JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim

Application Note

CROMATIC: *Cross*-Relationship *Map* of Cavities from Coronaviruses

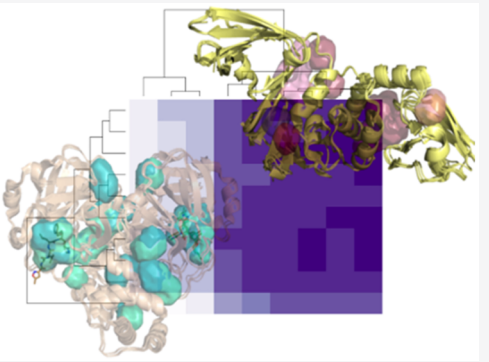
Lydia Siragusa,* Gabriele Menna, Fabrizio Buratta, Massimo Baroni, Jenny Desantis, Gabriele Cruciani, and Laura Goracci

Cite This: *J. Chem. Inf. Model.* 2022, 62, 2901−2908

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the etiologic agent of COVID-19 disease, has rapidly imposed an urgent need to identify effective drug candidates. In this context, the high resolution and non-redundant beta-Coronavirus protein cavities database is pivotal to help virtual screening protocols. Furthermore, the cross-relationship among cavities can lead to highlighting multitarget therapy chances. Here, we first collect all protein cavities on SARS-CoV-2, SARS-CoV, and MERS-CoV X-ray structures, and then, we compute a similarity map by using molecular interaction fields (MIFs). All the results come together in CROMATIC (CROss-relationship MAp of CaviTIes from Coronaviruses). CROMATIC encloses both a comprehensive and a non-redundant version of the cavities collection and a similarity map revealing, on the one hand, cavities that are conserved among the three Coronaviruses and, on the other hand, unexpected similarities among cavities that can represent a key starting point for multitarget therapy strategies. Similarity analysis was also performed for the available structures of SARS-CoV-2 spike variants, linking sequence mutations to three-dimensional interaction alterations. The CROMATIC repository is freely available to the scientific community at https://github.com/moldiscovery/sars-cromatic.



1. INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) was identified as the responsible agent for the COVID-19 pandemic. This virus spread very quickly, infecting millions of people and causing an estimated number of 5 million deaths.[1] SARS-CoV-2 belongs to the class of beta-coronaviruses, which also includes severe acute respiratory syndrome coronavirus (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS-CoV). SARS-CoV-2 shows a sequence similarity with SARS-CoV and with MERS-CoV of about 79% and 50%, respectively.[2]

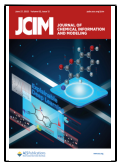
One of the most interesting starting points for exploring an organism is certainly its three-dimensional structure of the proteome obtained from crystallography experiments. Indeed, in the years 2020 and 2021, about 1700 new SARS-CoV-2 crystallographic structures were deposited in the Protein Data Bank (PDB), effectively creating a pool of extremely valuable knowledge and potential weapons to fight SARS-CoV-2.

For this reason, some annotated databases of three-dimensional structures of SARS-CoV-2 proteins have been published and are at the service of the scientific community, helping the exploration and understanding of potential therapeutic protein targets.

Among these, of notable importance are CoV3D[3] and SARS-CoV-2-3D.[4] CoV3D is focused on protein spike structures, while SARS-CoV-2-3D contains both experimentally solved structures and 3D models together with binding sites, protein−ligand docking, protein interactions with human proteins, and impacts of mutations. In other very interesting studies, the authors described druggability assessment for the SARS-CoV-2 proteome followed by docking of phytomolecules,[5] the mapping of druggable binding pockets on the experimental structures of 15 SARS-CoV-2 proteins genetic variabilities,[6] a web-based server for navigating possible interactions between SARS-CoV-2 and human proteins,[7] and a web server that predicts the binding modes between SARS-CoV-2 proteins and ligands.[8] In addition to annotated and systematic databases, numerous studies on SARS-CoV-2 potential drug candidate binding sites have been reported.[9−16]

In this scenario, focusing on the beta-Coronavirus family, we wanted to add and expand some missing information: (a) collecting and annotating all high quality X-ray structure potential binding sites and then (b) generating, for the first time, their similarity maps. We built CROMATIC (CROss-relationship MAp of CaviTIes from Coronaviruses), the first exhaustive cavities collection of SARS-CoV-2, SARS-CoV, and MERS-CoV, both in a comprehensive and non-redundant version (Figure 1). We included the three Coronaviruses in our analysis in order to have a general picture of the

Received: February 11, 2022
Published: June 13, 2022



https://doi.org/10.1021/acs.jcim.2c00169
J. Chem. Inf. Model. 2022, 62, 2901−2908

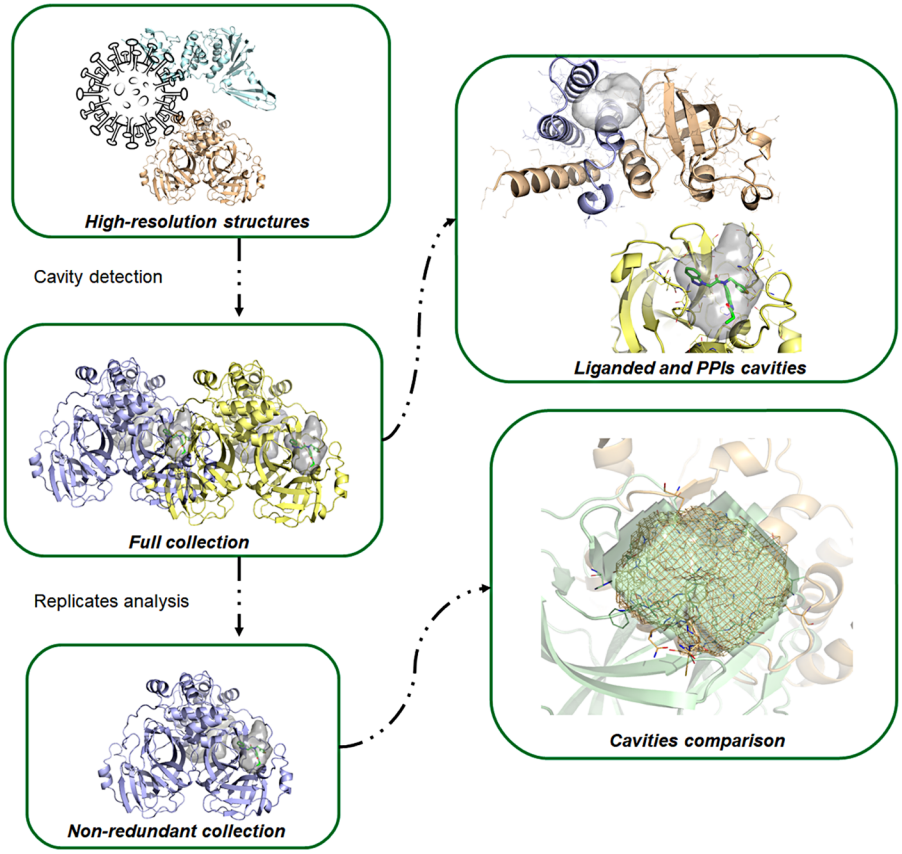


Figure 1. CROMATIC workflow.

variabilities of the active sites, especially for the design of potential broad spectrum inhibitors that, in the scenario of the development of a new pathogenic coronavirus, are certainly more promising. With respect to what is already available to the scientific community, CROMATIC proposes a cavities cross-relationship map by comparing their ligand "image" derived from the GRID molecular interaction fields (MIFs).[17,18] The map might help to reveal similarities and divergences among Coronavirus cavities, highlighting and confirming conserved sites and multitarget therapy opportunities. Furthermore, some useful information is added, such as the involvement in protein−protein interactions (PPIs), cases where the cavities have been targeted by ligands and the corresponding 3D structure of the ligand. Finally, a cross-relationship analysis of the SARS-CoV-2 spike variants is reported to highlight similarities and differences in the receptor binding domain (RBD) interacting with angiotensin converting enzyme-2 (ACE-2).

The set of this information and three-dimensional data represent an essential resource for the scientific community in order to have an overview of the Coronaviruses cavities collection and to exploit it for virtual screening protocols.

2. MATERIAL AND METHODS

The workflow employed for the development of CROMATIC is summarized in Figure 1.

Protein Data Set. The X-ray structures, with resolution ≤ 2.5, of SARS-CoV-2, SARS-CoV, and MERS-CoV proteins were downloaded from the Protein Data Bank on October 15, 2021. To complete the protein panel for SARS-CoV-2, we added also the cryo-EM structures of (I) closed and open states of spikes as a trimer (PDB IDs: 6vxx and 6vyb, respectively), (II) Nsp 12 (PDB ID: 7aap, chain A), and (III) ORF3a (PDB ID: 7kjr, chain A), since no X-ray structures were available for these proteins.

Here, 1028 biological units were thus collected. Each biological unit was subsequently pretreated, removing water molecules, any ligands, and retaining only structural metals. In order to consider each protein separately, the biological units were divided into chains, and all subsequent steps were performed on the individual chains and not on the multimer. A total number of 1271 chains was obtained, 1033 for SARS-CoV-2, 197 for SARS-CoV, and 41 for MERS-CoV.

Protein Cavities Detection. On each chain, the cavities were calculated using flapsite.[17,18] Each chain is embedded into a three-dimensional lattice, and at each point, a buriedness value is calculated and weighted using the GRID DRY probe. Points that are buried, and hydrophobic enough are retained and then subjected to an erosion/dilation algorithm to smooth the cavity region. Only cavities with sufficient sizes and hydrophobic volumes to accommodate a drug-like ligand are retained. The ideal ranges of size and hydrophobic volume were calculated from a data set of about 50,000 liganded cavities from the entire Protein Data Bank. The analysis ended up with 3928 cavities for SARS-CoV-2, 642 cavities for SARS-CoV, and 95 cavities for MERS-CoV for a total of 4665 cavities.

Sequence Clustering. In order to group the sequences of the single chains belonging to the same protein, we performed a clustering analysis using Clustal Omega.[19] The residue sequences, coming from the X-ray chains, were compared with an "all against all" approach, and the resulting matrix was

https://doi.org/10.1021/acs.jcim.2c00169
J. Chem. Inf. Model. 2022, 62, 2901−2908

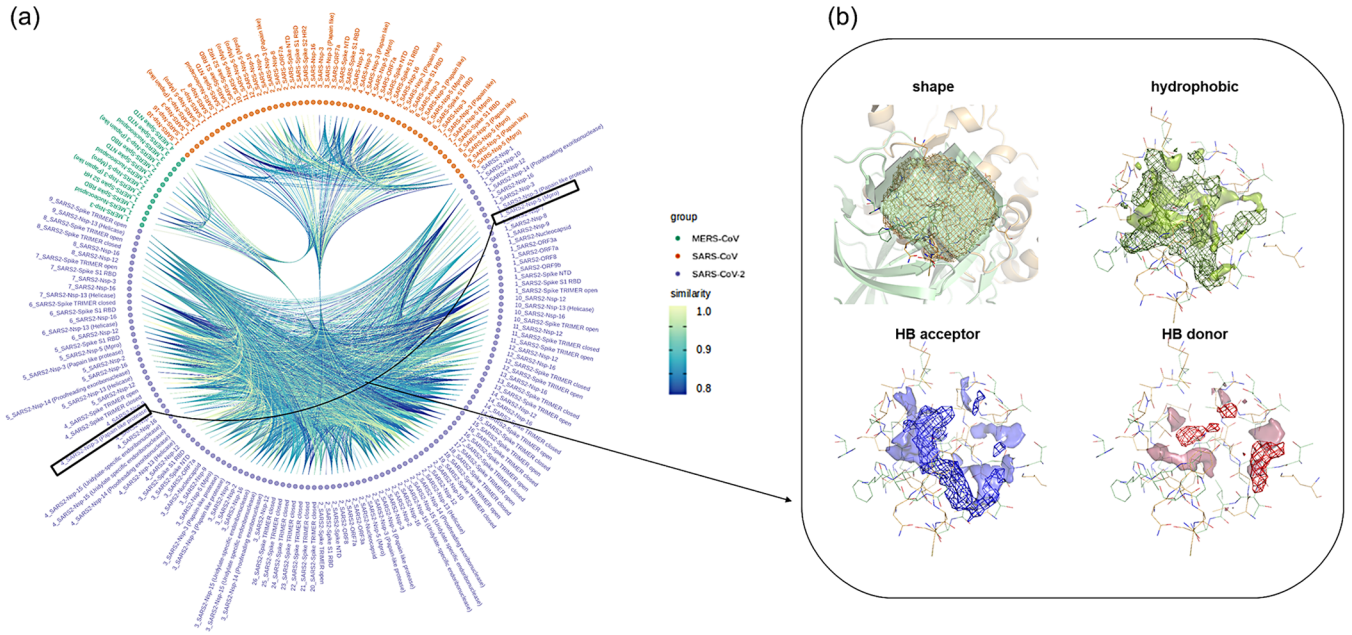


Figure 2. (a) Cross-relationship map, reporting similar pairs of cavities among and across Coronaviruses. M^{pro} and PL^{pro} active sites connections are highlighted. (b) Superposition of M^{pro} and PL^{pro} binding sites from SARS-CoV-2. M^{pro} residues are reported as green wireframes and PL^{pro} residues as orange wireframes. M^{pro} MIFs are reported as solid surfaces and PL^{pro} MIFs as wireframes. PL^{pro} PDB ID: 7ofs; M^{pro} PDB ID: 6yvf.

subjected to hierarchical clustering, using an "average" linkage approach. This procedure was applied for each of the the organisms separately, resulting in 27 clusters for SARS-CoV-2, 18 clusters for SARS-CoV, and 8 clusters for MERS-CoV. Each cluster, and therefore each sequence, was subsequently annotated following the classification reported by NCBI's analysis of the virus's genome sequence.[20]

Representative Cavities Selection. In order to reduce the redundancy coming from similar structures of the same protein and thus generating the same cavity on several chains, we proceeded with a redundancy analysis divided into two steps, both on the (a) chain level and on the (b) cavity level.

(a) First, within each sequence cluster, we tried to understand the conformational variability of the related protein chains to reduce the redundancy due to very similar chains. The RMSd between all the chain pairs was calculated with TMalign;[21] the resulting matrix was subjected to hierarchical clustering, applying a RMSd distance threshold equal to 1.5. For each cluster, a medoid chain was selected, excluding chains featured with nucleic acid segments, preferring chains provided with ligands, with a minimum amount of missing residues and with the lowest resolution value. Here, 36 representative chains were obtained for SARS-CoV-2, 26 for SARS-CoV, and 8 for MERS-CoV.

(b) Even if the cavities have been calculated on chains with RMSd values greater than 1.5, and therefore theoretically on different conformations, there could be cavities that are very similar to each other because they fall on portions of two chains that are well aligned and therefore with a local RMSd value lower than 1.5. This would lead to further redundancy at the cavity level. In a subsequent step, we then continued with a cluster analysis on cavity residues only. An *in-house* algorithm was used for this purpose. The residues defining cavities are three-dimensionally aligned: a similarity, calculated by means of a BLOSUM62 matrix,[22] greater than 3.0 and an RMSd lower than 2.2 are the criteria for defining two similar cavities and therefore belonging to the same cluster. Following this approach, we ended up with 149 representative cavities for SARS-CoV-2, 59 representative cavities for SARS-CoV, and 15 representative cavities for MERS-CoV, for a total of 223 cavities which constitute the non-redundant Coronavirus cavity collection.

Cavities Annotation. Each cavity was annotated with several pieces of information: (1) the PDB code of the protein chain to which it belongs, (2) the resolution of the protein chain to which it belongs, (3) the name of the protein to which it belongs, (4) the protein domain on which it is located, (5) any cocrystallized ligand, and (6) any interacting chain (cavity involved in protein−protein Interactions; PPIs).

Information 1 and 2 were deduced directly from the PDB,[23] information 3 and 4 from NCBI's analysis of the virus's genome sequence,[20] and information 5 and 6 by computing the occupancy volume of any ligand or chain, respectively, within each cavity of the same biological unit.

Cavities Similarities. In order to compare all the 223 representative cavities with an "all against all" approach, we used BioGPS approach,[17,18] which is based on a molecular interaction fields (MIFs) comparison. Each cavity is compared with all the other cavities in the data set, and a Tanimoto score is given, indicating the MIFs similarity (shape, hydrophobic, hydrogen bonding donor and acceptor). A square matrix of 223 × 223 cavities was thus obtained and used for extracting all possible similar pairs among Coronavirus protein cavities.

Spike Variants analysis. All SARS-CoV-2 spike variant crystallographic structures were searched in the literature. A manual selection of the best structures in terms of resolution led to a total of 11 SARS-Cov-2 variant structures (Table S2, Supporting Information). In this analysis, also CryoEM structures were considered in order to cover as many variants as possible. The wild type (WT) spike protein in complex with

https://doi.org/10.1021/acs.jcim.2c00169
J. Chem. Inf. Model. 2022, 62, 2901−2908

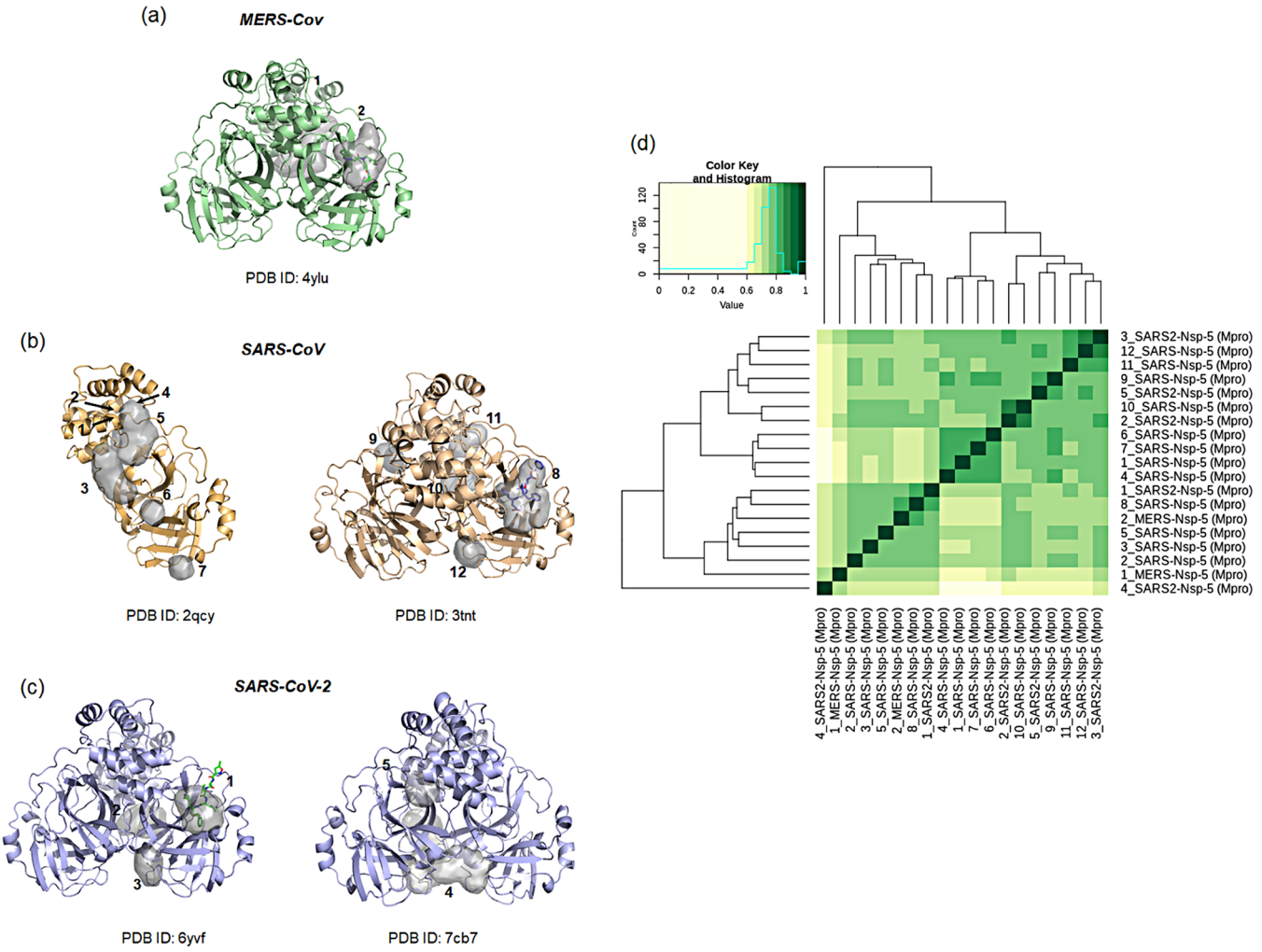


Figure 3. M^{pro} representative cavities for (a) MERS-CoV, (b) SARS-CoV, (c) SARS-CoV-2, and (d) their cross-relationship MIFs map. Dark green cells indicate similar cavity pairs. For SARS-CoV and SARS-CoV-2, two different conformations were found in the Protein Data Bank from the RMSd chain clustering.

angiotensin converting enzyme-2 (ACE-2) (PDB ID: 6lzg) was added to the following analysis, ending up with 12 structures.

For each variant only the receptor binding domain (RBD) was extracted for further analysis. All selected RBD domains were then aligned to the wild type spike protein in complex with ACE-2, by using Pymol.[25]

The RBD region interacting with ACE-2 was detected by using the flapsite tool.[17,18] Molecular interaction fields (MIFs) for the RBD interacting regions were computed by using BioGPS.[17,18] All RBD interacting regions were compared with an "all against all" approach to quantify similarities and differences in term of MIFs (shape, hydrophobic, hydrogen bonding donor and acceptor). The resulting similarity matrix was used to perform a principal component analysis (PCA), by using Rstudio, for helping with the interpretability of similarities and differences.

Cloud Computing. All the computation from the previous steps were performed on Amazon Web Services CPUs, using an Intel(R) Xeon(R) Platinum 8223CL CPU@3.00 GHz. Sequence and structure clustering took 4 h, cavity detection 30 min, and MIFs and their similarity computation 5 h.

3. RESULTS AND DISCUSSION

Cross-Relationship. The comprehensive and non-redundant collections of CROMATIC contain 4665 and 223 cavities, respectively. Detailed information about the numbers of proteins, structures, and cavities is available in the Supporting Information (Table S1).

The non-redundant cavities collection allows to abolish replicates but still maintain a certain degree of conformational variability useful for virtual screening protocols.

To reveal similarities and differences among the Coronaviruses cavities collection, both inter- and intraorganisms, we used MIFs comparison,[17,18] which can identify protein cavities sharing a similar three-dimensional arrangement of nonbonded interactions and, in particular, the shapes, hydrophobic interactions, and hydrogen bonding (HB) donor and acceptor interactions. We considered similar cavities as those having matching MIFs of at least 80% (similarity_value $\geq$ 0.8). Similar pairs were extracted and used to build a map where connections reveal similarities (Figure 2a). Following the principle that similar cavities bind similar ligands, this map can help in focusing attention on multitarget therapy development. Among all the potential similar pairs, we retrieved the binding sites of the main protease (Nsp 5, M^{pro}) and the papain-like

https://doi.org/10.1021/acs.jcim.2c00169
J. Chem. Inf. Model. 2022, 62, 2901−2908

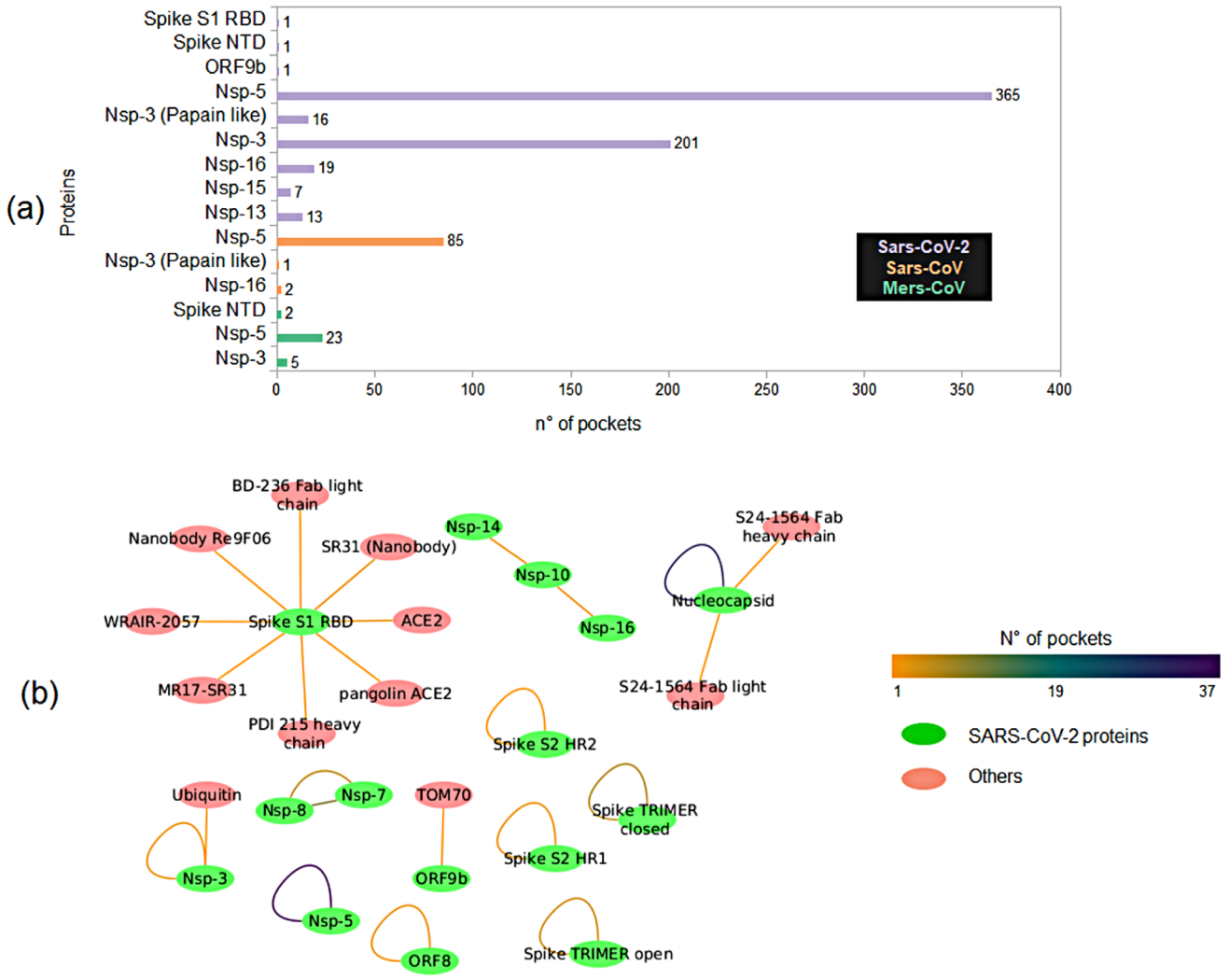


Figure 4. (a) Number of liganded cavities in full collection and (b) cavities involved in protein−protein interactions in SARS-CoV-2 full collection.

protease (Nsp 3, PL^{pro}) in SARS-CoV-2, already reported to bind the same ligand.[24] We proceeded to visualize the superposition of the two binding sites, revealing their similar MIFs (Figure 2b). Regions where cavity MIFs are matching represent regions where the two cavities are responsible for the same type of interaction with a putative ligand. The two sites are connected in the cross-relationship map, sharing a similar 3D arrangement of nonbonded interactions, confirming the importance of investigating this aspect for opening new multitarget therapy scenarios.

From the global map, the user can also extract information on one particular protein to understand which cavities are conserved among the three Coronaviruses. As an example, we extracted similar cross-relationships of M^{pro} from the three Coronaviruses to highlight which cavities are conserved and which are not (Figure 3). From the heatmap (Figure 3d), it is possible to note that the active sites are quite conserved, mostly between SARS-CoV and SARS-Cov-2 (8_SARS_Nsp-5 (M^{pro}) vs 1_SARS2-Nsp-5 (M^{pro})). Also, two cavities at the interface of the dimer are conserved in SARS-CoV and SARS-Cov-2 (12_SARS_Nsp-5 (M^{pro}) vs 3_SARS2-Nsp-5 (M^{pro}) and 10_SARS_Nsp-5 (M^{pro}) vs 2_SARS2-Nsp-5 (M^{pro})).

The previous results regarding the cross-relationship in both *inter-* and *intra*-Coronaviruses are also confirmed by some known cases of drug candidates acting according to promiscuous mechanisms.

The first concerns the drug candidate PF-07321332. This molecule came from an anti-SARS-CoV drug discovery campaign followed by hit-to-lead optimization, inhibiting the M^{pro} protein. It was shown to be active as inhibitors of M^{pro} in SARS-CoV-2 and MERS-CoV.[27] This confirms what has been previously demonstrated regarding the active site of M^{pro} which is quite conserved in the three Coronaviruses.

Another very interesting example concerns disulfiram, a candidate in a Phase 2 clinical trial, which seems to act according to a promiscuous mechanism. Indeed, it has been shown to inhibit both M^{pro} and PL^{pro} in SARS-COV-2,[28] confirming, also in this case, the intraorganism cross-relationship demonstrated by our analysis.

Liganded and PPIs Cavities. To help the navigation of interesting data, the presence of a crystallized ligand and the presence of any interacting protein are highlighted for each cavity. Therefore, it is possible to have an immediate synthesis both of the liganded and the protein−protein interaction (PPI) cavities collection (Figure 4). For all three Coronoviruses, as expected, the greatest effort in the search for potential ligands was made toward the M^{pro} protein (Nsp5) and in the case of SARS-CoV-2 also toward Nsp3. Interestingly, there is also room for research into potential protein−protein interaction

https://doi.org/10.1021/acs.jcim.2c00169
J. Chem. Inf. Model. 2022, 62, 2901−2908

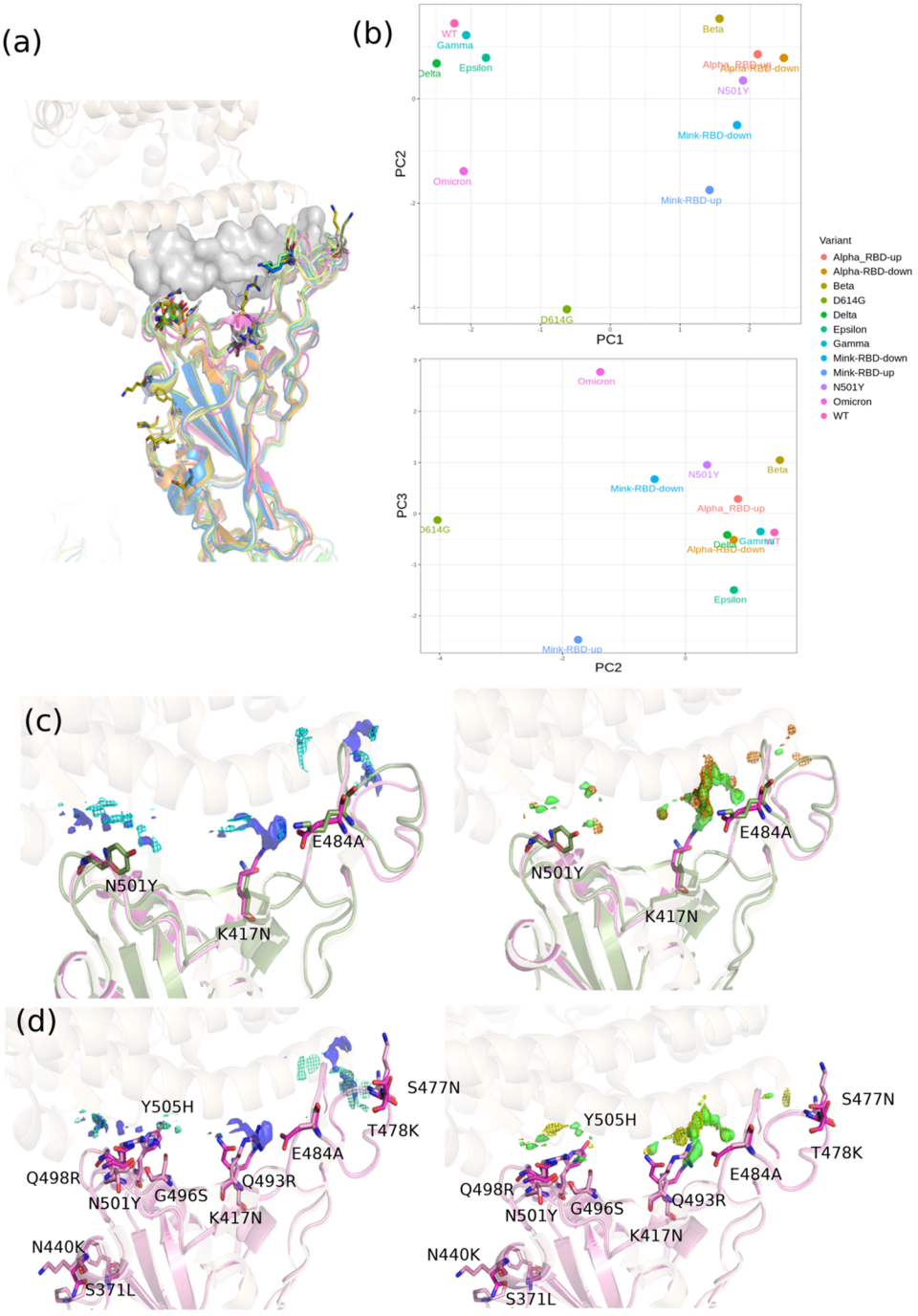


Figure 5. SARS-CoV-2 spike variants analysis. (a) RBD domains from 12 different variants aligned. Regions interacting with ACE-2 and ACE-2 are displayed as gray shapes and wheat cartoons, respectively. Mutations are reported as sticks. (b) PCA score plots (PC1 vs PC2 and PC2 vs PC3). (c) Superposition of WT and Beta variant, displayed as magenta and green cartoons, respectively. HB acceptor interactions are displayed as blue surfaces and cyan wireframes for WT and Beta variant, respectively. Hydrophobic interactions are displayed as green surfaces and orange wireframes for WT and Beta variant, respectively. (d) Superposition of WT and Omicron variant, displayed as magenta and pink cartoons, respectively. HB acceptor interactions are displayed as blue surfaces and light green wireframes for WT and Omicron variant, respectively. Hydrophobic interactions are displayed as green surfaces and yellow wireframes for WT and Omicron variant, respectively.

disruptors. PPI cavities include both those involved in the dimerization/trimerization of viral proteins (Nsp5, spike) and in the binding of human proteins or antibodies (spike, Nsp3).

SARS-CoV-2 Spike Variants. SARS-CoV-2 has accumulated several mutations. The emergence of new variants has become an essential event to study and understand. The first SARS-CoV-2 variant of concern (VOC), B.1.1.7 (also known as Alpha variant), was identified in the United Kingdom (UK) in September 2020. The last one, Omicron BA.5, was identified in February 2022. In the middle, there are about 40 more variants.[26]

CROMATIC also includes the crystallographic structures of the variants currently available in the Protein Data Bank. An accurate bibliographic search, combined with a manual selection of the best resolution structures, also useful for reducing redundancy, led to a total of 12 SARS-CoV-2 spike

https://doi.org/10.1021/acs.jcim.2c00169
J. Chem. Inf. Model. 2022, 62, 2901−2908

variant structures (Table S2, Supporting Information). The wild type (WT) structure was also considered to define similarities and differences.

For each structure, only the spike receptor binding domain (RBD), that enables the virus to infect human cells by interacting with angiotensin converting enzyme-2 (ACE-2), was considered.

The region interacting with ACE-2 was defined by using the flapsite algorithm[17,18] (Figure 5a). We then compared such a region from all the variant structures using an "all against all" approach by comparing their molecular interaction fields (MIFs). The resulting similarity matrix was used to perform a principal component analysis (PCA) in order to reveal the cross-relationships and the variability of SARS-CoV-2 variants (Figure 5b). The explained variance for each principal component was PC1 = 34.5%, PC2 = 57.3%, and PC3 = 71.9%. Interestingly, the Omicron variant RBD is positioned away from all the others, resulting in the most different one, consistent with the high number of mutations reported.

An interesting question that arises from this analysis concerns what kind of interaction actually makes these variants different and therefore which are the interactions responsible for the diverse affinity with the ACE-2 enzyme. We report an example where we compare the RBD in its wild type (WT), both with the Beta variant and the Omicron variant. The first lesson we can learn from this analysis is that the hydrophobic interactions (Figure 5c and d, right part) are very similar to each other by comparing the WT with both Beta and Omicron. What makes Omicron actually different from WT and Beta are the polar interactions (specifically hydrogen bonding acceptor; HB acceptor) (Figure 5c and d, left part). For example, Q493R and E484A mutations are responsible for very diverse HB acceptor interacions in Omicron RBD. In Beta RBD, instead, the N501Y mutation causes small differences in terms of HB acceptor interactions, having both Asparigine (N) and Tyrosine (Y) and the capability to accept and hydrogen bond through the carbonyl group.

We can then conclude that not all the mutations can lead to real differences in term of interactions with ACE-2, drugs, or vaccines. A three-dimensional analysis of the interactions made by these residues can help to define which are the most interesting anchoring points.

4. CONCLUSIONS

A freely available cavities collection named CROMATIC was built to collect the binding sites on SARS-CoV-2, SARS-CoV, and MERS-CoV targets, providing all 3D protein structures, cavities, ligands, and interactors. The similarities and divergences among the cavities allows, for the first time, to understand which ones are conserved among the three beta-Coronaviruses and, even more importantly, to find similar ones in order to design multitarget therapies. A comprehensive and systematical annotation of cavities helps in navigating and exploring the 3D data making the cavities collection a useful tool for drug investigation. In CROMATIC, the upload of new proteins will be manually curated to avoid adding too similar of structures that will generate redundancy.

5. DATA AND SOFTWARE AVAILABILITY

Files of proteins, cavities, and ligands can be downloaded from https://github.com/moldiscovery/sars-cromatic. Annotation data are also provided. A cross-relationship map is available at the same link both as a full square matrix and as a similar cavity pairs file.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.2c00169.

Table S1: Detailed information on the CROMATIC collection in terms of number of proteins, structures, and cavities. Table S2: SARS-COV-2 spike variant structures collection. (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Lydia Siragusa − *Molecular Horizon srl, Bettona 06084, Italy; Molecular Discovery, Ltd., Kinetic Business Centre, Borehamwood, Hertfordshire WD6 4PJ, United Kingdom;* orcid.org/0000-0003-4596-7242; Email: lydia@molhorizon.it

Authors

Gabriele Menna − *Molecular Discovery, Ltd., Kinetic Business Centre, Borehamwood, Hertfordshire WD6 4PJ, United Kingdom*

Fabrizio Buratta − *Molecular Discovery, Ltd., Kinetic Business Centre, Borehamwood, Hertfordshire WD6 4PJ, United Kingdom*

Massimo Baroni − *Molecular Discovery, Ltd., Kinetic Business Centre, Borehamwood, Hertfordshire WD6 4PJ, United Kingdom*

Jenny Desantis − *Laboratory for Chemometrics and Molecular Modeling, Department of Chemistry, Biology, and Biotechnology, University of Perugia, 06123 Perugia (PG), Italy*

Gabriele Cruciani − *Laboratory for Chemometrics and Molecular Modeling, Department of Chemistry, Biology, and Biotechnology, University of Perugia, 06123 Perugia (PG), Italy;* orcid.org/0000-0002-4162-8692

Laura Goracci − *Laboratory for Chemometrics and Molecular Modeling, Department of Chemistry, Biology, and Biotechnology, University of Perugia, 06123 Perugia (PG), Italy*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.2c00169

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Regione Umbria (Italy) for financial support within the PRO−CURA project and Daniela Trisciuzzi for fruitful discussions.

■ REFERENCES

(1) *WHO Coronavirus (COVID-19) Dashboard.* https://covid19.who.int/(accessed 2022-01-20).

(2) Lu, R.; Zhao, X.; Li, J.; Niu, P.; Yang, B.; Wu, H.; Wang, W.; Song, H.; Huang, B.; Zhu, N.; Bi, Y.; Ma, X.; Zhan, F.; Wang, L.; Hu, T.; Zhou, H.; Hu, Z.; Zhou, W.; Zhao, L.; Chen, J.; Meng, Y.; Wang, J.; Lin, Y.; Yuan, J.; Xie, Z.; Ma, J.; Liu, W. J; Wang, D.; Xu, W.; Holmes, E. C; Gao, G. F; Wu, G.; Chen, W.; Shi, W.; Tan, W. Genomic characterisation and epidemiology of 2019 novel

coronavirus: implications for virus origins and receptor binding. *Lancet* **2020**, *395*, 565−574.
(3) Gowthaman, R.; Guest, J. D.; Yin, R.; Adolf-Bryfogle, J.; Schief, W. R.; Pierce, B. G. CoV3D: a database of high resolution coronavirus protein structures. *Nucleic Acids Res.* **2021**, *49*, D282−D287.
(4) Alsulami, A. F.; Thomas, S. E.; Jamasb, A. R.; Beaudoin, C. A.; Moghul, I.; Bannerman, B.; Copoiu, L.; Vedithi, S. C.; Torres, P.; Blundell, T. L. SARS-CoV-2 3D database: understanding the coronavirus proteome and evaluating possible drug targets. Brief. *Bioinformatics* **2021**, *22*, 769−780.
(5) Kumar, A.; Shasany, A. K. SARS-CoV-2 Pocketome: Severe Acute Respiratory Syndrome Coronavirus 2, Pockets Identification for Antiviral & Antimicrobial Phytomolecules and Drug Repurposing. *ChemRxiv Preprint*, 2020. DOI: 10.26434/chemrxiv.12522515.v1.
(6) Yazdani, S.; De Maio, N.; Ding, Y.; Shahani, V.; Goldman, N.; Schapira, M. Genetic Variability of the SARS-CoV-2 Pocketome. *J. Proteome Res.* **2021**, *20*, 4212−4215.
(7) Ovek, D.; Taweel, A.; Abali, Z.; Tezsezen, E.; Koroglu, Y. E.; Tsai, C. J.; Nussinov, R.; Keskin, O.; Gursoy, A. Web interface for 3D visualization and analysis of SARS-CoV-2-human mimicry and interactions. *Bioinformatics* **2021**, btab799.
(8) Kong, R.; Yang, G.; Xue, R.; Liu, M.; Wang, F.; Hu, J.; Guo, X.; Chang, S. COVID-19 Docking Server: a meta server for docking small molecules, peptides and antibodies against potential targets of COVID-19. *Bioinformatics* **2020**, *36*, 5109−5111.
(9) Gervasoni, S.; Vistoli, G.; Talarico, C.; Manelfi, C.; Beccari, A. R.; Studer, G.; Tauriello, G.; Waterhouse, A. M.; Schwede, T.; Pedretti, A. A Comprehensive Mapping of the Druggable Cavities within the SARS-CoV-2 Therapeutically Relevant Proteins by Combining Pocket and Docking Searches as Implemented in Pockets 2.0. *Int. J. Mol. Sci.* **2020**, *21*, 5152.
(10) Mohebbi, A.; Askari, F. S.; Sammak, A. S.; Ebrahimi, M.; Najafimemar, Z. Druggability of cavity pockets within SARS-CoV-2 spike glycoprotein and pharmacophore-based drug discovery. *Future Virol* **2021**, *16*, 389.
(11) Ugur Marion, I.; Marion, A. Molecular Modelling Reveals Eight Novel Druggable Binding Sites in SARS-CoV-2's Spike Protein. *ChemRxiv Preprint*, 2020. DOI: 10.26434/chemrxiv.13292768.v.
(12) Toelzer, C.; Gupta, K.; Yadav, S.; Borucu, U.; Davidson, A. D.; Kavanagh Williamson, M.; Shoemark, D. K.; Garzoni, F.; Staufer, O.; Milligan, R.; Capin, J.; Mulholland, A. J.; Spatz, J.; Fitzgerald, D.; Berger, I.; Schaffitzel, C. Free fatty acid binding pocket in the locked structure of SARS-CoV-2 spike protein. *Science* **2020**, *370*, 725−730.
(13) Sztain, T.; Amaro, R.; McCammon, J. A. Elucidation of cryptic and allosteric pockets within the SARS-CoV-2 protease. *bioRxiv Preprint*, 2020. DOI: 10.1101/2020.07.23.218784.
(14) Tang, J.; Tsigelny, I.; Greenberg, J.; Miller, M.; Kouznetsova, V. Potential SARS-CoV-2 Nonstructural Protein 15 (NSP15) Inhibitors: Repurposing FDA-Approved Drugs. *ChemRxiv Preprint*, 2021 . DOI: 10.26434/chemrxiv-2021-ht9sx.
(15) Zuzic, L.; Samsudin, F.; Shivgan, A. T.; Raghuvamsi, P. V.; Marzinek, J. K.; Boags, A.; Pedebos, C.; Tulsian, N. K.; Warwicker, J.; MacAry, P.; Crispin, P.;Khalid, S.; Anand, G. S.; Bond, P. J. Uncovering cryptic pockets in the SARS-CoV-2 spike glycoprotein. *BioRxiv Preprint*, 2021 DOI: 10.1101/2021.05.05.442536.
(16) Farias, A. B.; Candiotto, G.; Siragusa, L.; Goracci, L.; Cruciani, G.; Oliveira, E. R. A.; Horta, B. A. C. Targeting Nsp9 as an anti-SARS-CoV-2 strategy. *New J. Chem.* **2021**, *45*, 522−525.
(17) Siragusa, L.; Cross, S.; Baroni, M.; Goracci, L.; Cruciani, G. BioGPS: navigating biological space to predict polypharmacology, off-targeting, and selectivity. *Proteins* **2015**, *83*, 517−532.
(18) Baroni, M.; Cruciani, G.; Sciabola, S.; Perruccio, F.; Mason, J. S. A common reference framework for analyzing/comparing proteins and ligands. Fingerprints for Ligands and Proteins (FLAP): theory and application. *J. Chem. Inf. Model* **2007**, *47*, 279−294.
(19) Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J. D.; Higgins, D. G. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539.
(20) *NCBI National Center for Biotechnology Information SARS-CoV-2 structure page*. https://www.ncbi.nlm.nih.gov/Structure/SARS-CoV-2.html (accessed 2021−11−25).
(21) Zhang, Y.; Skolnick, J. TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res.* **2005**, *33*, 2302−2309.
(22) Henikoff, S.; Henikoff, J. G. Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. U.S.A* **1992**, *89*, 10915−10919.
(23) Rose, Y.; Duarte, J. M.; Lowe, R.; Segura, J.; Bi, C.; Bhikadiya, C.; Chen, L.; Rose, A. S.; Bittrich, S.; Burley, S. K.; Westbrook, J. D. RCSB Protein Data Bank: Architectural Advances Towards Integrated Searching and Efficient Access to Macromolecular Structure Data from the PDB Archive. *J. Mol. Biol.* **2021**, *433*, 166704.
(24) Schrödinger, L.; DeLano, W. *PyMOL*, 2020. http://www.pymol.org/pymol (accessed June 2022).
(25) Di Sarno, V.; Lauro, G.; Musella, S.; Ciaglia, T.; Vestuto, V.; Sala, M.; Scala, M. C.; Smaldone, G.; Di Matteo, F.; Novi, S.; Tecce, M. F.; Moltedo, O.; Bifulco, G.; Campiglia, P.; Gomez-Monterrey, I. M.; Snoeck, R.; Andrei, G.; Ostacolo, C.; Bertamino, A. Identification of a dual acting SARS-CoV-2 proteases inhibitor through in silico design and step-by-step biological characterization. *Eur. J. Med. Chem.* **2021**, *226*, 113863.
(26) *European Centre for Disease Prevention and Control SARS-CoV-2 variants of concern page*. https://www.ecdc.europa.eu/en/covid-19/variants-concern (accessed 2022−04−13).
(27) Winkler, E. S.; Chen, R. E.; Alam, F.; Yildiz, S.; Case, J. B.; Uccellini, M. B.; Holtzman, M. J.; Garcia-Sastre, A.; Schotsaert, M.; Diamond, M. S. SARS-CoV-2 Causes Lung Infection without Severe Disease in Human ACE2 Knock-In Mice. *J. Virol* **2022**, *96*, e0151121.
(28) Ma, C.; Hu, Y.; Townsend, J. A.; Lagarias, P.; Marty, M. T.; Kolocouris, A.; Wang, J. Ebselen, Disulfiram, Carmofur, PX-12, Tideglusib, and Shikonin Are Nonspecific Promiscuous SARS-CoV-2 Main Protease Inhibitors. *ACS Pharmacol. Transl. Sci.* **2020**, *3*, 1265−1277.