Thirteen Potential TMPRSS2 Inhibitors

Subjects: Biochemistry & Molecular Biology

Contributor: Xiaoqiang Huang

We identified a set of 13 approved or clinically investigational drugs with positively charged guanidinobenzoyl and/or aminidinobenzoyl groups, including the experimentally verified TMPRSS2 inhibitors Camostat and Nafamostat. Molecular docking suggested that the guanidinobenzoyl or aminidinobenzoyl group in all the drugs could form putative salt bridge interactions with the side-chain carboxyl group of Asp435 located in the S1 pocket of TMPRSS2. Molecular dynamics simulations further revealed the high stability of the putative salt bridge interactions over long-time simulations. These results suggest that the proposed compounds, in addition to Camostat and Nafamostat, could be effective TMPRSS2 inhibitors for COVID-19 treatment by occupying the S1 pocket with the hallmark positively charged groups.

Keywords: SARS-CoV-2 ; COVID-19 ; TMPRSS2 ; drug ; docking ; molecular dynamics

1. Introduction

TMPRSS2 is a viable anti-SARS-CoV-2 host protein target for the following four reasons. First, it is not mutation-prone. Second, TMPRSS2 is used by other coronaviruses (e.g., SARS-CoV and MERS-CoV) and by influenza A viruses for the activation of surface glycoproteins; therefore, a specific TMPRSS2 inhibitor may treat a whole class of diseases caused by different pathogens[1], including SARS-CoV-2 variants, during this pandemic and in coming years. Third, TMPRSS2 does not appear to play an essential role in any organ, as other proteases may provide a degree of redundancy; thus, TMPRSS2 inhibition may have few on-target side effects. In TMPRSS2-knockout mice, TMPRSS2 appeared dispensable for normal development, growth, and organ function[2]. Fourth, since TMPRSS2 is a member of the serine protease family for which many inhibitors are available, finding a suitable drug to target it should be feasible.

Soon after the outbreak of COVID-19, Hoffmann et al. demonstrated that SARS-CoV-2's dependence on TMPRSS2 for cell entry can be blocked by a clinically proven protease inhibitor, Camostat[3]. A metabolite of Camostat, 4-(4-guanidinobenzoyloxy)phenylacetic acid (GBPA, known as FOY-251) also inhibited TMPRSS2 but with reduced efficiency compared to Camostat[4]. Later, numerous research groups proved that Nafamostat has about 10-fold greater potency than Camostat for preventing SARS-CoV-2 infection through in vitro and in vivo studies[5][6][7][8]. Hempel et al. carried out a systematic analysis to compare Nafamostat, Camostat, and GBPA to determine how these compounds could effectively inhibit TMPRSS2[8]. Their computational studies suggested that the three compounds contain the positively charged guanidinobenzoyl and/or aminidinobenzoyl moiety, which can form stable salt bridge interactions with the negatively charged aspartic acid Asp435 in the S1 pocket of TMPRSS2, occupying the binding site and leading to the inhibition.

Inspired by these studies, we hypothesized that some other guanidinobenzoyl- or aminidinobenzoyl-containing drugs may act as TMPRSS2 inhibitors. We identified from DrugBank a narrowed list of 13 compounds (three FDA-approved drugs and 10 investigational drugs) that contain guanidinobenzoyl or aminidinobenzoyl groups. We computationally evaluated their potency for inhibiting TMPRSS2 through molecular docking, molecular dynamics (MD) simulation, and post-MD analysis.

2. TMPRSS2 Sequence and Structural Model

TMPRSS2 (UniProt ID: O15393) has two isoforms produced by alternative splicing. Isoform 1 (O15393-1) has been chosen as the canonical sequence and was used in this study. The full-length TMPRSS2 isoform 1 is composed of 492 amino acids, containing two topological domains (amino acids 1-84, cytoplasmic domain and 106-492, extracellular domain), one transmembrane domain (amino acids 85-105), and one trypsin-like catalytic domain (amino acids 256-492) (**Figure 1**a). Isoform 2 does not have a catalytic domain and was excluded from the analysis.

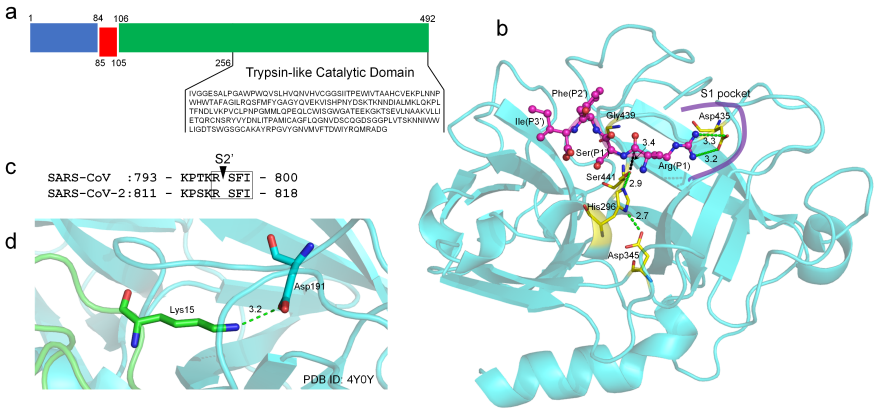

Figure 1. Sequence, topology, structural model, and function of TMPRSS2. (**a**) Sequence and domain topology of TMPRSS2 (amino acids 1-84: cytoplasmic domain, 85-105: transmembrane domain, 106-492: extracellular domain, and 256-492: catalytic domain). (**b**) C-I-TASSER model of the TMPRSS2 catalytic domain (shown in cyan cartoon). The conserved, catalytic triad (Ser441, His296, and Asp345); oxyanion holes (mainchain amide groups of Ser441 and Gly439); and the conserved aspartic acid (Asp435) in the S1 pocket are shown in yellow sticks. A tetrapeptide (RSFI, shown in a magenta ball-and-stick model) extracted from the SARS-CoV-2 S2′ cleavage site is docked into the binding site using HPEPDOCK. Hydrogen bonds and salt bridge interactions are illustrated in dashed green lines. The distance between the atom Oγ of Ser441 and the carbonyl C atom of P1 arginine is illustrated in dashed black lines. The distances are shown around the lines (unit: Å). (**c**) The S2′ cleavage sites of SARS-CoV and SARS-CoV-2. (**d**) An example of the trypsin inhibitor interaction in the S1 pocket.

There is no experimental structure available for TMPRSS2 and its domains. We used a deep-learning contact-guided protein structure assembly approach, C-I-TASSER[9], to model the structure of the catalytic domain of TMPRSS2 (**Figure 1**b). The model had a C-score[10] of 0.45, which corresponded to the estimated TM-score[11] of 0.89. Here, the C-score was a confidence score for estimating the global quality of predicted models by C-I-TASSER; based on large-scale benchmark tests, C-I-TASSER models with a C-score > −2.5 correspond to a correct fold with a TM-score > 0.5. The model also had high local structure quality, with a MolProbity[12] score of 0.91, which ranked at the 100th percentile. This puts the structure models amongst the best structures of a comparable solution by comparison with a representative set of experimental structures collected from the Protein Data Bank (PDB). Our model exhibited a very high structural similarity (e.g., TM-score > 0.95) to the reported models generated by homology modeling approaches. Compared with the homology models built on a single template, our C-I-TASSER model was constructed by considering the consensus of multiple templates (PDB IDs: 7meq, 3w94, 4dgj, and 6eso) and, thus, avoided the modeling bias toward a single experimental structure.

A typical feature of trypsin or a trypsin-like protease is the deeply buried negatively charged aspartic acid in the S1 pocket, which specifically recognizes the positively charged arginine or lysine at the P1 site of a protein substrate. In TMPRSS2, such an aspartic acid residue is Asp435 (**Figure 1**b); there is no other aspartic acid residue in the S1 pocket. Besides, the catalytic elements of TMPRSS2 include a well-established catalytic triad (Ser441–His296–Asp345) indicated by the hydrogen-bonding network and two oxyanion holes (i.e., the main-chain amide groups of Ser441 and Gly439) (**Figure 1**b). The ideal configuration of these catalytic elements also suggested a good quality of the TMPRSS2 model. TMPRSS2 can prime the spike proteins at the S2′ cleavage site for SARS-CoV and SARS-CoV-2 (**Figure 1**c). We used a protein–peptide docking tool HPEPDOCK[13] to predict the binding mode of the P1-P1′-P2′-P3′ tetrapeptide (RSFI, **Figure 1**c) extracted from SARS-CoV/SARS-CoV-2 bound to the binding pocket of TMPRSS2. In the top one pose, the P1 arginine was predicted to form bidentate salt bridge interactions with Asp435, while the Oγ atom of the catalytic Ser441 is 3.4 Å to the carbonyl atom of the P1 arginine (**Figure 1**b) within the van der Waals contact distance (3.5 Å) that is critical for the subsequent bond-breaking catalysis. The predicted binding interactions mimic those made between trypsin and its natural substrate, in which the lysine side-chain amino group interacts with the conserved aspartic acid in the S1 pocket (**Figure 1**d).

3. Guanidinobenzoyl- or Aminidinobenzoyl-Containing Drugs

Previous studies revealed that a positively charged group that mimics arginine or lysine in a natural substrate of trypsin was important for a drug acting as an inhibitor to the trypsin-like TMPRSS2[8][14][15]. Specifically, the guanidinobenzoyl and/or aminidinobenzoyl group in Camostat or Nafamostat could form stable binding interactions with the conserved Asp435 in the S1 pocket in TMPRSS2 and lead to inhibition. Compared with Camostat and Nafamostat, Gabexate, which

contains an arginine-like side-chain, showed only a weak inhibitory potency[5][6]. Considering that a drug's rigidity is crucial for high-affinity binding due to its low conformational entropy effect[6][16][17], we preferentially considered drugs with guanidinobenzoyl or aminidinobenzoyl rather than an arginine- or lysine-like side-chain group; the former two groups have far fewer degrees of freedom and are, hence, more rigid. We searched DrugBank for FDA-approved or investigational drugs that contain guanidinobenzoyl or guanidinobenzoyl and obtained a small library of 13 drugs (**Figure 2**).

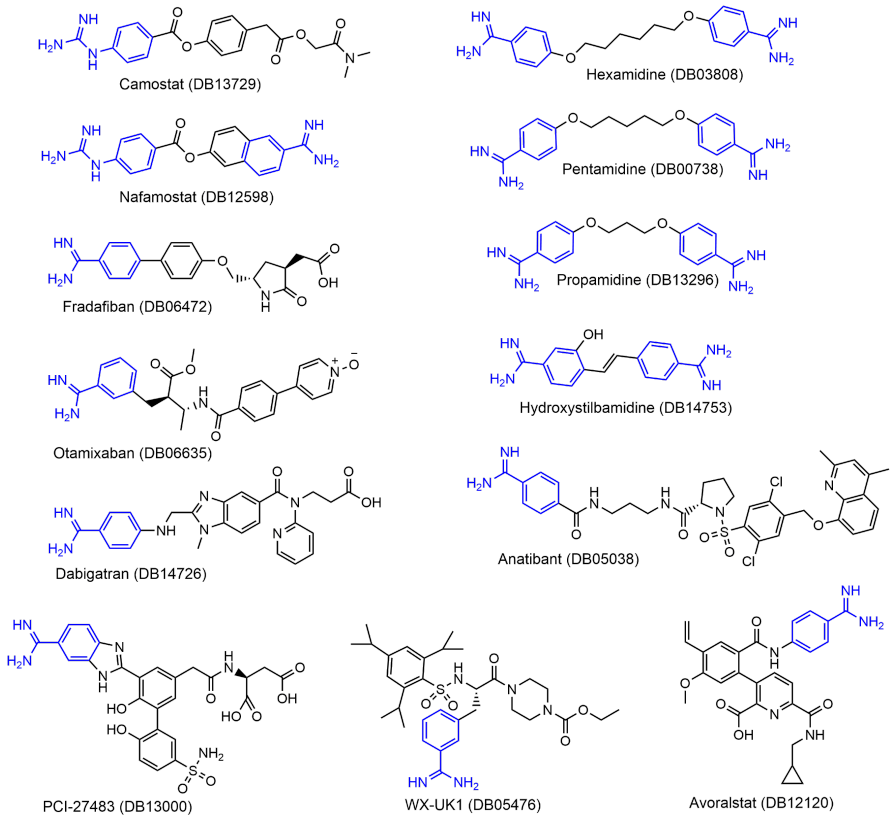

Figure 2. Molecular formulas of the 13 drugs investigated in this work. The guanidinobenzoyl and aminidinobenzoyl groups are highlighted in blue.

Among the 13 drugs, Camostat is a serine protease inhibitor approved in Japan for the treatment of chronic pancreatitis and postoperative reflux esophagitis. Nafamostat is a synthetic serine protease inhibitor approved as an anticoagulant therapy for patients undergoing continuous renal replacement therapy due to acute kidney injury and used for the treatment of acute pancreatitis in Japan. Camostat and Nafamostat were demonstrated to be effective TMPRSS2 inhibitors and are in clinical trials for COVID-19 treatment. It was speculated that Camostat and Nafamostat are covalent TMPRSS2 inhibitors, because their ester bonds can be cleaved by serine proteases; this speculation was supported by their low nanomolar-level inhibitory behaviors. In contrast, the other drugs do not contain a cleavable ester bond adjacent to the guanidinobenzoyl or aminidinobenzoyl group and may only function as a noncovalent TMPRSS2 inhibitor. Except for Camostat and Nafamostat, none of the other drugs have been clinically investigated for COVID-19 treatment.

4. Molecular Docking Suggests Salt Bridge Interactions between Guanidinobenzoyl or Aminidinobenzoyl and Asp435

Each of the 13 drugs was docked into the putative binding pocket of TMPRSS2 using LeDock[18] and the poses were rescored using our physics- and knowledge-based energy function EvoEF2[19], because the default LeDock score function may tolerate severe intermolecular steric clashes in a few top-ranked poses. The molecular docking results indicated that all the drugs had at least one pose that could form salt bridge interactions with the negatively charged Asp435. The top one pose with the lowest EvoEF2 score is shown in **Figure 3** for each of the drugs. All the drugs are well docked into the binding pocket of TMPRSS2, with their guanidinobenzoyl or aminidinobenzoyl groups aligning in the S1 pocket and forming salt bridge interactions with the side-chain carboxyl group of Asp435, supporting favorable binding scores.

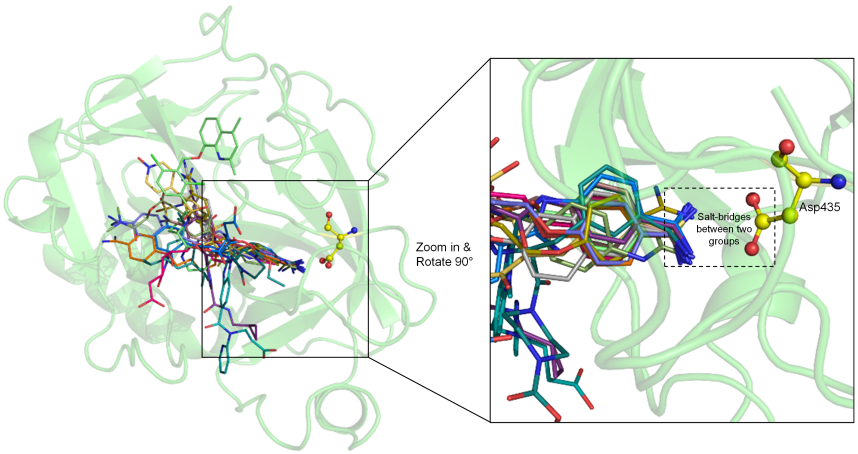

Figure 3. Superposition and comparison of the ligand poses with the lowest EvoEF2 scores for all 13 drugs. TMPRSS2 is shown in the green cartoon model, with residue Asp435 depicted in the yellow ball-and-stick model. The zoom-in inset shows that guanidinobenzoyl or aminidinobenzoyl can form salt bridge interactions with the Asp435 carboxyl group (shown in the dashed box).

5. MD Simulations Reveal High Stability of the Putative Salt Bridge Interactions

We carried out MD simulations to examine the binding stability between TMPRSS2 and the drugs. Before MD, the top ten poses (if they existed) ranked by EvoEF2 for each drug were parameterized using the ACPYPE[20] program with the AM1-BCC[21] charge model. The number of poses that can be successfully applied to MD for the 13 drugs were 10 (DB00738), 10 (DB03808), 10 (DB05038), 9 (DB05476), 8 (DB06472), 10 (DB06635), 7 (DB12120), 6 (DB12598), 5 (DB13000), 10 (DB13296), 10 (DB13729), 10 (DB14726), and 4 (DB14753), respectively. TMPRSS2 in complex with each suitable drug pose was subjected to a long-time (100 ns) MD simulation using GROMACS v2020.4[22].

According to docking models, the guanidinobenzoyl or aminidinobenzoyl groups were docked into the deep S1 pocket and formed salt bridge interactions with Asp435. We examined the stability of the putative salt bridge interactions by measuring the minimum distance between the positively charged guanidinobenzoyl or aminidinobenzoyl group and the negatively charged carboxyl group of Asp435 (denoted as d_{ON}^{min}); only the distances between the nitrogen and oxygen atoms were calculated. All the drugs had at least one pose with a mean and median d_{ON}^{min} fluctuating around 2.8 Å, an ideal salt bridge distance, with small deviations (**Figure 4**). Therefore, a long-time MD simulation indicated the high stability of the putative salt bridge interactions between the guanidinobenzoyl or aminidinobenzoyl group and Asp435, which should be important for TMPRSS2 inhibition.

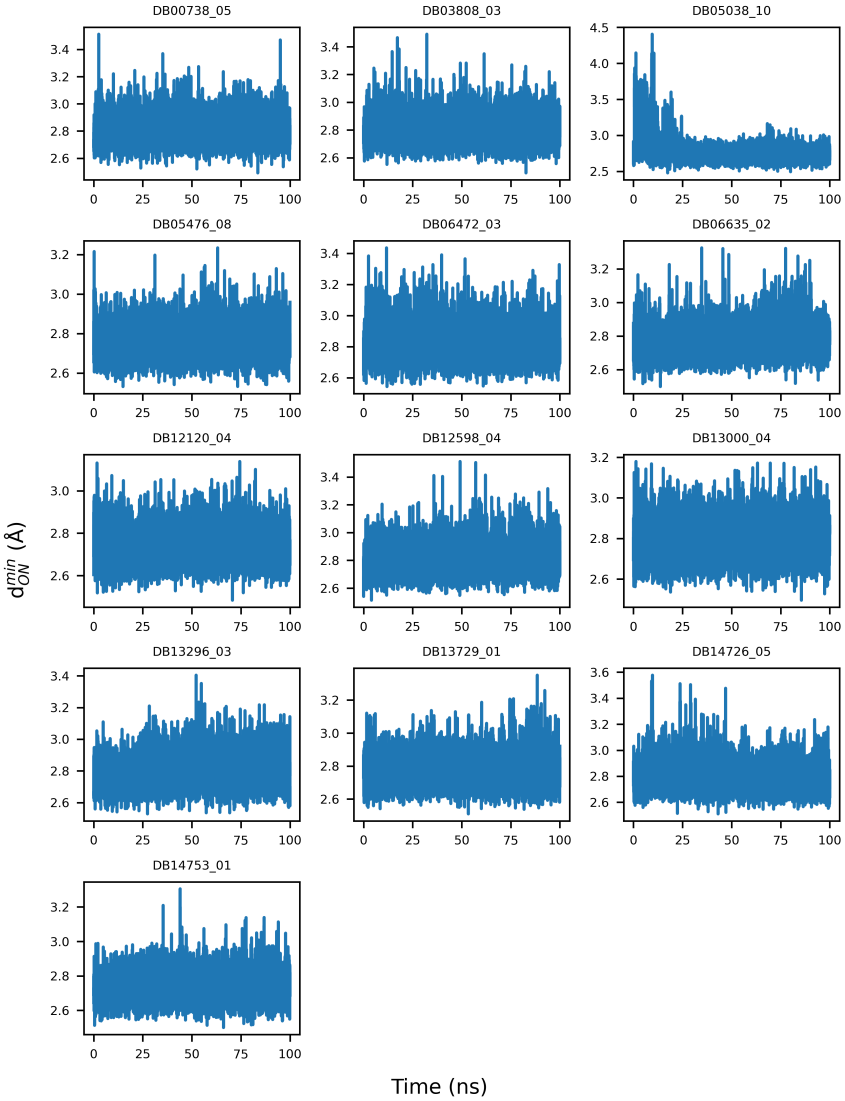

Figure 4. Example illustration of the variations of the minimum distance between guanidinobenzoyl or aminidinobenzoyl and the Asp435 carboxyl (d_{ON}^{min}) over a 100-ns simulation time.

6. Conclusions

Building on the recent finding that the positively charged groups in Camostat and Nafamostat play a critical role in inhibiting TMPRSS2 by stable binding with the conserved aspartic acid Asp435 in the S1 pocket of TMPRSS2, we identified a narrowed set of 13 compounds (three FDA-approved and 10 investigational drugs) with positively charged guanidinobenzoyl or aminidinobenzoyl groups and computationally assessed their potency for inhibiting TMPRSS2. This work differed from virtual screening studies that focus on identifying TMPRSS2 inhibitors from huge drug databases. Usually, a virtual screening study suggests a long list of candidates for experimental tests but, finally, comes up with few positive hits; instead, here, we tried to evaluate and repurpose only a few very promising candidates. The molecular docking studies showed that all the 13 drugs indeed utilized guanidinobenzoyl or aminidinobenzoyl to form favorable salt bridge interactions with the Asp435 carboxyl, and a series of long-time (100 ns) MD simulations revealed the high stability of the salt bridge interactions between each drug and TMPRSS2, although each whole ligand may undergo large conformational changes. Collectively, the computational data supported these drugs as potential TMPRSS2 inhibitors for treating COVID-19.

References

1. Rolando Cannalire; Irina Stefanelli; Carmen Cerchia; Andrea R. Beccari; Sveva Pelliccia; Vincenzo Summa; SARS-Co V-2 Entry Inhibitors: Small Molecules and Peptides Targeting Virus or Host Cells. *International Journal of Molecular Sci ences* **2020**, *21*, 5707, 10.3390/ijms21165707.
2. Konrad H. Stopsack; Lorelei A. Mucci; Emmanuel S. Antonarakis; Peter S. Nelson; Philip W. Kantoff; TMPRSS2 and C OVID-19: Serendipity or Opportunity for Intervention?. *Cancer Discovery* **2020**, *10*, 779-782, 10.1158/2159-8290.cd-20

-0451.

3. Markus Hoffmann; Hannah Kleine-Weber; Simon Schroeder; Nadine Krüger; Tanja Herrler; Sandra Erichsen; Tobias S. Schiergens; Georg Herrler; Nai-Huei Wu; Andreas Nitsche; et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **2020**, *181*, 271-280.e8, 10.1016/j.cell.2020.02.052.
4. Markus Hoffmann; Heike Hofmann-Winkler; Joan C. Smith; Nadine Krüger; Prerna Arora; Lambert K. Sørensen; Ole S. Søgaard; Jørgen Bo Hasselstrøm; Michael Winkler; Tim Hempel; et al. Camostat mesylate inhibits SARS-CoV-2 activation by TMPRSS2-related proteases and its metabolite GBPA exerts antiviral activity. *EBioMedicine* **2021**, *65*, 103255, 10.1016/j.ebiom.2021.103255.
5. Markus Hoffmann; Simon Schroeder; Hannah Kleine-Weber; Marcel A. Müller; Christian Drosten; Stefan Pöhlmann; Nafamostat Mesylate Blocks Activation of SARS-CoV-2: New Treatment Option for COVID-19. *Antimicrobial Agents and Chemotherapy* **2020**, *64*, e00754-20, 10.1128/aac.00754-20.
6. Jonathan H. Shrimp; Stephen Kales; Philip E. Sanderson; Anton Simeonov; Min Shen; Matthew D. Hall; An Enzymatic TMPRSS2 Assay for Assessment of Clinical Candidates and Discovery of Inhibitors as Potential Treatment of COVID-19. *ACS Pharmacology & Translational Science* **2020**, *3*, 997-1007, 10.1021/acsptsci.0c00106.
7. Mizuki Yamamoto; Maki Kiso; Yuko Sakai-Tagawa; Kiyoko Iwatsuki-Horimoto; Masaki Imai; Makoto Takeda; Noriko Kinoshita; Norio Ohmagari; Jin Gohda; Kentaro Semba; et al. The Anticoagulant Nafamostat Potently Inhibits SARS-CoV-2 S Protein-Mediated Fusion in a Cell Fusion Assay System and Viral Infection In Vitro in a Cell-Type-Dependent Manner. *Viruses* **2020**, *12*, 629, 10.3390/v12060629.
8. Tim Hempel; Lluís Raich; Simon Olsson; Nurit P. Azouz; Andrea M. Klingler; Markus Hoffmann; Stefan Pöhlmann; Marc E. Rothenberg; Frank Noé; Molecular mechanism of inhibiting the SARS-CoV-2 cell entry facilitator TMPRSS2 with camostat and nafamostat. *Chemical Science* **2020**, *12*, 983-992, 10.1039/d0sc05064d.
9. Wei Zheng; Yang Li; Chengxin Zhang; Robin Pearce; S. M. Mortuza; Yang Zhang; Deep-learning contact-map guided protein structure prediction in CASP13. *Proteins: Structure, Function, and Bioinformatics* **2019**, *87*, 1149-1164, 10.1002/prot.25792.
10. Yang Zhang; I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **2008**, *9*, 1-8, 10.1186/1471-2105-9-40.
11. Yang Zhang; Jeffrey Skolnick; Scoring function for automated assessment of protein structure template quality. *Proteins: Structure, Function, and Bioinformatics* **2004**, *57*, 702-710, 10.1002/prot.20264.
12. Vincent Chen; W. Bryan Arendall; Jeffrey J. Headd; Daniel Keedy; Robert M. Immormino; Gary J. Kapral; Laura W. Murray; Jane Richardson; David C. Richardson; MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D Biological Crystallography* **2009**, *66*, 12-21, 10.1107/s0907444909042073.
13. Pei Zhou; Bowen Jin; Hao Li; Sheng-You Huang; HPEPDOCK: a web server for blind peptide–protein docking based on a hierarchical algorithm. *Nucleic Acids Research* **2018**, *46*, W443-W450, 10.1093/nar/gky357.
14. David J. Huggins; Structural analysis of experimental drugs binding to the SARS-CoV-2 target TMPRSS2. *Journal of Molecular Graphics and Modelling* **2020**, *100*, 107710-107710, 10.1016/j.jmgm.2020.107710.
15. Diego E. Escalante; David M. Ferguson; Structural modeling and analysis of the SARS-CoV-2 cell entry inhibitor camostat bound to the trypsin-like protease TMPRSS2. *Medicinal Chemistry Research* **2021**, *30*, 399-409, 10.1007/s00044-021-02708-7.
16. Alastair D. G. Lawson; Malcolm MacCoss; Jag P. Heer; Importance of Rigidity in Designing Small Molecule Drugs To Tackle Protein–Protein Interactions (PPIs) through Stabilization of Desired Conformers. *Journal of Medicinal Chemistry* **2017**, *61*, 4283-4289, 10.1021/acs.jmedchem.7b01120.
17. Christophe Verlinde; Wim Gj Hol; Structure-based drug design: progress, results and challenges. *Structure* **1994**, *2*, 577-587, 10.1016/s0969-2126(00)00060-5.
18. Na Zhang; Hongtao Zhao; Enriching screening libraries with bioactive fragment space. *Bioorganic & Medicinal Chemistry Letters* **2016**, *26*, 3594-3597, 10.1016/j.bmcl.2016.06.013.
19. Xiaoqiang Huang; Robin Pearce; Yang Zhang; EvoEF2: accurate and fast energy function for computational protein design. *Bioinformatics* **2019**, *36*, 1135-1142, 10.1093/bioinformatics/btz740.
20. Alan W Sousa Da Silva; Wim F Vranken; ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Research Notes* **2011**, *5*, 367-367, 10.1186/1756-0500-5-367.
21. Araz Jakalian; Bruce L. Bush; David B. Jack; Christopher I. Bayly; Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method. *Journal of Computational Chemistry* **2000**, *21*, 132-146, 10.1002/(sici)1096-987x(20000130)21:2<132::aid-jcc5>3.0.co;2-p.

22. Mark James Abraham; Teemu Murtola; Roland Schulz; Szilárd Páll; Jeremy Smith; Berk Hess; Erik Lindahl; GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1-2*, 19-25, 10.1016/j.softx.2015.06.001.
