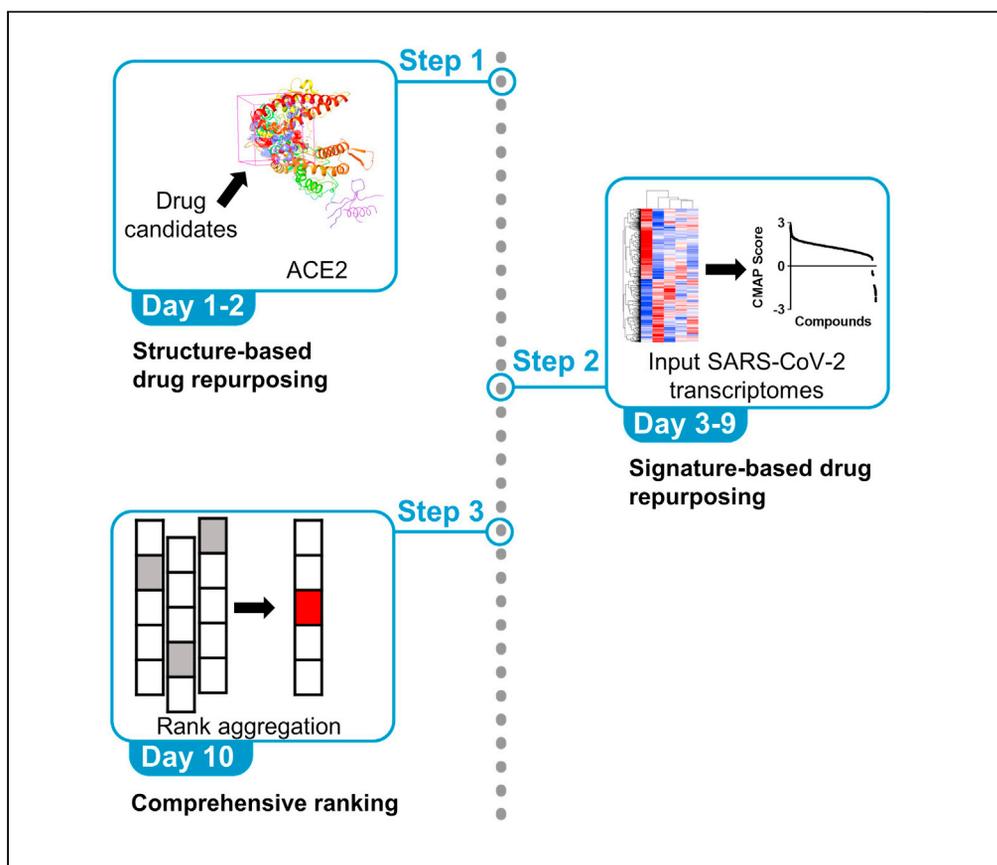


Protocol

Protocol for high-throughput screening of ACE2 enzymatic activators to treat COVID-19-induced metabolic complications



Pin Chen, Chenshu Liu, Zhongyu Zhang, Zilun Li, Sifan Chen, Yutong Lu

chensf26@mail.sysu.edu.cn (S.C.)
yutong.lu@nscg-gz.cn (Y.L.)

Highlights

Structure-based high-throughput virtual screening for ACE2 enzymatic activators

Signature-based drug repositioning for COVID-19-induced metabolic complications

Workflow applicable for other diseases with ACE2 as a potential target

Drug repositioning represents a cost- and time-efficient strategy for drug development. Here, we present a workflow of *in silico* screening of ACE2 enzymatic activators to treat COVID-19-induced metabolic complications. By using structure-based virtual screening and signature-based off-target effect identification via the Connectivity Map database, we provide a ranked list of the repositioning candidates as potential ACE2 enzymatic activators to ameliorate COVID-19-induced metabolic complications. The workflow can also be applied to other diseases with ACE2 as a potential target.

Publisher's note: Undertaking any experimental protocol requires adherence to local institutional guidelines for laboratory safety and ethics.

Chen et al., STAR Protocols 3, 101641
September 16, 2022 © 2022
The Author(s).
<https://doi.org/10.1016/j.xpro.2022.101641>



Protocol

Protocol for high-throughput screening of ACE2 enzymatic activators to treat COVID-19-induced metabolic complications

Pin Chen,^{1,6,7} Chenshu Liu,^{2,3,6} Zhongyu Zhang,^{4,5} Zilun Li,^{2,3} Sifan Chen,^{4,5,*} and Yutong Lu^{1,8,*}

¹National Supercomputer Center in Guangzhou, School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou, Guangdong 510006, China

²Division of Vascular Surgery, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, Guangdong 510080, China

³National-Guangdong Joint Engineering Laboratory for Diagnosis and Treatment of Vascular Diseases, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, Guangdong 510080, China

⁴Guangdong Provincial Key Laboratory of Malignant Tumor Epigenetics and Gene Regulation, Guangdong-Hong Kong Joint Laboratory for RNA Medicine, Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University, Guangzhou, Guangdong 510120, China

⁵Medical Research Center, Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University, Guangzhou, Guangdong 510120, China

⁶These authors contributed equally

⁷Technical contact

⁸Lead contact

*Correspondence: chensf26@mail.sysu.edu.cn (S.C.), yutong.lu@nscg-gz.cn (Y.L.)
<https://doi.org/10.1016/j.xpro.2022.101641>

SUMMARY

Drug repositioning represents a cost- and time-efficient strategy for drug development. Here, we present a workflow of *in silico* screening of ACE2 enzymatic activators to treat COVID-19-induced metabolic complications. By using structure-based virtual screening and signature-based off-target effect identification via the Connectivity Map database, we provide a ranked list of the repositioning candidates as potential ACE2 enzymatic activators to ameliorate COVID-19-induced metabolic complications. The workflow can also be applied to other diseases with ACE2 as a potential target.

For complete details on the use and execution of this protocol, please refer to Li et al. (2022).

BEFORE YOU BEGIN

Overview

The repositioning of commercial drugs may reduce costs of early-stage research and clinical trials, as safety and toxicity of these drugs have been well studied before and after initial approval (Ma et al., 2013). The core concept of drug repositioning is polypharmacology, which is defined as that small molecules would bind several different targets and interconnect different cell signal transduction pathways. Based on this concept, structure-based drug repositioning is usually feasible for certain diseases with approved treatment targets (Parisi et al., 2020). Though several studies reported that COVID-19 would induce metabolic complications (Guo et al., 2021; Gupta et al., 2020; He et al., 2021; Singh and Singh, 2020), the lack of approved targets for these complications tackled efficient structure-based drug repositioning. Our recent study verified ACE2 as a potential target for COVID-19 induced metabolic complications, and a comprehensive workflow with both structure-based and signature-based drug repositioning approaches was constructed (Li et al., 2022). This workflow applied high-throughput virtual docking screening with ACE2 as the potential target, and the small molecules off-target effect identification based on transcriptome signature of COVID-19 via the Connectivity Map (CMAP) (Subramanian et al., 2017) (Methods video S1). We believe the



abovementioned approaches potentially represent a highly efficient methodology for the development of new therapeutics. This protocol could also be applied for other diseases with ACE2 as a potential target, like hypertension, atherosclerosis and non-alcohol fatty liver disease (Crackower et al., 2002; Jiang et al., 2014; Moreira de Macêdo et al., 2014).

Software setup and installation

⌚ Timing: ~ 1 day

A high-performance computing (HPC) cluster with more than 10 nodes operated Linux- or Unix-based system is required to execute this protocol. It costs about 48 core hours for screening DrugBank library by molecular docking method, and 6,912,000 core hours for running molecular dynamics on Tianhe-2 supercomputer with the 2.6 GHz Intel Xeon processors. We recommend that users use 6,912,000 core hours divided by the CPU core number to estimate a rough computation time before running a molecular dynamics model. The prerequisite software can be downloaded from the corresponding websites as shown in the [key resources table](#). The accompanied user manuals provide detailed usage method.

Preparation for application of the CMAP

⌚ Timing: ~1 day

1. Creating an account.
 - a. Prepare a validated personal/office email address from academic institution or non-profit organization, like xxx@xxx.edu.com.
 - b. Access the official website of the CMAP (<https://clue.io>) using common web browsers, i.e., Chrome or Firefox. Press the 'Log in' button and then 'Create an account'.
 - c. Fill in the details required for account setting, as full name, email account, institution, research role and academic training field.
2. Account validation.
 - a. By using email address from academic institution or non-profit organization, the validation process usually takes 1 day, and you will receive a validation email after getting approval from the website.
 - b. Commercial usage requires further contact with the website.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Approved drug list	DrugBank database	https://go.drugbank.com
Predicted compounds and ranked list of virtual binding with ACE2	Table S1	Table_S1.xlsx
Predicted compounds and ranked list from CMAP	Table S2	Table_S2.xlsx
Predicted compounds and ranked list of repositioned drugs by combining method	Table S3	Table_S3.xlsx
Scripts for preparing input and analyzing output of molecular docking	This study	https://doi.org/10.5281/zenodo.6865580
Software and algorithms		
SiteMap	Schrödinger	https://www.schrodinger.com
Glide	Schrödinger	https://www.schrodinger.com
GROMACS	GROMACS website	http://www.gromacs.org
Antechamber	AmberTools	https://ambermd.org
The CMAP	Broad Institute	https://clue.io
Gaussian 03. Revision C.02	Gaussian, Inc.	https://gaussian.com

STEP-BY-STEP METHOD DETAILS

Structure-based virtual screening method

⌚ Timing: ~1 day

Drugs obtained from the DrugBank library were virtually docked with preprocessed ACE2 protein structure (PDB code: 1R42) in certain grid for enzymatic activators screening. The whole scripts for preparing and analyzing molecular docking could be found on <https://doi.org/10.5281/zenodo.6865580>.

1. Determine the docking site. [Troubleshooting 1](#).
 - a. Download the ACE2 protein files (PDB code: 1R42) from RCSB PDB website (<https://www.rcsb.org>) in PDB format (named as "1R42.pdb").
 - b. Open graphical user interface of Schrödinger software and input the "1R42.pdb" file ([Figure 1A](#)).
 - c. Select solvents by clicking the "S" button on the toolbar and then use "delete" on the keyboard to delete them ([Figure 1B](#)).
 - d. Preprocess the protein by clicking the "protein preparation" button on the toolbar, and then a popup will appear. Click the "Process" button on the popup with default settings ([Figure 1C](#)).
 - e. Determine potential receptor binding sites with SiteMap ([Figure 1D](#)). Click the "TASKS" button on the toolbar and search "SiteMap", then a popup will appear. Keep all the default parameters and click "Run" button.
 - f. SiteMap calculates 5 potential binding sites and displays them on the graph interface ([Figure 2](#)). and select "site 2" located in the hinge-bending region of ACE2 protein based on the hypothesis that the potential activators could keep the ACE2 protein pocket open ([Kulemina and Ostrov, 2011](#)) ([Figure 1E](#)).
2. Prepare the receptor grid generation.
 - a. Open the "Receptor Grid Generation".
 - b. Set the docking site box with the centroid of "site 2" (center_x = 68.2 Å, center_y = 64.8 Å, center_z = 39.5 Å) and box length equal to 20 Å ([Figure 1F](#)).
3. Process ligand library.
 - a. Obtain the script from website (<https://doi.org/10.5281/zenodo.6865580>).
 - b. Download the DrugBank library from website (<https://go.drugbank.com>).
 - c. Run the bash script "bash run_process_drugbak.sh" to prepare the input files for docking.
4. Running molecular docking.
 - a. Run the bash script "bash initial_glide_input.sh" to prepare the input files for Glide.
 - b. Run the script "bash run_glide.sh" to start molecular docking.

Note: We recommend standard precision (SP) glide docking, which performs exhaustive sampling and balances between speed and accuracy. Another extra-precision scoring function is optional, named XP, which employs a more sophisticated evaluation for protein-ligand interactions than SP and costs more than double time than SP.

Prioritize repositioned drug candidates

⌚ Timing: ~3 h

Repositioned drug candidates are ranked primarily according to their binding affinity with ACE2 protein. We first selected 7,767 compounds with the potential to directly bind to and activate ACE2.

5. Prioritize the drug candidates using Score term in glide output file. [Troubleshooting 2](#).

The "Score" is defined by the following formula:

$$\text{Score} = a \times \text{vdW} + b \times \text{Coul} + \text{Lipo} + \text{Hbond} + \text{Metal} + \text{Rewards} + \text{RotB} + \text{Site}$$

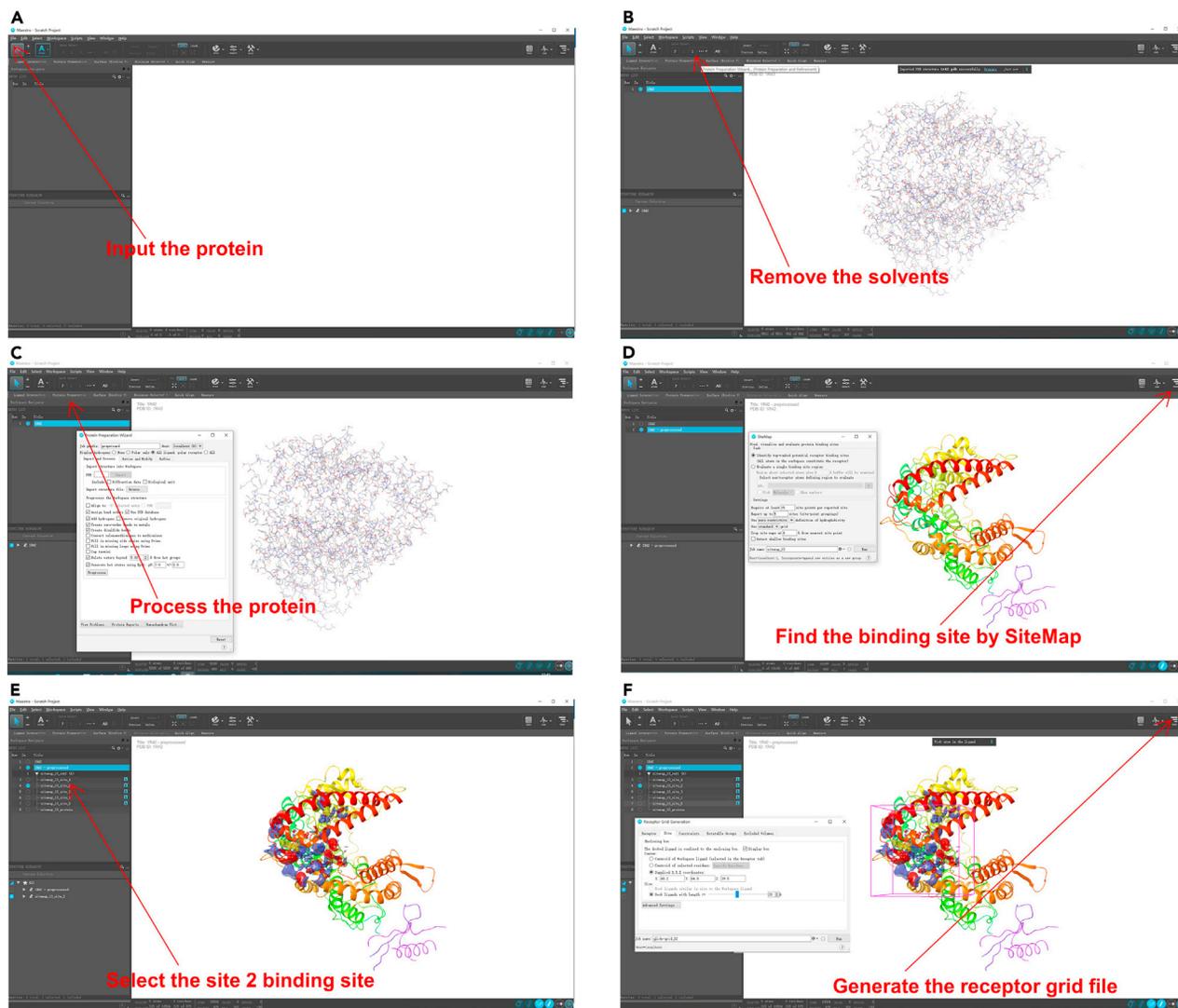


Figure 1. Pre-processing procedures of molecular docking analysis

- (A) Snapshot of “Input the protein” step in Maestro (Schrödinger UI package).
 (B) Snapshot of “Remove the solvents” step in Maestro.
 (C) Snapshot of “Process the protein” step in Maestro.
 (D) Snapshot of “Find the binding sites by SiteMap” step in Maestro.
 (E) Snapshot of “Select the site 2 binding site” step in Maestro.
 (F) Snapshot of “Generate the receptor grid file” step in Maestro.

Where the coefficients of a and b are 0.050 and 0.150, respectively. And the other terms are:

vdW = van der Waals interaction energy.

$Coul$ = Coulomb interaction energy.

$Lipo$ = Lipophilic-contact plus phobic-attractive term.

$HBond$ = Hydrogen-bonding term.

$Metal$ = Metal-binding term (usually a reward).

$Rewards$ = Various reward or penalty terms.

$RotB$ = Penalty for freezing rotatable bonds.

$Site$ = Polar interactions in the active site.

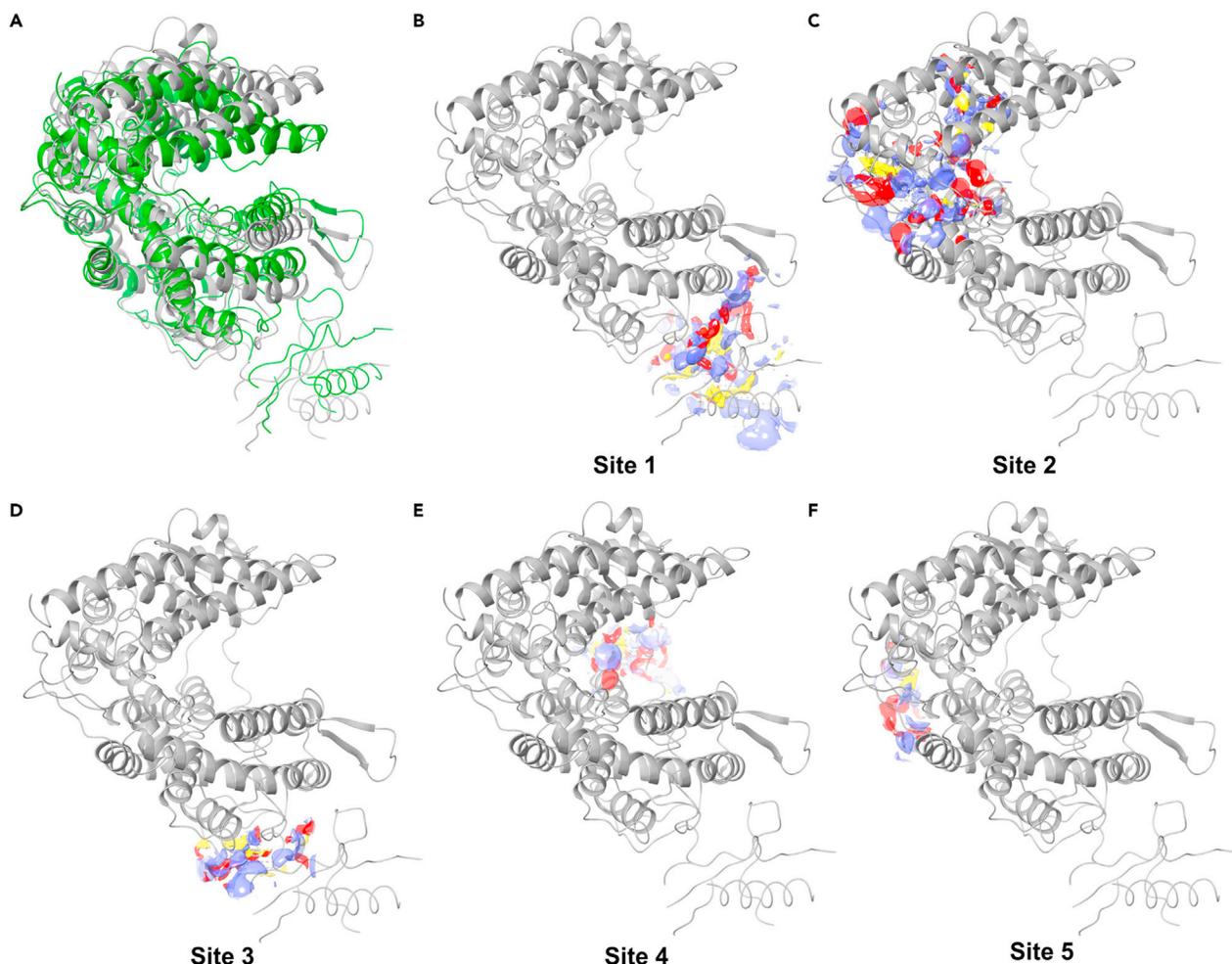


Figure 2. ACE2 structure and its binding sites calculated by Sitemap tool

(A) The native (grey color) and inhibitive (green color, PDB code: 1R4L) state of ACE2 enzyme structure.

(B–F) The site 1 to site 5 calculated by Sitemap tool. Here, regions within the binding site suitable for occupancy by hydrophobic groups or by ligand hydrogen-bond donors, acceptors, or metal-binding functionality, were highlighted.

- a. Run the command of “glide_sort -R glide_result_lib.sdfgz > top_ligand_ranked.dat” to obtain ranked compounds.
- b. Here, we empirically selected 7,767 compounds based on “Score” value no more than -4 kcal/mol (Table S1).

Molecular dynamics analysis of top ranked drugs (optional)

⌚ Timing: ~30 days

To get more accurate binding affinity between compounds and the ACE2 protein and its influence on ACE2-Spike binding, we also carried out molecular dynamics simulations for protein-ligand complex. Here, imatinib and methazolamide were experimentally verified to treat COVID-19 induced metabolic complications in mice via activation of ACE2 enzyme (Li et al., 2022). We chose these two drugs for further molecular dynamics simulation for detailed docking binding structure.

6. Process structure of SARS-CoV-2 spike receptor-binding domain bound with ACE2.

- a. Obtain the crystal structure file of SARS-CoV-2 spike receptor-binding domain bound with ACE2 file (PDB code: 6M0J) from website (<https://www.rcsb.org>) in PDB file (named as "6M0J.pdb").
 - b. Open graphical user interface of Schrödinger software and load the "6M0J.pdb" and "1R42.pdb" file.
 - c. Process structure by removing solvents, ions and other molecules following the same step of "Determine the docking site".
 - d. Align the 6M0J protein to the previous processed 1R42 protein. Select the 6M0J and 1R42 in the workspace and then click the "Quick Align" button.
 - e. Export the 6M0J protein for the next step.
7. Prepare the input files for molecular dynamics.
- a. Prepare ligand. We first calculated the charges using the quantum software (Gaussian 03, revision C. 02) and then used antechamber to obtain a mol2 file containing the RESP charges. The GAFF force field was then applied to the ligand using AmberTools.

```

>antechamber -i ligand.out -fi gout -c resp -o ligand_resp.mol2 -fo mol2 -pfy
>parmchk -i ligand_resp.mol2 -f mol2 -o ligand.frcmod
  
```

- b. Open the tleap tool in AmberTools, and then execute the following command.

```

>source leaprc.gaff
>loadamberparams ligand.frcmod
>lig = loadmol2 ligand_resp.mol2
>check lig
>save amberparm lig lig.prmtop lig.inpcrd
  
```

- c. Use the acpype tool in AmberTools to convert file format for GROMACS.

```

>acpype -p lig.prmtop -x lig.inpcrd
  
```

- d. Prepare protein.

```

>gmx pdb2gmx -f protein.pdb -o protein.gro -water spc -igh
  
```

- e. Run the bash script "bash run_gromacs.sh" to carry out molecular dynamics.
8. Run the bash script "bash mmpbsa_run.sh" to calculate binding affinity by MM-GBSA method. Here, the last 100 ns trajectories were used for analysis. [Troubleshooting 3](#).

Note: Some compounds were not available for simulation due to lack of certain structure information in the GROMACS.

Signature-based drug repositioning screening for treating COVID-19 from the CMAP

⌚ Timing: ~7 days

The CMAP represents a classic analysis tool for signature-based drug repositioning, and is an online platform conducting pattern-matching algorithms to compare input gene signatures with its profiled gene signatures in diverse contexts of over 30,000 chemical perturbagens ([Subramanian et al., 2017](#)). Transcriptome profile from public database or unpublished data could be applied for this step, with differentially expressed genes (DEGs) signature of SARS-CoV-2 infection. A transcriptome profile from public database ([Liao et al., 2020](#)) was applied as the walk-through example in the

following steps. The 10–150 most up-regulated DEGs were required for CMAP analysis, with optional up to 150 most down-regulated DEGs. This analysis aimed to obtain potential candidate compounds for treating COVID-19 induced metabolic complications based on transcriptional signature.

9. Collect transcriptome profile of SARS-CoV-2 infection. [Troubleshooting 4](#).
 - a. Search references reporting RNA-seq of SARS-CoV-2 infection from PubMed (<https://pubmed.ncbi.nlm.nih.gov>), Web of Science (<https://www.webofscience.com>) and in some cases from preprint archive website like Biorxiv (<https://www.biorxiv.org>). For instance, search terms 'transcriptome', 'RNA-seq', 'COVID-19' and 'SARS-CoV-2' were applied in PubMed, and the corresponding bulk RNA-seq focusing on messenger RNA library from Gene Expression Omnibus (<https://www.ncbi.nlm.nih.gov/geo>), National Genomics Data Center (<https://ngdc.cncb.ac.cn>) and et cetera was further collected.
 - b. Process the collected RNA-seq profiles before application in CMAP. Three types of RNA-seq data were usually provided by the references. Namely, original Fastq or Sam file, matrix of RPKM or raw counts of each sample and DEGs tabulations. Process the Fastq/Sam file or matrix of RPKM/counts data with classic transcriptome analysis pipeline to obtain the DEGs tabulation of SARS-CoV-2 infection vs. mock.
 - c. Filter the DEGs with p adjusted value < 0.05, and further rank the DEGs with fold change of SARS-CoV-2 infection vs. mock. The 10–150 most up-regulated DEGs were required for CMAP analysis, with optional up to 150 most down-regulated DEGs.
10. Input transcriptome profile and analyze potential off-target effect of compounds in CMAP. [Troubleshooting 5](#).
 - a. Log in the CMAP official website, and press 'Tools' button in the CMAP official site. Use the 'query' section in the drop-down menu. And you would enter the query interface.
 - b. Name your project and then select 'Gene expression (L1000)' from the drop-down menu of 'Query parameters'. Leave the remaining query parameters as default settings.
 - c. As CMAP prioritizes potential compounds with similarity against the input transcriptomic gene set, input the 10–150 most up-regulated DEGs in the 'UP-regulated genes' box, while input the most down-regulated DEGs, up to 150, in the 'DOWN-regulated genes' box optionally ([Figure 3A](#)). Then press submit, and wait for the query result.
 - d. Once the query was finished, select the query and press 'HEAT MAP'. This would lead you to another webpage called 'CLUE MORPHEUS' and might require a few minutes for data loading ([Figure 3B](#)).
 - e. Once the webpage was loaded, select the 'query result' tab, and the full details of the CMAP analysis would be shown in an online interactive tabulation. To download this online data, press 'File' and use 'Save Dataset' ([Figure 3C](#)). This would generate a GCT file that could be open by Excel or Python.
 - f. CMAP archived several different kinds of perturbations, like compounds treatment, overexpression or knockdown gene manipulations. In this protocol, we focused on the compounds, which was named as 'trt_cp' in the 'pert_type' column. The candidate compounds could be ranked with the 'norm_cs', which was the similarity scores of transcriptional signatures of candidate compounds against the signature of our input transcriptomic gene set, meaning the compounds with the highest positive scores held the most similar transcriptional signatures when comparing with our input downregulated and upregulated genes, whereas the compounds with the lowest negative scores held the most opposite transcriptional signatures. To obtain compounds that could potentially reverse SARS-CoV-2 infection gene alterations, filter and rank the compounds with negative score. You could use certain cutoff value or limit compound numbers as the output criteria for compounds list, in the walk-through instance, we applied -1.35 as the threshold ([Table S2](#)).

Optional: You could input gene signatures based on global proteomics data in the CMAP.

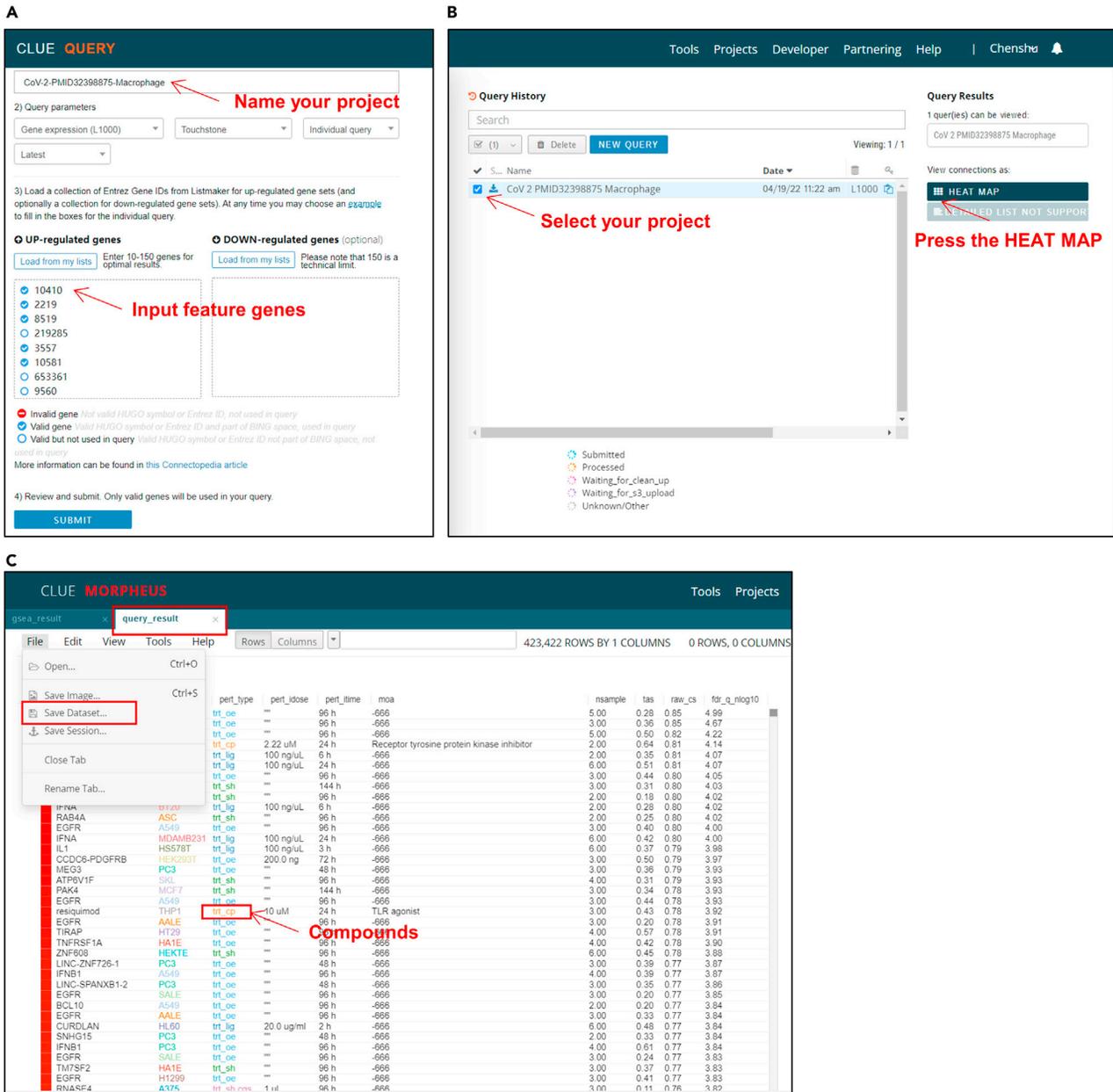


Figure 3. Application of the CMAP
(A) Snapshot of “Clue query” step in the CMAP.
(B) Snapshot of “HEAT MAP” step in the CMAP.
(C) Snapshot of “Query result” step in the CMAP.

Note: The P adjusted value used in DEGs filtering could be tailored in certain cases. Official HUGO symbol or Entrez ID were required in the gene input box, we recommended to use Entrez ID to avoid misinterpretation.

Comprehensive ranking of compounds based on virtual screening and CMAP

⊙ Timing: ~2 h

With the help of virtual docking analysis of potential ACE2 enzymatic activators and CMAP analysis of potential SARS-CoV-2 treating compounds, two lists were obtained respectively. This following step aimed to obtain comprehensive ranking of the overlapped compounds from these two lists to set up the final candidates list for further *in vitro* or *in vivo* validation, which in potential might treat COVID-19 induced metabolic complications via activating ACE2.

11. Overlapped compounds selection from two lists.
 - a. Filter two lists with same ID. As virtual screening would give generic names, database string ID from the compounds dataset and CMAP would give generic names, ID conversion was required for intersection selection. We recommend the use of Canonical Simplified Molecular Input Line Entry System (SMILES) or International Chemical Identifier (InChI) for ID conversion.
 - b. Intersection selects overlapped compounds from two lists using online Venn diagram (<https://bioinfogp.cnb.csic.es/tools/venny/index.html>) or function based on R or python.
 - c. Form the common compounds list with compound ID, virtual docking energy and CMAP score.
12. Comprehensive ranking of overlapped compounds.
 - a. Rank each compound based on virtual docking energy and CMAP score separately.
 - b. Let r_{vir} denote the rank of compounds list based on virtual docking energy, r_{cmap} denote the rank of compounds list based on CMAP score and ρ_{compre} denote the product of two ranks. The formula can be expressed as

$$\rho_{compre} = r_{vir} \times r_{cmap}$$

- c. Prioritize the common compounds based on the new ρ_{compre} value to obtain the final compounds list (Table S3).

Optional: If multiple compounds lists were obtained and comprehensive ranking was required, other more complex rank aggregation algorithm could be applied, like TopKLists or RobustRankAggreg based on R package.

Note: This step gives a comprehensive ranking list with candidates that would potentially treat COVID-19 induced metabolic complications via activating ACE2 enzymatic activity, the further *in vitro* or *in vivo* analysis were required for final verification.

EXPECTED OUTCOMES

In this protocol, we described an *in silico* drug repositioning workflow to identify potential drugs against COVID-19 induced metabolic complications via direct enzymatic activation of ACE2. A complete table listing the predicted binding affinity for each drug-target pair, and CMAP analysis of potential compounds was generated. Finally, a ranked list of the repositioned drug candidates was provided by combining two methods. These results may expedite the drug development for COVID-19 induced metabolic complications. Though the off-target effects of these drugs shall be further verified via biological experiments on an individual basis, this strategy should be helpful to reposition “old drug” for novel uses by facilitating the selection of lead compound for in-depth experimental and clinical evaluation.

LIMITATIONS

There are several limitations for this protocol. First, the high throughput screening applied in this protocol could only give a ranked compounds list with potential to treat COVID-19 induced metabolic complications via activating ACE2, further experimental verification is required. Second, the

compounds in CMAP database are limited and cannot cover the compounds of DrugBank database in molecular docking result list, which might limit the compounds screening set to CMAP database.

TROUBLESHOOTING

Problem 1

Understanding the active sites of a protein and exploiting their function is the cornerstone of structure-based drug design. In this protocol, Sitemap can recommend 5 potential binding sites based the ACE2 structure. However, how to choose suitable bindings sites for screening potential drugs for activating ACE2 enzyme was still difficult for most users (step 1 in [step-by-step method details](#)).

Potential solution

We displayed the ACE2 protein structure in the native and inhibited state observed by experimental equipment in [Figure 2A](#), and aligned these two structures together. It could be clearly seen that the active pocket of the ACE2 protein is like a “crocodile mouth”. When it is in the inhibited state, the “crocodile mouth” is slightly closed, while open in the active state. Therefore, we proposed that a potential activator for ACE2 would be more likely to keep the “crocodile mouth” open. Based on this hypothesis, we further screened the binding sites calculated by Sitemap tool. As shown [Figures 2B–2F](#), the site 1 and site 3 were outside of the active regions. The site 4 was inside the “crocodile mouth”, which was a typical binding site for inhibitors. The site 5 was behind the “crocodile mouth” and had weak influence on structure. The last site 2 was in the hinge-bending region of the “crocodile mouth”, which would be able to prevent the structure from closing. As a result, we chose site 2 as our binding target.

Problem 2

Based on the scores of molecular docking, how to select a certain number of potential drugs for the next experimental verification or fine-grained calculation (for instance, molecular dynamics) largely depends on the user experience. Selecting too many potential drugs will result in long validation time and high costs. Conversely, small selection might result in missing potential drugs (step 5 in [step-by-step method details](#)).

Potential solution

In this protocol, -4 kcal/mol was chosen as the threshold to select potential drugs based on experience. Molecular docking is still a rough screening step that treats proteins as rigid and ignore the role of water molecules. However, this method can help us to remove some unreasonable drugs for the next step. We recommended to establish a threshold based on two aspects. On one hand, the drug and protein should be potentially bound, of which the binding score shall be negative. On the other hand, a certain proportion of drugs (such as 50%), according to the scoring value from low to high, shall be considered based on the specific need for the next step, as well as the capacity of following biological verification.

Problem 3

Molecular dynamics method can obtain more accurate binding energy and kinetic trajectory, but its calculation is very time-consuming. We used Tianhe-2 supercomputer to apply for 240 CPU physical cores for each protein-ligand pair, and 1 μ s kinetic trajectory was obtained. We spent about 30 days to obtain the final data, which may limit the efficiency of this method (step 8 in [step-by-step method details](#)).

Potential solution

Leasing larger-scale computing resources in HPC clusters or supercomputing centers will further accelerate the process of molecular dynamics.

Problem 4

When searching RNA-seq profiles, RNA-seq profiles focusing on different RNA might be found. For instance, miRNA sequencing, long non-coding RNA sequencing, single cell mRNA sequencing or mRNA sequencing could all be found in the searching (step 9 in [step-by-step method details](#)).

Potential solution

In order to input differentially expressed genes of transcriptome profile to CMAP, bulk-RNA sequencing focusing on mRNA was required. The detailed transcriptome analysis pipeline based on mRNA-seq could be found on ENCODE project (<https://www.encodeproject.org/data-standards/ma-seq/long-rnas>) or Galaxy (<https://usegalaxy.org>).

Problem 5

When inputting several transcriptome profiles in CMAP, different compound lists with lack of consistency might be obtained (step 10 in [step-by-step method details](#)).

Potential solution

We recommend inputting at least three transcriptome profiles in CMAP, and compounds that presented in more than one list could be included for further study.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Yutong Lu (yutong.lu@nscg-gz.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The code generated during this study is available at the Zenodo repository: <https://doi.org/10.5281/zenodo.6865580>.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.xpro.2022.101641>.

ACKNOWLEDGMENTS

This work is supported by the Program for Guangdong Introducing Innovative and Entrepreneurial Teams (No. 2016ZT06D211), Guangdong Basic and Applied Basic Research Foundation (No. 2021B1515020005), Guangdong Science and Technology Department (No. 2020B1212060018, No. 2020B1212030004), National Natural Science Foundation of China (No. 81970406), Guangdong Basic and Applied Basic Research Foundation (No. 2020A1515011471).

AUTHOR CONTRIBUTIONS

Conceptualization, Y.L., S.C.; investigation, P.C., C.L.; Methodology; P.C., C.L.; Software, P.C., C.L.; Writing – Original Draft, P.C., C.L.; Writing – Review & Editing, Y.L., S.C., Z.Z., Z.L.; Funding Acquisition, Y.L., S.C.; Supervision, Y.L., S.C.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

- Crackower, M.A., Sarao, R., Oudit, G.Y., Yagil, C., Kozieradzki, I., Scanga, S.E., Oliveira-dos-Santos, A.J., da Costa, J., Zhang, L., Pei, Y., et al. (2002). Angiotensin-converting enzyme 2 is an essential regulator of heart function. *Nature* *417*, 822–828.
- Guo, Z., Jiang, S., Li, Z., and Chen, S. (2021). Metabolic syndrome "interacts" with COVID-19. *BIO Integration* *1*, 168–177.
- Gupta, A., Madhavan, M.V., Sehgal, K., Nair, N., Mahajan, S., Sehrawat, T.S., Bikdeli, B., Ahluwalia, N., Ausiello, J.C., Wan, E.Y., et al. (2020). Extrapulmonary manifestations of COVID-19. *Nat. Med.* *26*, 1017–1032.
- He, X., Liu, C., Peng, J., Li, Z., Li, F., Wang, J., Hu, A., Peng, M., Huang, K., Fan, D., et al. (2021). COVID-19 induces new-onset insulin resistance and lipid metabolic dysregulation via regulation of secreted metabolic factors. *Signal Transduct. Target. Ther.* *6*, 427.
- Jiang, F., Yang, J., Zhang, Y., Dong, M., Wang, S., Zhang, Q., Liu, F.F., Zhang, K., and Zhang, C. (2014). Angiotensin-converting enzyme 2 and angiotensin 1-7: novel therapeutic targets. *Nat. Rev. Cardiol.* *11*, 413–426.
- Kulemina, L.V., and Ostrov, D.A. (2011). Prediction of off-target effects on angiotensin-converting enzyme 2. *J. Biomol. Screen* *16*, 878–885.
- Li, Z., Peng, M., Chen, P., Liu, C., Hu, A., Zhang, Y., Peng, J., Liu, J., Li, Y., Li, W., et al. (2022). Imatinib and methazolamide ameliorate COVID-19-induced metabolic complications via elevating ACE2 enzymatic activity and inhibiting viral entry. *Cell Metab.* *34*, 424–440.e7.
- Liao, M., Liu, Y., Yuan, J., Wen, Y., Xu, G., Zhao, J., Cheng, L., Li, J., Wang, X., Wang, F., et al. (2020). Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat. Med.* *26*, 842–844.
- Ma, D.L., Chan, D.S.H., and Leung, C.H. (2013). Drug repositioning by structure-based virtual screening. *Chem. Soc. Rev.* *42*, 2130–2141.
- Moreira de Macêdo, S., Guimarães, T.A., Feltenberger, J.D., and Sousa Santos, S.H. (2014). The role of renin-angiotensin system modulation on treatment and prevention of liver diseases. *Peptides* *62*, 189–196.
- Parisi, D., Adasme, M.F., Sveshnikova, A., Bolz, S.N., Moreau, Y., and Schroeder, M. (2020). Drug repositioning or target repositioning: a structural perspective of drug-target-indication relationship for available repurposed drugs. *Comput. Struct. Biotechnol. J.* *18*, 1043–1055.
- Singh, A.K., and Singh, R. (2020). Hyperglycemia without diabetes and new-onset diabetes are both associated with poorer outcomes in COVID-19. *Diabetes Res. Clin. Pract.* *167*, 108382.
- Subramanian, A., Narayan, R., Corsello, S.M., Peck, D.D., Natoli, T.E., Lu, X., Gould, J., Davis, J.F., Tubelli, A.A., Asiedu, J.K., et al. (2017). A next generation connectivity Map: L1000 platform and the first 1,000,000 profiles. *Cell* *171*, 1437–1452.e17.