.   Anwar, A.; Khan, N. A.; Siddiqui, R., Repurposing of Drugs Is a Viable Approach to Develop Therapeutic Strategies against Central Nervous System Related Pathogenic Amoebae. *ACS Chem Neurosci* **2020**.

30.   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,** *30* (3), 269-271.

31.   Sigrist, C. J.; de Castro, E.; Cerutti, L.; Cuche, B. A.; Hulo, N.; Bridge, A.; Bougueleret, L.; Xenarios, I., New and continuing developments at PROSITE. *Nucleic Acids Res* **2013,** *41* (Database issue), D344-7.

32.   Zhao, Z.; Xie, L.; Bourne, P. E., Insights into the binding mode of MEK type-III inhibitors. A step towards discovering and designing allosteric kinase inhibitors across the human kinome. *PLoS One* **2017,** *12* (6), e0179936.

33.   Zhao, Z.; Xie, L.; Bourne, P. E., Structural Insights into Characterizing Binding Sites in Epidermal Growth Factor Receptor Kinase Mutants. *J. Chem. Inf. Model.* **2019,** *59* (1), 453-462.

34.   Zhao, Z.; Xie, L.; Xie, L.; Bourne, P. E., Delineation of Polypharmacology across the Human Structural Kinome Using a Functional Site Interaction Fingerprint Approach. *Journal of medicinal chemistry* **2016,** *59* (9), 4326-4341.

35.   Zhao, Z.; Bourne, P. E., Revealing Acquired Resistance Mechanisms of Kinase-Targeted Drugs Using an on-the-Fly, Function-Site Interaction Fingerprint Approach. *J Chem Theory Comput* **2020,** *16* (5), 3152-3161.

36.   Zhang, Y.; Skolnick, J., TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res* **2005,** *33* (7), 2302-9.

37.   Xie, L.; Bourne, P. E., A robust and efficient algorithm for the shape description of protein structures and its application in predicting ligand binding sites. *BMC bioinformatics* **2007,** *8 Suppl 4*, S9.

38.   Xie, L.; Bourne, P. E., Detecting evolutionary relationships across existing fold space, using sequence order-independent profile-profile alignments. *Proceedings of the National Academy of Sciences of the United States of America* **2008,** *105* (14), 5441-6.

39.   Kellenberger, E.; Muller, P.; Schalon, C.; Bret, G.; Foata, N.; Rognan, D., sc-PDB: an annotated database of druggable binding sites from the Protein Data Bank. *J Chem Inf Model* **2006,** *46* (2), 717-27.

40.   Da Silva, F.; Desaphy, J.; Rognan, D., IChem: A Versatile Toolkit for Detecting, Comparing, and Predicting Protein-Ligand Interactions. *ChemMedChem* **2018,** *13* (6), 507-510.

41.   Wishart, D. S.; Knox, C.; Guo, A. C.; Cheng, D.; Shrivastava, S.; Tzur, D.; Gautam, B.; Hassanali, M., DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res* **2008,** *36* (Database issue), D901-6.