

Virus-CKB 2.0: Viral-Associated Disease-Specific Chemogenomics Knowledgebase

Yixuan Hao,[#] Maozi Chen,[#] Yasmin Othman, Xiang-Qun Xie,* and Zhiwei Feng*



ABSTRACT: Transmissible and infectious viruses can cause large-scale epidemics around the world. This is because the virus can constantly mutate and produce different variants and subvariants to counter existing treatments. Therefore, a variety of treatments are urgently needed to keep up with the mutation of the viruses. To facilitate the research of such treatment, we updated our Virus-CKB 1.0 to Virus-CKB 2.0, which contains 10 kinds of viruses, including enterovirus, dengue virus, hepatitis C virus, Zika virus, herpes simplex virus, *Andes orthohantavirus*, human immunodeficiency virus, Ebola virus, Lassa virus, influenza virus, coronavirus, and norovirus. To date, Virus-CKB 2.0 archived at least 65 antiviral drugs (such as remdesivir, telaprevir, acyclovir, boceprevir, and nelfinavir) in the market, 178 viral-related targets with 292 available 3D crystal or cryo-EM structures, and 3766 chemical agents reported for these target proteins. Virus-CKB 2.0 is integrated with established tools for target prediction and result visualization; these include HTDocking, TargetHunter, blood-brain barrier (BBB) predictor, Spider Plot, etc. The Virus-CKB 2.0 server is accessible at https://www.cbligand.org/g/virus-ckb. By using the established chemogenomic tools and algorithms and newly developed tools, we can screen FDA-approved drugs and chemical compounds that may bind to these proteins involved in viral-associated disease regulation. If the virus strain mutates and the vaccine loses its effect, we can still screen drugs that can be used to treat the mutated virus in a fleeting time. In some cases, we can even repurpose FDA-approved drugs through Virus-CKB 2.0.

■ INTRODUCTION

The goal of antiviral treatment is the eradication of the virus. Direct-acting antiviral agents play a significant role in antiviral therapy such as anti-HIV and anti-HCV therapy.¹ However, viral resistance proves to be a major obstacle preventing direct-acting antiviral drugs from eliminating the virus, because once disease strains mutate and acquire viral resistance, direct-acting antiviral drugs may render ineffective.² Thus, it is very necessary to investigate the life cycle of viruses and track the structure of different viruses continuously to meet the research purpose.

Viruses have different protein domains related to the invasion or replication of the viruses, so drugs targeting these different protein domains may provide therapeutic potential for the treatment of diseases caused by these viruses. SARS-CoV-2, the causative agent of COVID-19, has four structural proteins: the spike protein, envelope protein, membrane protein, and nucleocapsid protein. The spike protein plays a

significant role in the attachment of viruses during the process of virus infection.³ The SARS-CoV-2 genome encodes several polyproteins that can be processed by viral proteases and form nonstructural proteins. These nonstructural proteins can form a replicase complex that plays a significant role during the transcription and replication of the viral genome.⁴ For example, remdesivir has been shown to inhibit RNA-dependent RNA polymerase (RdRp) composed of nonstructural protein 12 (NS12) with high potency.^{5,6} Galidesivir, a directacting antiviral drug, has been confirmed to treat the Zika virus

Received:July 6, 2022Accepted:September 26, 2022Published:October 10, 2022





infection by inhibiting the nonstructural protein 5 RNAdependent RNA polymerase (RdRp).⁷ Therefore, computational tools can greatly boost the search for new drug molecules and aid the repurposing of the drugs and modification of existing medicines.

To facilitate the COVID-19 research and discovery of new antiviral treatments in general, we have established an integrated viral-associated disease-specific chemogenomics knowledgebase published in July 2020.8 The former Virus-CKB 1.0 has archived a total of 65 antiviral drugs (such as remdesivir, telaprevir, acyclovir, boceprevir, and nelfinavir) in the market and 107 viral-related targets, including 6 HIVrelated targets, 7 BCV/HCV-related targets, 49 influenzarelated targets, 39 coronavirus-related targets, and another 6 unclassified viral-related targets with 189 available 3D crystal or cryo-EM structures. Moreover, 2698 chemical agents reported for these target proteins are also included in Virus-CKB 1.0. As for writing, Virus-CKB 1.0 has been published for more than a year. Many target proteins collected for Virus-CKB 1.0 did not make their way into Virus-CKB 1.0, because when constructing Virus-CKB 1.0, the resolution of the targets available was generally lower than it is now, and the number of targets was much lower than the current targets. In addition, many target information, 3D structures, and related active compounds of SARS-CoV-2 have been released during the past 1 year. For example, by far, there are more than 240 3CL protease crystal structures of SARS-CoV-2 and more than 8000 reported compounds that can bind with or inhibit SARS-CoV-2.9 There are more than 250 nonstructural protein crystal structures of SARS-CoV-2 and more than 800 reported compounds that can bind with or inhibit nonstructural protein of SARS-CoV-2.¹⁰ There are more than 170 crystal structures of the spike protein of SARS-CoV-2 and more than 800 reported compounds that can bind with or inhibit the spike protein of SARS-CoV-2.11

Virus-CKB can also be applied to the research of other viruses that can cause viral infectious diseases. As the years go by, information and related structures of many other viruses have also been published, such as dengue virus,^{12,13} Zika virus,^{14,15} Ebola virus,^{16,17} enterovirus,^{18,19} Lassa virus,^{20,21} norovirus,^{22,23} herpes simplex virus,^{24–27} and *Andes orthohantavirus.*²⁸ Therefore, we incorporated these viruses into Virus-CKB 2.0. Table 1 shows a detailed comparison between Virus-CKB 1.0 and Virus-CKB 2.0.

Table 1. Comparison between Virus-CKB 1.0 and Virus-CKB 2.0

features	Virus-CKB 1.0	Virus-CKB 2.0
protein targets	104	178
3D crystal or cryo-EM structures	189	292
chemical agents	2698	3766
docking method	idock	jdock

Currently, there are several viral-related chemogenomic tools or computational algorithms available. For example, ExCAPE-DB is a large-scale chemical structure dataset that can facilitate big data analysis in chemogenomic, and the COVID-19 Docking Server is an interactive server for docking small molecules and peptides against potential targets of COVID-19.^{29,30} However, ExCAPE-DB cannot enable computational prediction of the binding affinity between the small molecules and the viral protein targets; COVID-19 Docking Server only

focuses on the viral targets of COVID-19. Therefore, we constructed Virus-CKB 2.0 to facilitate viral-related research.

MATERIALS AND METHODS

Genes and Domain Structures. The viral genome bears the information about the functional and structural proteins of the virus. Virus proteins undergo post-translational modifications to form different protein domains of the virus with contrasting functions. Direct-acting antiviral agents are drugs that act directly on viral proteins, which are considered the most promising and effective antiviral treatment.³¹ Screening small molecules that can directly act on viral proteins with distinct functions is a common method for antiviral drug discovery. Therefore, we focus on the interaction between small molecules and viral proteins and collect protein structures based on the different domain information of viruses. For example, dengue virus genome encodes three structural proteins (pre-membrane (prM), envelope (E), and the capsid (C)), which constitute the components of the virion, and seven nonstructural proteins (NS1, NS2A/B, NS3, NS4A/B, and NS5), which are involved in viral RNA replication.¹² When we collect domain structures of dengue virus from Protein Data Bank³² (https://www.rcsb.org/), there are more than 24,000 PDB files. First, we divided the PDB files according to the domain information of dengue virus such as pre-membrane (prM), envelope (E), capsid (C), and seven nonstructural proteins: NS1, NS2A/B, NS3, NS4A/B, and NS5. We will select the PDB files with the most protein residues and the highest resolution to display the structure of the protein domain as completely as possible and incorporate them into Virus-CKB 2.0.

We collect data such as chemical molecules, antibodies, genes, and proteins involved in viral-associated disease regulation from several public databases such as Protein Data Bank³² (https://www.rcsb.org/), UniProt³³ (https://www.uniprot.org/), ChEMBL³⁴ (https://www.ebi.ac.uk/chembl/), Therapeutic Target Database^{35–39} (http://db.idrblab.net/ttd/), and recent literature.

Drugs and Chemicals. Using "dengue", "Zika", or other virus name as keywords, we searched the DrugBank⁴⁰ database and ChEMBL³⁴ database and retrieved 3766 antiviral drugs or viral-related chemical agents and incorporated them into Virus-CKB 2.0. Antiviral drugs or chemical agents with an IC₅₀ lower than 1 μ M toward a viral-related target are defined as active chemical molecules, while those chemical agents with an IC₅₀ larger than 10 μ M are defined as inactive chemical compounds. Therefore, we only retrieved chemical substances with an IC₅₀ lower than 10 μ M into Virus-CKB 2.0. Figure S1 shows the distribution of these 3766 chemical agents among the virused

HTDocking. High-throughput docking (HTDocking)⁴¹⁻⁴³ is a web-based high-throughput computing method that can predict the potential interactions between the user-input compounds and target proteins. jdock is the docking engine of HTDocking at the backend. The definition of HTDocking was first proposed by Liu et al. in 2014.⁴¹ AutoDock Vina⁴⁴ was used as the docking engine at the backend of HTDocking at that time. When the Virus-CKB 1.0 was published in 2020, idock was used as the docking engine at the backend of HTDocking.⁸ Compared with AutoDock Vina, idock has faster scoring function and more efficient optimization algorithm. idock achieves a 3.3 speedup in CPU time and an average speedup of 7.5 in elapsed time.⁴⁵ We have updated the docking engine of HTDocking from idock to jdock⁴⁶ in Virus-CKB 2.0.



Figure 1. Spider Plot of viruses archived in the Virus-CKB 2.0.

jdock saves the fixed time overhead of calculating the Random Forest score. The improvement ratio is related to the structural complexity of the docking and the machine's performance. The time saved for each docking run on mainstream computers is about 2 to 20 s.⁴⁶

Workflow of the Virus-CKB 2.0. The Virus-CKB 2.0 server is implemented with several web applications and existing algorithms including high-throughput docking (HTDocking), TargetHunter,⁴⁷ blood-brain barrier (BBB) predictor,^{41,48,49} Spider Plot,⁵⁰ and other third-party software including Open Babel,^{51,52} JSME Molecular Editor,⁵³ jdock, and NGL Viewer,⁵⁴ which enables the prediction of the protein-ligand interaction and the visualization of the results.

Our platform mainly focuses on the 3D crystal or cryo-EM structures of viral proteins and their related compounds. The workflow of the Virus-CKB 2.0 server consists of three major steps. The first step is to input compounds or chemical agents that may have the potential to bind with the protein structures of the viruses. The job creation page of Virus-CKB 2.0 (Figure S2) can be found in the Supporting Information. The user can submit up to five compounds or chemical agents through the job creation page. The user can define the name of the job and the name of the molecule. The user can define the molecule structure by drawing it with JSME Molecular Editor⁵³ or paste any pre-existing molecule structure in MOL, SDF, or SMILES format via the drop-down menu of the blue double-triangle icon on the toolbar. In this way, the user needs to be aware that the pasted code should not contain any leading or trailing spaces. If the user does not want to reveal the job to the public on Virus-CKB 2.0, they can choose to click Private job so as to hide the input compounds and the test results. After submitting the structures of query compounds, our platform will automatically convert the input format from MOL, SDF, or SMILES into PDB and PDBQT formats with the help of Open Babel and conduct further computational prediction.

The second step is to conduct in silico BBB prediction, highthroughput docking (HTDocking), and fingerprint-based similarity search simultaneously by our established algorithms implemented in Virus-CKB 2.0. The in silico BBB prediction can be used to predict whether the query compounds can pass

through the BBB and therefore cause side effects in the central nervous system (CNS).^{41,48,49} In Virus-CKB 2.0, each virus domain includes up to three conformations with the highest resolution and the most diverse structures. The protein structures of different viruses were collected the from Protein Data Bank, most of which bind to ligands to form specialized complexes. Only a small fraction of viral protein structures don't have bound ligands. Then, we removed water molecules, small solvent molecules, and ligands bound to the viral protein structure to extract the viral protein structure. We archived these viral protein structures and the binding pockets of these viral protein structures in Virus-CKB 2.0 and used these structures as protein targets to conduct docking. HTDocking will dock up to five user compounds into binding pockets of up to three different conformations and generate docking scores, respectively. HTDocking can provide up to nine potential docking scores with different binding poses of the same protein pocket and generate the average docking score. By the definition of free energy, the docking score is negative, and the smaller value (greater absolute value) represents the better binding pose.⁴³ The user can compare the docking score of the query ligand with a known ligand that can bind with the particular target and produce a therapeutic effect to assess whether the query ligand is a good ligand for a specific target. TargetHunter is a web-based target prediction tool that can predict the potential biotargets of submitted compounds based on targets associated with its most similar counterpart (Tanimoto Coefficients) algorithm.⁴⁷ The prediction is based on an important principle of medicinal chemistry: structurally similar compounds have similar physicochemical properties and may cause similar biological effects. When the query compound is uploaded, TargetHunter will first generate molecular fingerprints using Tanimoto coefficients for the query compound and compounds in the public database such as the ChEMBL, then calculate the similarity scores based on Tanimoto coefficients, and finally list the most similar ones as the potential biotargets.

The third step is to conduct systems pharmacology target mapping for potential drug repurposing and drug combination. Upon completion of all the in silico computation processes, the



Figure 2. Spider Plot for the analysis of four antiviral drugs, including telaprevir, boceprevir, narlaprevir, and remdesivir. Potentially active binding sites (viral protein targets) for the query compounds are shown as purple discs, and the binding sites that have been validated by a bioassay are shown as green discs. The average docking scores are marked on the line connecting the query compounds to the viral protein targets.

docking scores and similarity scores are automatically analyzed and classified, and the results are thereafter fed to Spider Plot for the visualization of the compound-target interaction network. Spider Plot is an online tool that can visualize the molecule-protein interaction network based on the target classification,⁵⁰ with which users can intuitively see the in silico computational prediction results generated by Virus-CKB 2.0.

RESULTS AND DISCUSSION

Data and Software Availability. The Virus-CKB 2.0 server is accessible at https://www.cbligand.org/g/virus-ckb. The Virus-CKB 2.0 targets data is accessible at https://www. cbligand.org/g/virus-ckb/targets.

The Virus-CKB 2.0 website is compatible with modern web browsers such as Microsoft Edge, Firefox, Chrome, and Safari, provided that JavaScript and cookies are enabled. The latest release version of these web browsers is recommended for better rendering. **Virus-CKB 2.0 Overview.** Currently, we have collected 178 viral-related targets with 292 available 3D crystal or cryo-EM structures, which include 6 HIV-related targets; 5 herpes simplex virus-related targets; 3*Andes orthohantavirus*-related targets; 4 Lassa virus-related targets; 15 dengue virus-related targets; 9 Zika virus-related targets; 11 Ebolavirus-related targets; 9 Zika virus-related targets; 37 coronavirus-related targets; 12 enterovirus-related targets; 37 coronavirus-related targets that include 6 targets from SARS, 8 targets from SARS-CoV-2, and 23 targets from other coronaviruses; and 49 influenza virus-related targets, as shown in Figure S3. Figure 1 shows all the viruses archived in Virus-CKB 2.0.

Job Output. Once the job is created, the platform will display it on the job listing page. To view the job details, one can click on the listing item. Then, four buttons (JOB, BBB, OUTPUT, and SPIDER PLOT) will be added in the menu bar between the ALL JOBS button and the TARGETS button. Clicking on the BBB button will bring the user to the Blood-

Article



Figure 3. Spider Plot for the analysis of the drug modification, including remdesivir and remdesivir triphosphate. The prediction results showed that remdesivir and remdesivir triphosphate can bind with the RNA-dependent RNA polymerases of SARS-CoV-2.

Brain Barrier Predictor page (Figure S4) where the BBB predictions are computed with two algorithms (AdaBoost and SVM) and a total of eight algorithm–fingerprint combinations for each query compound will be displayed.⁴¹ Clicking on the OUTPUT button will bring the user to the output page (Figure S5), which displays the structure of viral protein targets and their binding pockets, the submitted ligand name, the viral protein target name, the docking scores calculated by HTDocking between the input compounds and the protein target Conformations, the similarity scores calculated by TargetHunter between the input compounds and the most similar active or inactive ChEMBL compound, and finally the

best match compound's ChEMBL ID. Clicking on the ChEMBL ID will bring the user to the comparison page (Figure S6) which provides a detailed comparison between the structural information of the input compound and the most similar compound. The two blocks of numbers at the bottom of the comparison page are the hexadecimal representation of the molecular fingerprints (FP2) computed in our platform to be used in the search of the most similar compound. The numbers marked in red indicate the fingerprint differences between the two compounds, so the more numbers marked in red, the lower the structural similarity between the two

Article

compounds. The final similarity score is computed with Tanimoto, which is simply the quotation of the number of identical numbers (in binary) shared in two fingerprints and the number of nonzero numbers (in binary) appeared in either fingerprint. The output page has two display modes: graphic-based mode and text-based mode. The user can switch the display mode by clicking on the \blacksquare button in the upper right corner. For a specific job, clicking on the SPIDER PLOT will bring the user to the Spider Plot page (Figure 2), which displays the predicted interactions network between the input compounds and the protein targets in Virus-CKB 2.0. Both the text-based output page and the Spider Plot page have a "Gene/ Protein" switch button in the upper right corner. Clicking on this button will switch between gene name and protein name displayed on the output page and Spider Plot page.

As we can see from the Spider Plot generated by our platform (Figure 2), we submitted four antiviral drugs in one job, including telaprevir, boceprevir, narlaprevir, and remdesivir. The structure and name of the query compounds defined by the user are placed in a rectangular box. The discs around the query compound represent the predicted viral protein targets with which the query compound has the potential to bind. The number on the dotted line connecting the predicted viral protein targets and the query compound represents its average docking score. The user can also toggle the display of any docking score on a connecting line. If the query compound has been verified by a bioassay to bind with the predicted viral protein target and produce corresponding therapeutic effects, or the query compound has a highly similar compound (similarity score ≥ 0.95) that has been verified by a bioassay to do so, the predicted viral protein target will be marked as green discs. Otherwise, the predicted targets will be shown as purple discs. For example, our results indicated that the antiviral drug telaprevir can bind to nonstructural protein 3 protease (discs: NS3_HCV1B) of the hepatitis C virus, which is consistent with a bioassay that verifies that telaprevir can be used to treat HCV infection.55

In Silico Drug Modification. When creating a job, the user can modify the submitted compound according to the structure-activity relationship. For example, the user can add functional groups such as benzene ring, hydroxyl, or carbonyl. Our platform will perform in silico prediction on the modified structures, and users can compare the modified structures and modify the compound according to the prediction. For example, remdesivir is a prodrug that has broad-spectrum action against several RNA viruses. Remdesivir can change to the active triphosphate form, which inhibits the RNAdependent RNA polymerases of SARS-CoV-2.56 The platform allows the comparison of the binding affinities of the modified compound derivatives entered by the user. As shown in Figure 3, our results showed that both remdesivir and remdesivir triphosphate can bind with the RNA-dependent RNA polymerases of SARS-CoV-2, which is consistent with a biochemical analysis that confirmed that remdesivir triphosphate can target RNA-dependent RNA polymerases of SARS-CoV-2 and cause the termination of RNA synthesis.⁵

Drug Repurposing for Dengue Virus Infection. To further validate our platform, we predicted several FDA-approved nonviral drugs that can bind with the dengue virus and may have the potential to produce therapeutic effects toward dengue virus infections.

According to statistics, antiviral research on the dengue virus mainly focused on targeting structural and nonstructural proteins. Among them, research on structural proteins has mainly focused on the envelope protein because it plays a significant role in the process of virus entry into cells.⁵⁷ Research on nonstructural proteins has mainly focused on NS3 and NS5, because NS5 is the largest and most conserved nonstructural protein and has viral RNA-dependent RNA polymerase and methyltransferase activity^{58,59} and NS3 has serine protease activity.⁵⁷ We conducted a retrospective analysis (positive control) for the FDA-approved nonviral drug ivermectin, as shown in Figure S7; our results showed that ivermectin was predicted to bind with NS3 of dengue virus (discs: NS3 DEN4T) and NS3 of Zika virus (discs: NS3 ZIKV), which are consistent with a study that conformed that ivermectin is a potent inhibitor of flavivirus replication specifically targeting NS3 helicase activity.⁶⁰ As shown in Figure S8, we also conducted a retrospective analysis (positive control) for policresulen, which is commonly used for the treatment of gynecological infections;⁶¹ our prediction results showed that policresulen was predicted to bind with viral NS3 of dengue virus (discs: NS3 DEN4T), which is consistent with a cell assay that confirmed that policresulen is a NS2B/ NS3 protease inhibitor and can inhibit the replication of dengue virus.⁶² All these findings could provide computational prediction-based guidance for repurposing FDA-approved drugs to deal with sudden outbreaks of viral infections. We also recommend that the results based on computational predictions need to be validated by biological experiments.

Drug Combination for the Treatment of Viral-Related Disease. Combination therapies are becoming increasingly popular in treating viral-related diseases because combination therapies have the advantages of high efficacy and low toxicity and can prevent the development of drug resistance among viruses. In addition to guiding drug repurposing, our platform can also provide computational prediction-based recommendations for drug combinations to treat viral-related diseases. For example, as shown in Figure S9, our retrospective analysis (positive control) results showed that brequinar (anticancer) and sofosbuvir (anti-HCV) can bind with HIV (discs: RT HIV1) and HCV (discs: NS5B HCV1A, NS5B_HCV2A), which are consistent with a cell assay that confirmed that the combination of sofosbuvir and brequinar can produce a significant therapeutic effect toward HCV infection with low toxicity.⁶³ The findings could provide computational prediction-based guidance for combination therapies to improve efficacy, lower toxicity, and deal with drug resistance issues of viral-related diseases. We also recommend that the results based on computational predictions need to be validated by biological experiments.

Antiviral Activity of Aspirin. Aspirin, also known as acetylsalicylic acid (ASA), is well known for its antiinflammatory, pain-relieving, and fever-reducing activities.⁶⁴ The antiviral activities of aspirin have been confirmed by several reports. The main antiviral mechanism of aspirin involves blocking nuclear factor kappa B activation and inhibiting hepatitis C virus entry by downregulating claudin-1 rather than acting directly on viral proteins.^{65,66} Our computational prediction result shown in Figure S10 indicates that aspirin cannot bind directly to any viral protein target in Virus-CKB 2.0, which is consistent with these reports.

CONCLUSIONS

In this work, we developed a viral-associated disease-specific chemogenomics knowledgebase, Virus-CKB 2.0, which archived 65 antiviral drugs in the market, 178 viral-related targets with 292 available 3D crystal or cryo-EM structures, and 3766 chemical agents reported for these target proteins. Virus-CKB 2.0 provides the user with a multifunctional platform, which can conduct computational prediction of direct-acting antiviral effects of submitted compounds. Specifically, our platform will first conduct in silico BBB prediction, high-throughput docking (HTDocking), and fingerprint-based similarity search simultaneously for the submitted compounds with our established algorithms integrated in Virus-CKB 2.0. Then, the docking scores, similarity scores, and all the other computation results will be analyzed and used to generate the visualization of the compound-target interaction network in Spider Plot. Finally, the results can be used to guide drug repurposing and drug combination.

To validate our platform, we applied Virus-CKB 2.0 to predict the direct-acting antiviral effects of the FDA-approved nonviral drugs ivermectin and policresulen. The retrospective analysis (positive control) results showed that ivermectin and policresulen can bind with dengue virus proteins and exhibit significant anti-dengue virus effects, which have been confirmed by relevant bioassays. We also predicted that the drug combination brequinar and sofosbuvir can bind with HIV and HCV and exhibit anti-HIV and anti-HCV effects, which have also been confirmed by relevant experimental data. We also conduct computational prediction for remdesivir and remdesivir triphosphate; the results showed that both remdesivir and remdesivir triphosphate can bind with the RNA-dependent RNA polymerases of SARS-CoV-2, which is consistent with a biochemical analysis that confirmed that remdesivir triphosphate can target RNA-dependent RNA polymerases of SARS-CoV-2. Therefore, our platform can provide meaningful computationally based recommendations for drug modification, drug repurposing, and drug combination therapy.

In the future, we plan to develop more techniques such as protein—protein docking, which can enable the docking of proteins against virus protein targets. We also plan to continue the collection of viral targets and screen in new protein structures, as well as add new navigation and computation options for the domains of protein structures. We have a vision that these will make Virus-CKB 2.0 a particularly useful knowledgebase for the research of viral disease and drug repurposing.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.2c04258.

Figures S1–S10 mentioned in the text (PDF)

AUTHOR INFORMATION

Corresponding Authors

Xiang-Qun Xie – Department of Pharmaceutical Sciences and Computational Chemical Genomics Screening Center, School of Pharmacy; National Center of Excellence for Computational Drug Abuse Research; Drug Discovery Institute; Departments of Computational Biology and Structural Biology, School of Medicine, University of *Pittsburgh, Pittsburgh, Pennsylvania 15261, United States;* orcid.org/0000-0002-6881-6175; Phone: 412-383-5276; Email: xix15@pitt.edu; Fax: 412-383-7436

Zhiwei Feng – Department of Pharmaceutical Sciences and Computational Chemical Genomics Screening Center, School of Pharmacy; National Center of Excellence for Computational Drug Abuse Research; Drug Discovery Institute; Departments of Computational Biology and Structural Biology, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania 15261, United States;
orcid.org/0000-0001-6533-8932; Phone: 412-419-4896; Email: zhf11@pitt.edu

Authors

- Yixuan Hao Department of Pharmaceutical Sciences and Computational Chemical Genomics Screening Center, School of Pharmacy; National Center of Excellence for Computational Drug Abuse Research; Drug Discovery Institute; Departments of Computational Biology and Structural Biology, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania 15261, United States
- Maozi Chen Department of Pharmaceutical Sciences and Computational Chemical Genomics Screening Center, School of Pharmacy; National Center of Excellence for Computational Drug Abuse Research; Drug Discovery Institute; Departments of Computational Biology and Structural Biology, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania 15261, United States
- Yasmin Othman Department of Pharmaceutical Sciences and Computational Chemical Genomics Screening Center, School of Pharmacy; National Center of Excellence for Computational Drug Abuse Research; Drug Discovery Institute; Departments of Computational Biology and Structural Biology, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania 15261, United States

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.2c04258

Author Contributions

[#]These authors contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by NIH NIDA (grant numbers R01DA052329 and P30 DA035778A1).

Notes

The authors declare no competing financial interest.

The Virus-CKB 2.0 server is accessible at https://www. cbligand.org/g/virus-ckb. All the 3D protein structures, ligand structures, and binding pockets can be accessed at https:// www.cbligand.org/g/virus-ckb/targets.

ACKNOWLEDGMENTS

The authors acknowledge the funding support to the Xie laboratory from the NIH NIDA (grant P30 DA035778A1 and grant R01DA052329).

REFERENCES

(1) Patel, S. V.; Jayaweera, D. T.; Althoff, K. N.; Eron, J. J.; Radtchenko, J.; Mills, A.; Moyle, G.; Santiago, S.; Sax, P. E.; Gillman, J.; Mounzer, K.; Elion, R. A.; Huhn, G. D. Real-world efficacy of direct acting antiviral therapies in patients with HIV/HCV. *PLoS One* **2020**, *15*, No. e0228847. (2) Li, D. K.; Chung, R. T., Overview of Direct-Acting Antiviral Drugs and Drug Resistance of Hepatitis C Virus. In *Hepatitis C Virus Protocols;* Law, M., Ed. Springer New York: New York, NY, 2019; pp. 3–32.

(3) Wu, C.; Liu, Y.; Yang, Y.; Zhang, P.; Zhong, W.; Wang, Y.; Wang, Q.; Xu, Y.; Li, M.; Li, X.; Zheng, M.; Chen, L.; Li, H. Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. *Acta Pharm. Sin. B* **2020**, *10*, 766–788.

(4) Douangamath, A.; Fearon, D.; Gehrtz, P.; Krojer, T.; Lukacik, P.; Owen, C. D.; Resnick, E.; Strain-Damerell, C.; Aimon, A.; Ábrányi-Balogh, P.; Brandão-Neto, J.; Carbery, A.; Davison, G.; Dias, A.; Downes, T. D.; Dunnett, L.; Fairhead, M.; Firth, J. D.; Jones, S. P.; Keeley, A.; Keserü, G. M.; Klein, H. F.; Martin, M. P.; Noble, M. E. M.; O'Brien, P.; Powell, A.; Reddi, R. N.; Skyner, R.; Snee, M.; Waring, M. J.; Wild, C.; London, N.; von Delft, F.; Walsh, M. A. Crystallographic and electrophilic fragment screening of the SARS-CoV-2 main protease. *Nat. Commun.* **2020**, *11*, 5047.

(5) Gordon, C. J.; Tchesnokov, E. P.; Woolner, E.; Perry, J. K.; Feng, J. Y.; Porter, D. P.; Götte, M. Remdesivir is a direct-acting antiviral that inhibits RNA-dependent RNA polymerase from severe acute respiratory syndrome coronavirus 2 with high potency. *J. Biol. Chem.* **2020**, *295*, 6785–6797.

(6) Hasan, A.; Paray, B. A.; Hussain, A.; Qadir, F. A.; Attar, F.; Aziz, F. M.; Sharifi, M.; Derakhshankhah, H.; Rasti, B.; Mehrabi, M.; Shahpasand, K.; Saboury, A. A.; Falahati, M. A review on the cleavage priming of the spike protein on coronavirus by angiotensin-converting enzyme-2 and furin. *J. Biomol. Struct. Dyn.* **2021**, *39*, 3025–3033.

(7) Lim, S.-Y.; Osuna, C. E.; Best, K.; Taylor, R.; Chen, E.; Yoon, G.; Kublin, J. L.; Schalk, D.; Schultz-Darken, N.; Capuano, S.; Safronetz, D.; Luo, M.; MacLennan, S.; Mathis, A.; Babu, Y. S.; Sheridan, W. P.; Perelson, A. S.; Whitney, J. B. A direct-acting antiviral drug abrogates viremia in Zika virus—infected rhesus macaques. *Sci. Transl. Med.* **2020**, *12*, No. eaau9135.

(8) Feng, Z.; Chen, M.; Liang, T.; Shen, M.; Chen, H.; Xie, X.-Q. Virus-CKB: an integrated bioinformatics platform and analysis resource for COVID-19 research. *Briefings Bioinf.* **2021**, *22*, 882–895. (9) Liu, Y.; Liang, C.; Xin, L.; Ren, X.; Tian, L.; Ju, X.; Li, H.; Wang, Y.; Zhao, Q.; Liu, H.; Cao, W.; Xie, X.; Zhang, D.; Wang, Y.; Jian, Y. The development of Coronavirus 3C-Like protease (3CLpro) inhibitors from 2010 to 2020. *Eur. J. Med. Chem.* **2020**, *206*, No. 112711.

(10) Littler, D. R.; Gully, B. S.; Colson, R. N.; Rossjohn, J. Crystal Structure of the SARS-CoV-2 Non-structural Protein 9, Nsp9. *iScience* **2020**, *23*, No. 101258.

(11) Sternberg, A.; Naujokat, C. Structural features of coronavirus SARS-CoV-2 spike protein: Targets for vaccination. *Life Sci.* **2020**, 257, No. 118056.

(12) Perera, R.; Kuhn, R. J. Structural proteomics of dengue virus. *Curr. Opin. Microbiol.* **2008**, *11*, 369–377.

(13) Morens, D. M.; Fauci, A. S. Dengue and Hemorrhagic FeverA Potential Threat to Public Health in the United States. *JAMA* 2008, 299, 214–216.

(14) Sirohi, D.; Kuhn, R. J. Zika Virus Structure, Maturation, and Receptors. J. Infect. Dis. 2017, 216, S935-s944.

(15) Sirohi, D.; Chen, Z.; Sun, L.; Klose, T.; Pierson, T. C.; Rossmann, M. G.; Kuhn, R. J. The 3.8 Å resolution cryo-EM structure of Zika virus. *Science* **2016**, 352, 467–470.

(16) Cantoni, D.; Rossman, J. S. Ebolaviruses: New roles for old proteins. *PLoS Negl. Trop. Dis.* **2018**, *12*, No. e0006349.

(17) Wan, W.; Kolesnikova, L.; Clarke, M.; Koehler, A.; Noda, T.; Becker, S.; Briggs, J. A. G. Structure and assembly of the Ebola virus nucleocapsid. *Nature* **2017**, *551*, 394–397.

(18) Yee, P. T. I.; Poh, C. L. Development of Novel Vaccines against Enterovirus-71. *Viruses* **2016**, *8*, 1.

(19) Yi, E.-J.; Shin, Y.-J.; Kim, J.-H.; Kim, T.-G.; Chang, S.-Y. Enterovirus 71 infection and vaccines. *Clin. Exp. Vaccine Res.* **2017**, *6*, 4–14.

(20) Li, S.; Sun, Z.; Pryce, R.; Parsy, M. L.; Fehling, S. K.; Schlie, K.; Siebert, C. A.; Garten, W.; Bowden, T. A.; Strecker, T.; Huiskonen, J. T. Acidic pH-Induced Conformations and LAMP1 Binding of the Lassa Virus Glycoprotein Spike. *PLoS Pathog.* **2016**, *12*, No. e1005418.

(21) Tang-Huau, T. L.; Feldmann, H.; Rosenke, K. Animal models for Lassa virus infection. *Curr. Opin. Virol.* **2019**, *37*, 112–117.

(22) Graziano, V. R.; Wei, J.; Wilen, C. B. Norovirus Attachment and Entry. *Viruses* **2019**, *11*, 495.

(23) de Graaf, M.; van Beek, J.; Koopmans, M. P. Human norovirus transmission and evolution in a changing world. *Nat. Rev. Microbiol.* **2016**, *14*, 421–433.

(24) Enright, A. M.; Prober, C. G. Herpesviridae infections in newborns: varicella zoster virus, herpes simplex virus, and cytomegalovirus. *Pediatr. Clin.* **2004**, *51*, 889–908.

(25) Whitley, R. J.; Roizman, B. Herpes simplex virus infections. *Lancet* **2001**, 357, 1513–1518.

(26) McElwee, M.; Vijayakrishnan, S.; Rixon, F.; Bhella, D. Structure of the herpes simplex virus portal-vertex. *PLoS Biol.* 2018, *16*, No. e2006191.

(27) Copeland, A. M.; Newcomb, W. W.; Brown, J. C. Herpes simplex virus replication: roles of viral proteins and nucleoporins in capsid-nucleus attachment. *J. Virol.* **2009**, *83*, 1660–1668.

(28) Vera-Otarola, J.; Solis, L.; Lowy, F.; Olguín, V.; Angulo, J.; Pino, K.; Tischler, N. D.; Otth, C.; Padula, P.; López-Lastra, M. The Andes orthohantavirus NSs protein antagonizes the type I interferon response by inhibiting MAVS signaling. *J. Virol.* **2020**, *94*, e00454– e00420.

(29) Sun, J.; Jeliazkova, N.; Chupakhin, V.; Golib-Dzib, J.-F.; Engkvist, O.; Carlsson, L.; Wegner, J.; Ceulemans, H.; Georgiev, I.; Jeliazkov, V.; Kochev, N.; Ashby, T. J.; Chen, H. ExCAPE-DB: an integrated large scale dataset facilitating Big Data analysis in chemogenomics. *Aust. J. Chem.* **2017**, *9*, 17.

(30) Kong, R.; Yang, G.; Xue, R.; Liu, M.; Wang, F.; Hu, J.; Guo, X.; Chang, S. COVID-19 Docking Server: an interactive server for docking small molecules, peptides and antibodies against potential targets of COVID-19. *Bioinformatics* **2020**, 5109.

(31) Kim, J.; Park, S.-J.; Park, J.; Shin, H.; Jang, Y.-S.; Woo, J.-S.; Min, D.-H. Interfaces, Identification of a Direct-Acting Antiviral Agent Targeting RNA Helicase via a Graphene Oxide Nanobiosensor. *ACS Appl. Mater. Interfaces* **2021**, *13*, 25715–25726.

(32) Burley, S. K.; Berman, H. M.; Bhikadiya, C.; Bi, C.; Chen, L.; Di Costanzo, L.; Christie, C.; Dalenberg, K.; Duarte, J. M.; Dutta, S.; Feng, Z.; Ghosh, S.; Goodsell, D. S.; Green, R. K.; Guranovic, V.; Guzenko, D.; Hudson, B. P.; Kalro, T.; Liang, Y.; Lowe, R.; Namkoong, H.; Peisach, E.; Periskova, I.; Prlic, A.; Randle, C.; Rose, A.; Rose, P.; Sala, R.; Sekharan, M.; Shao, C.; Tan, L.; Tao, Y. P.; Valasatava, Y.; Voigt, M.; Westbrook, J.; Woo, J.; Yang, H.; Young, J.; Zhuravleva, M.; Zardecki, C. RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic Acids Res.* **2019**, *47*, D464–D474.

(33) Apweiler, R.; Bairoch, A.; Wu, C. H.; Barker, W. C.; Boeckmann, B.; Ferro, S.; Gasteiger, E.; Huang, H.; Lopez, R.; Magrane, M.; Martin, M. J.; Natale, D. A.; O'Donovan, C.; Redaschi, N.; Yeh, L. S. UniProt: the Universal Protein knowledgebase. *Nucleic Acids Res.* **2004**, *32*, D115–D1119.

(34) Gaulton, A.; Bellis, L. J.; Bento, A. P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; Overington, J. P. ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* **2012**, *40*, D1100–D1107.

(35) Zhu, F.; Shi, Z.; Qin, C.; Tao, L.; Liu, X.; Xu, F.; Zhang, L.; Song, Y.; Liu, X.; Zhang, J.; Han, B.; Zhang, P.; Chen, Y. Therapeutic target database update 2012: a resource for facilitating target-oriented drug discovery. *Nucleic Acids Res.* **2012**, *40*, D1128–D1136.

(36) Qin, C.; Zhang, C.; Zhu, F.; Xu, F.; Chen, S. Y.; Zhang, P.; Li, Y. H.; Yang, S. Y.; Wei, Y. Q.; Tao, L.; Chen, Y. Z. Therapeutic target database update 2014: a resource for targeted therapeutics. *Nucleic Acids Res.* **2014**, *42*, D1118–D1123.

(37) Yang, H.; Qin, C.; Li, Y. H.; Tao, L.; Zhou, J.; Yu, C. Y.; Xu, F.; Chen, Z.; Zhu, F.; Chen, Y. Z. Therapeutic target database update 2016: enriched resource for bench to clinical drug target and targeted pathway information. *Nucleic Acids Res.* **2016**, *44*, D1069–D1074.

(38) Li, Y. H.; Yu, C. Y.; Li, X. X.; Zhang, P.; Tang, J.; Yang, Q.; Fu, T.; Zhang, X.; Cui, X.; Tu, G.; Zhang, Y.; Li, S.; Yang, F.; Sun, Q.; Qin, C.; Zeng, X.; Chen, Z.; Chen, Y. Z.; Zhu, F. Therapeutic target database update 2018: enriched resource for facilitating bench-toclinic research of targeted therapeutics. *Nucleic Acids Res.* **2018**, *46*, D1121–d1127.

(39) Wang, Y.; Zhang, S.; Li, F.; Zhou, Y.; Zhang, Y.; Wang, Z.; Zhang, R.; Zhu, J.; Ren, Y.; Tan, Y.; Qin, C.; Li, Y.; Li, X.; Chen, Y.; Zhu, F. Therapeutic target database 2020: enriched resource for facilitating research and early development of targeted therapeutics. *Nucleic Acids Res.* **2019**, *48*, D1031–D1041.

(40) Wishart, D. S.; Feunang, Y. D.; Guo, A. C.; Lo, E. J.; Marcu, A.; Grant, J. R.; Sajed, T.; Johnson, D.; Li, C.; Sayeeda; Assempour; Iynkkaran; Liu; Maciejewski; Gale; Wilson; Chin; Cummings; le; Pon; Knox; Wilson, M. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* **2018**, *46*, D1074– D1082.

(41) Liu, H.; Wang, L.; Lv, M.; Pei, R.; Li, P.; Pei, Z.; Wang, Y.; Su, W.; Xie, X.-Q. AlzPlatform: An Alzheimer's Disease Domain-Specific Chemogenomics Knowledgebase for Polypharmacology and Target Identification Research. *J. Chem. Inf. Model.* **2014**, *54*, 1050–1060.

(42) Cheng, J.; Wang, S.; Lin, W.; Wu, N.; Wang, Y.; Chen, M.; Xie, X.-Q.; Feng, Z. Computational systems pharmacology-target mapping for fentanyl-laced cocaine overdose. *ACS Chem. Neurosci.* **2019**, *10*, 3486–3499.

(43) Chen, M.; Jing, Y.; Wang, L.; Feng, Z.; Xie, X.-Q. DAKB-GPCRs: An Integrated Computational Platform for Drug Abuse Related GPCRs. J. Chem. Inf. Model. 2019, 59, 1283–1289.

(44) Trott, O.; Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461.

(45) Li, H.; Leung, K. S.; Wong, M. H. In idock: A multithreaded virtual screening tool for flexible ligand docking. 2012 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB); IEEE, 9–12 May 2012; 2012; pp. 77–84.

(46) Chen, M.; Feng, Z.; Wang, S.; Lin, W.; Xie, X.-Q. MCCS, a novel characterization method for protein–ligand complex. *Briefings Bioinf.* **2021**, *22*, No. bbaa239.

(47) Wang, L.; Ma, C.; Wipf, P.; Liu, H.; Su, W.; Xie, X.-Q. TargetHunter: An In Silico Target Identification Tool for Predicting Therapeutic Potential of Small Organic Molecules Based on Chemogenomic Database. *AAPS J.* **2013**, *15*, 395–406.

(48) Zhang, Y.; Wang, L.; Feng, Z.; Cheng, H.; McGuire, T. F.; Ding, Y.; Cheng, T.; Gao, Y.; Xie, X.-Q. StemCellCKB: An Integrated Stem Cell-Specific Chemogenomics KnowledgeBase for Target Identification and Systems-Pharmacology Research. J. Chem. Inf. Model. 2016, 56, 1995–2004.

(49) Zhang, H.; Ma, S.; Feng, Z.; Wang, D.; Li, C.; Cao, Y.; Chen, X.; Liu, A.; Zhu, Z.; Zhang, J.; Zhang, G.; Chai, Y.; Wang, L.; Xie, X.-Q. Cardiovascular Disease Chemogenomics Knowledgebase-guided Target Identification and Drug Synergy Mechanism Study of an Herbal Formula. *Sci. Rep.* **2016**, *6*, 33963.

(50) Gareau, T. P.; Smith, R. G.; Barbercheck, M. E.; Mortensen, D. A. J. Spider plots: a tool for participatory extension learning. *J. Extension* **2010**, *48*, No. STOT8.

(51) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An open chemical toolbox. *Aust. J. Chem.* **2011**, *3*, 33.

(52) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open babel. *Aust. J. Chem.* **2013**, *3*, 33–2011.

(53) Ertl, P. Molecular structure input on the web. Aust. J. Chem. 2010, 2, 1.

(54) Rose, A. S.; Bradley, A. R.; Valasatava, Y.; Duarte, J. M.; Prlić, A.; Rose, P. W. Web-based molecular graphics for large complexes. *Bioinformatics* **2018**, *34*, 3755–3758.

(55) Zeuzem, S.; Andreone, P.; Pol, S.; Lawitz, E.; Diago, M.; Roberts, S.; Focaccia, R.; Younossi, Z.; Foster, G. R.; Horban, A.; Ferenci, P.; Nevens, F.; Müllhaupt, B.; Pockros, P.; Terg, R.; Shouval, D.; van Hoek, B.; Weiland, O.; Van Heeswijk, R.; De Meyer, S.; Luo, D.; Boogaerts, G.; Polo, R.; Picchio, G.; Beumont, M. Telaprevir for Retreatment of HCV Infection. *N. Engl. J. Med.* **2011**, *364*, 2417– 2428.

(56) Celik, I.; Erol, M.; Duzgun, Z. J. M. D. In silico evaluation of potential inhibitory activity of remdesivir, favipiravir, ribavirin and galidesivir active forms on SARS-CoV-2 RNA polymerase. *Mol. Diversity* **2022**, *26*, 279–292.

(57) Troost, B.; Smit, J. M. Recent advances in antiviral drug development towards dengue virus. *Curr. Opin. Virol.* **2020**, *43*, 9–21. (58) Kaptein, S. J. F.; Neyts, J. Towards antiviral therapies for treating dengue virus infections. *Curr. Opin. Pharmacol.* **2016**, *30*, 1–7.

(59) Chan, C. Y.; Ooi, E. E. Dengue: an update on treatment options. 2015, *10* (12), 2017–2031, Future Microbiol., DOI: 10.2217/fmb.15.105.

(60) Mastrangelo, E.; Pezzullo, M.; De Burghgraeve, T.; Kaptein, S.; Pastorino, B.; Dallmeier, K.; de Lamballerie, X.; Neyts, J.; Hanson, A. M.; Frick, D. N.; Bolognesi, M.; Milani, M. Ivermectin is a potent inhibitor of flavivirus replication specifically targeting NS3 helicase activity: new prospects for an old drug. *J. Antimicrob. Chemother.* **2012**, *67*, 1884–1894.

(61) Renner, A. Albothyl, a substance with a new mechanism of action in the treatment of gynecological diseases. *Med. Klin.* **1954**, *49*, 1998–1999.

(62) Wu, D.-W.; Mao, F.; Ye, Y.; Li, J.; Xu, C.-L.; Luo, X.-M.; Chen, J.; Shen, X. Policresulen, a novel NS2B/NS3 protease inhibitor, effectively inhibits the replication of DENV2 virus in BHK-21 cells. *Acta Pharmacol. Sin.* **2015**, *36*, 1126–1136.

(63) Ianevski, A.; Yao, R.; Biza, S.; Zusinaite, E.; Mannik, A.; Kivi, G.; Planken, A.; Kurg, K.; Tombak, E.-M.; Ustav, M.; Shtaida, N.; Kulesskiy, E.; Jo, E.; Yang, J.; Lysvand, H.; Løseth, K.; Oksenych, V.; Aas, P. A.; Tenson, T.; Vitkauskienė, A.; Windisch, M. P.; Fenstad, M. H.; Nordbø, S. A.; Ustav, M.; Bjørås, M.; Kainov, D. E. Identification and Tracking of Antiviral Drug Combinations. 2020, *12* (10), 1178, DOI: 10.3390/v12101178.

(64) Morris, T.; Stables, M.; Hobbs, A.; de Souza, P.; Colville-Nash, P.; Warner, T.; Newson, J.; Bellingan, G.; Gilroy, D. W. Effects of Low-Dose Aspirin on Acute Inflammatory Responses in Humans. *J. Immunol* **2009**, *183*, 2089–2096.

(65) Liao, C.-L.; Lin, Y.-L.; Wu, B.-C.; Tsao, C.-H.; Wang, M.-C.; Liu, C.-I.; Huang, Y.-L.; Chen, J.-H.; Wang, J.-P.; Chen, L.-K. Salicylates inhibit flavivirus replication independently of blocking nuclear factor kappa B activation. *J. Virol.* **2001**, *75*, 7828–7839.

(66) Yin, P.; Zhang, L. Aspirin inhibits hepatitis C virus entry by downregulating claudin-1. *J Viral Hepatitis* **2016**, *23*, 62–64.