Structure-based phylogeny identifies avoralstat as a TMPRSS2 inhibitor that prevents SARS-CoV-2 infection in mice

Young Joo Sun,[1] Gabriel Velez,[1,2] Dylan E. Parsons,[1,3] Kun Li,[4] Miguel E. Ortiz,[4] Shaunik Sharma,[4] Paul B. McCray Jr.,[4,5] Alexander G. Bassuk,[4,6,7] and Vinit B. Mahajan[1,8]

[1]Molecular Surgery Lab, Byers Eye Institute, Department of Ophthalmology, Stanford University, Palo Alto, California, USA. [2]Medical Scientist Training Program, University of Iowa, Iowa City, Iowa, USA. [3]Stanford ChEM-H Medicinal Chemistry Knowledge Center, Stanford University, Palo Alto, California, USA. [4]Department of Pediatrics, [5]Department of Microbiology and Immunology, [6]Department of Neurology, and [7]Iowa Neuroscience Institute, University of Iowa, Iowa City, Iowa, USA. [8]Veterans Affairs Palo Alto Health Care System, Palo Alto, California, USA.

Drugs targeting host proteins can act prophylactically to reduce viral burden early in disease and limit morbidity, even with antivirals and vaccination. Transmembrane serine protease 2 (TMPRSS2) is a human protease required for SARS coronavirus 2 (SARS-CoV-2) viral entry and may represent such a target. We hypothesized that drugs selected from proteins related by their tertiary structure, rather than their primary structure, were likely to interact with TMPRSS2. We created a structure-based phylogenetic computational tool named 3DPhyloFold to systematically identify structurally similar serine proteases with known therapeutic inhibitors and demonstrated effective inhibition of SARS-CoV-2 infection in vitro and in vivo. Several candidate compounds, avoralstat, PCI-27483, antipain, and soybean trypsin inhibitor, inhibited TMPRSS2 in biochemical and cell infection assays. Avoralstat, a clinically tested kallikrein-related B1 inhibitor, inhibited SARS-CoV-2 entry and replication in human airway epithelial cells. In an in vivo proof of principle, avoralstat significantly reduced lung tissue titers and mitigated weight loss when administered prophylactically to mice susceptible to SARS-CoV-2, indicating its potential to be repositioned for coronavirus disease 2019 (COVID-19) prophylaxis in humans.

Introduction

Coronavirus disease 2019 (COVID-19), caused by SARS coronavirus 2 (SARS-CoV-2), has spread globally, and prophylactic and early-stage therapies are needed for high-risk populations. Even with vaccines, adjunctive therapies that mitigate viral entry or replication may attenuate disease severity and reduce viral spread by asymptomatic and early-stage patients. In response to the urgent need for therapeutics, there is investigation into repositioning existing drugs toward viral proteins (e.g., Remdesivir). An alternative strategy is to target human proteins utilized by viruses with small molecules. This approach can work synergistically with vaccination and may be especially important for individuals where vaccination is contraindicated or deferred, to prevent viral transmission that may occur after vaccination, for front-line workers exposed to repeated high viral load, in countries where sophisticated vaccine delivery and storage is unavailable, and potentially for protection from viral mutation and other viruses using similar host mechanisms.

Transmembrane serine protease 2 (TMPRSS2) is a human serine protease that is a priming protease for the spike glycoprotein found on the surface of all coronaviruses (1, 2). TMPRSS2's S1-peptidase domain is required for SARS-CoV-2 entry into host epithelial cells in the upper- and lower-respiratory tract (3, 4), but it is not necessary for development or homeostasis in mice, making it an attractive drug target (5). It is yet to be determined whether TMPRSS2 inhibition mitigates SARS-CoV-2 infection in vivo. Camostat, a serine protease inhibitor originally developed for acute pancreatitis, inhibits TMPRSS2 in vitro and is in a clinical trial for the treatment of COVID-19 (ClinicalTrials.gov NCT04321096) (6, 7). However, camostat's plasma half-life is less than 1 minute, and its efficacy for COVID-19 is yet to be determined (4, 8, 9). Thus, identification of inhibitors targeting TMPRSS2 with improved pharmacokinetic properties remains important.

Conventional methods for identifying drug candidates typically employ high-throughput screening (HTS) or in silico screening using compound libraries previously tested in humans (10). In silico screening for TMPRSS2 has been challenging because there is no high-resolution TMPRSS2 molecular structure. Although HTS methods can rapidly screen thousands of compounds, there are certain limitations. HTS methods utilize only a few generalized experimental parameters with technical limitations, such as narrow dose range and experimental conditions, which may not account for the unique features of each compound. This can lead to false positives and negatives. Although false positives are filtered out in subsequent experiments, false negatives may overlook valuable compounds. Because HTS uses shotgun rather than hypothesis-driven approaches, it may be difficult to ascertain the mechanism of action, and this may slow the downstream development of candidate drugs into human therapies. Hypothesis-driven screening methods, utilizing protein structures and a limited num-
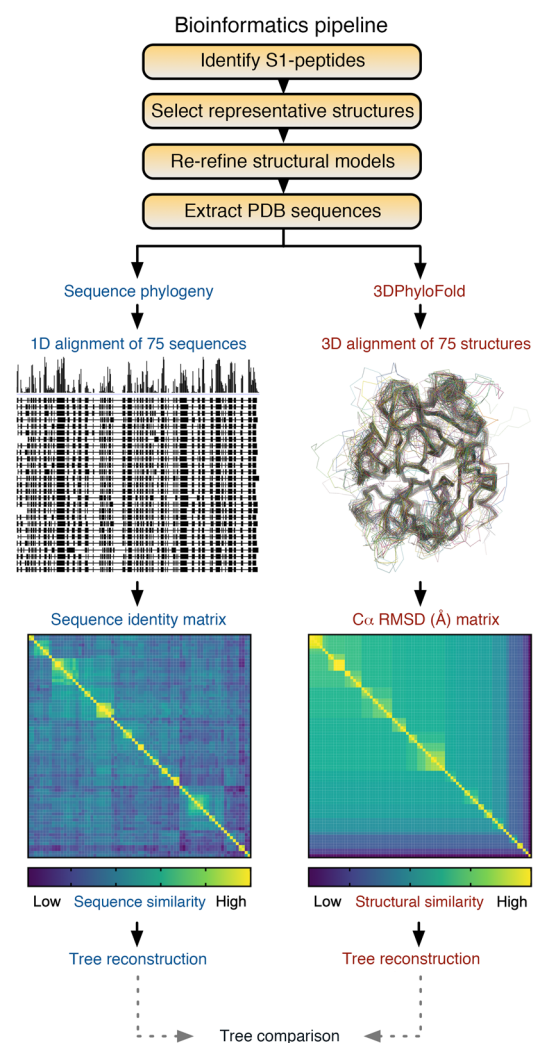

Figure 1. Bioinformatics workflow for sequence-based phylogenetic analysis and structure-based phylogenetic analysis. Conventional sequence-based phylogenetic analysis relies on alignment of the primary sequence. 3DPhyloFold calculates a structural dissimilarity matrix (SDM) based on the overlay of 75 representative mammalian S1-peptidase domains with high-resolution structures. The workflow can determine the differences and similarities between the sequence-based phylogeny and structure-based phylogeny of a domain family.

ber of compounds, remain a valuable and complementary strategy for drug repositioning. One approach to rational-based drug repositioning is to identify proteins with preexisting drugs that are similar to the target protein.

Results

In silico drug repositioning by 3DPhyloFold. To identify drug repositioning candidates, we created a computational/hypothesis-driven drug repurposing method called 3DPhyloFold that identifies structurally similar proteins to rationally select candidate inhibitors. A comprehensive phylogenetic analysis of 600 S1-peptidases sequentially related to the TMPRSS2 S1-peptidase domain (TMPRSS2-S1P) showed TMPRSS2-S1P clustered closely to canonical TMPRSS family members, like hepsin, as well as proteases outside of the TMPRSS subfamily: coagulation factor XI

and kallikrein-related B1 (KLKB1; Supplemental Figure 1, A and B, and Supplemental Table 1; supplemental material available online with this article; https://doi.org/10.1172/JCI147973DS1). TMPRSS2-S1P was closest to hepsin, and a homology-based model was generated (Supplemental Figure 1C). Next, 3DPhylo-Fold determined the 3D relationship of TMPRSS2-S1P to other S1-peptidase structures (Figure 1 and Supplemental Table 2). Using structural quality metrics (see Methods), 74 S1-peptidases and TMPRSS2-S1P were aligned by conventional sequence phylogenetic analysis. TMPRSS2-S1P clustered closely with KLKB1, factor XI, and complement factor I (CFAI), whereas the kallikrein- and trypsin-like clades clustered further away (Figure 2A). In 3DPhyloFold, pairwise structural comparisons of the representative tertiary structures were used to calculate a structural dissimilarity matrix (SDM) based on the root mean square deviation between protein alpha-carbons (Cα RMSD; Figure 1). A structure-based phylogenetic tree was then generated from the SDM (Figure 2B). Clustering of the structure-based tree was distinct from that of the sequence-based tree. Proteins close in the primary sequence analysis (e.g., CFAI and CTRB1) were much farther away in the 3DPhyloFold structure tree (Supplemental Figure 2G). Although distant in the sequence-based tree, the trypsin-like clade and factor VII moved much closer to TMPRSS2-S1P in the structure-based tree (Figure 2B). This suggested that, while divergent in sequence, TMPRSS2-S1P adopts a 3D fold closer to trypsin and factor VII. We prioritized the 6 S1-peptidases with the highest structural similarity to TMPRSS2-S1P: hepsin, acrosin, trypsin, factor VII, factor XI, and KLKB1.

Using these 6 proteases, we sought known small molecules and peptidomimetic inhibitors containing a guanidine or structurally related groups (see Methods) because S1-peptidases are inhibited by compounds containing a 4-amindinobenzylamide moiety. The 4-amindinobenzylamide moiety mimics the key specificity feature of S1-peptidase substrates: an arginine residue at the first N-terminal residue of the substrates' cleavage site (P1) that forms a strong interaction with an aspartate in the corresponding S1 subpocket of the peptidase (Figure 3A) (11). This search curated 90 experimental compounds and 4 small molecules previously tested in human clinical trials, which docked well to TMPRSS2-S1P (Figure 3, B and C, Supplemental Tables 3 and 4, and Supplemental Figure 3). In addition, 3DPhyloFold analysis revealed a natural trypsin-inhibiting protein based on the structure of porcine trypsin with soybean trypsin inhibitor (SBTI; PDBID 1AVW; Figure 3C). Since the porcine trypsin binding pocket was similar to that of TMPRSS2-S1P (~68% sequence identity), we modeled the TMPRSS2-S1P/SBTI complex and identified a conserved inhibitory motif (PYRIRF) with favorable docking interactions, suggesting SBTI might bind and inhibit TMPRSS2-S1P (Supplemental Figure 4 and Supplemental Tables 4 and 5).

Biochemical evaluation of 3DPhyloFold inhibitors. We focused on the inhibitory potential of human drugs available for repositioning, including avoralstat, PCI-27483, and antipain, along with SBTI. A biochemical inhibition assay using the extracellular region of purified recombinant TMPRSS2 (residues 106–492) was utilized to test inhibition by these compounds (7). The rank order potency against TMPRSS2 was avoralstat (IC_{50} = 2.73 ± 0.19 nM), SBTI (IC_{50} = 121 ± 4 nM), antipain (IC_{50} = 748 ± 63 nM), and PCI-

A Conventional sequence-based tree
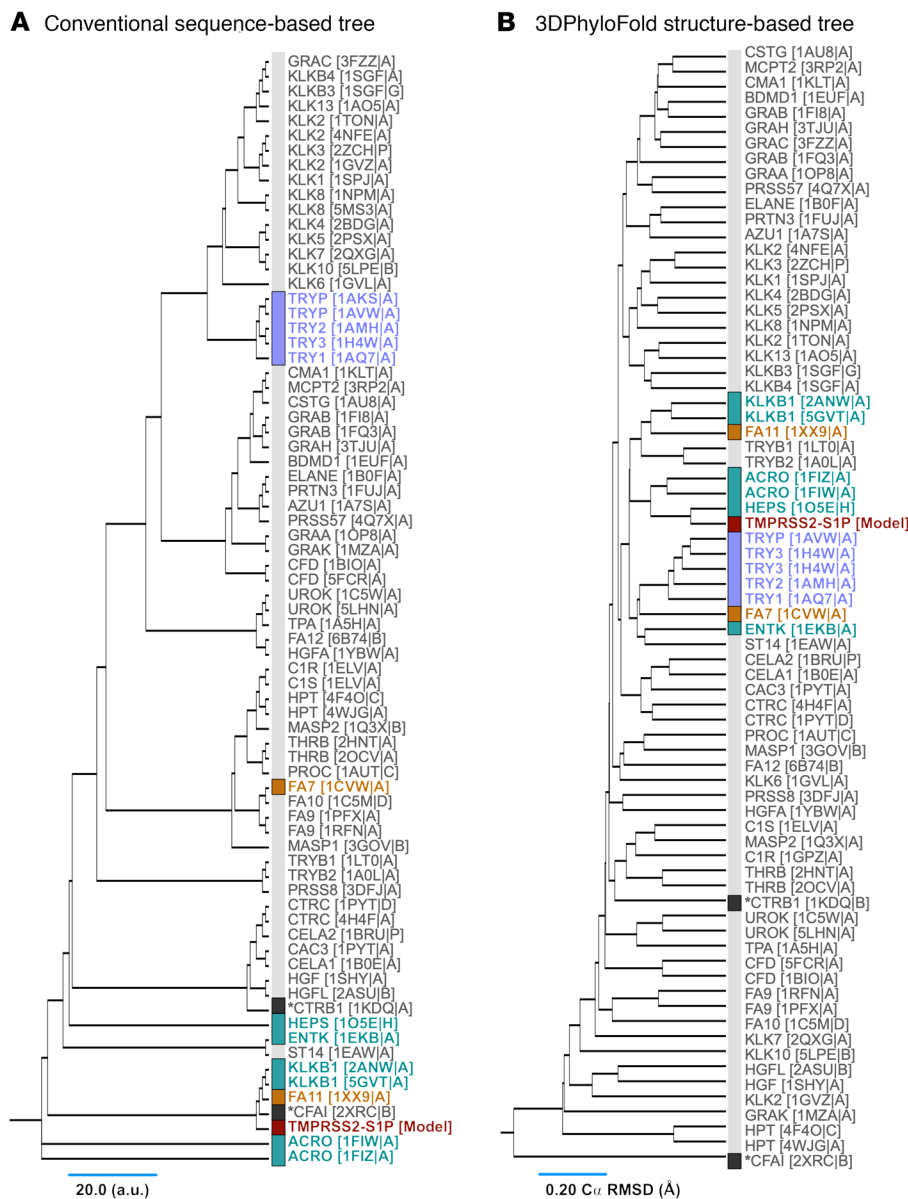
B 3DPhyloFold structure-based tree

Figure 2. Structure-based phylogenetic analysis identifies sequentially divergent serine proteases with similar folds to TMPRSS2. (A) Sequence-based phylogenetic tree of 75 representative peptidase domains with the main clusters highlighted in different colors. Branches are labeled according to the corresponding structures and Protein Data Bank IDs. **(B)** Structure-based phylogenetic tree of mammalian S1-peptidase domain structures. Evolutionary distance was inferred using the UPGMA method. Branch lengths correspond to the Cα RMSD (Å) of the pairwise structural alignments calculated in 3DPhyloFold. The proteases with the highest structural similarity to TMPRSS2-S1P were hepsin, acrosin, trypsin, factor VII, factor XI, and KLKB1.

original target proteases, and there was no inhibition of non–S1 proteases. Strikingly, avoralstat was more than 18-fold more selective toward TMPRSS2 than other S1 proteases (Figure 3E and Supplemental Table 6). Camostat was not as selective as avoralstat. We further characterized avoralstat's specificity by expanding the protease screen to include an additional 60 structurally distant proteases, including MMPs, caspases, cathepsins, and cysteine or aspartyl proteases (4, 12–14). Avoralstat displayed potent inhibition of other S1 proteases, consistent with their proximity to TMPRSS2 in the 3DPhyloFold tree, while displaying no inhibition of non–S1 proteases (Figure 4A and Supplemental Table 7), suggesting inhibition was specific and not due to protein aggregation effects. Notably, avoralstat inhibited several proteins that were structurally similar to TMPRSS2, including factor VIIa and tryptase b2, despite being distant in primary sequence (Figure 4A). Conversely, avoralstat was less effective at inhibiting proteases that clustered further from TMPRSS2 on the 3DPhyloFold tree: chymotrypsin ($IC_{50} \geq 1 \mu M$) and elastase ($IC_{50} \geq 1 \mu M$; Figure 4B), despite their proximity in the primary sequence phylogenetic tree. To further confirm that the compounds target the protease domain of TMPRSS2, we tested inhibition using catalytically active recombinant S1P domain (residues 252–489; Supplemental Figure 5) and found similar inhibition trends (Figure 4C). Taken together, these results suggested avoralstat was highly selective for TMPRSS2, consistent with the predictions by structural phylogenetic analysis.

Cellular evaluation of 3DPhyloFold inhibitors. Inhibition of full-length TMPRSS2 (TMPRSS2-FL) proteolytic activity was then tested in cells. TMPRSS2-FL contains an autoproteolysis motif (residues 252–257), which is subject to cleavage and can be used to probe the activity of TMPRSS2 in cells (15). Cells were transfected with either WT or loss-of-function TMPRSS2-S441A mutant (Figure 5, A and B, and Supplemental Figure 6). Compared with the inactive S441A mutant, TMPRSS2-WT showed reduced signal by immunoblot as previously reported (Figure 5B) (15). Inhibitor treatment prevented TMPRSS2-FL autoproteolysis and significantly increased the TMPRSS2-FL band intensity (Figure 5B).

To test whether the compounds specifically inhibited viral entry, we used vesicular stomatitis virus–based (VSV-based)
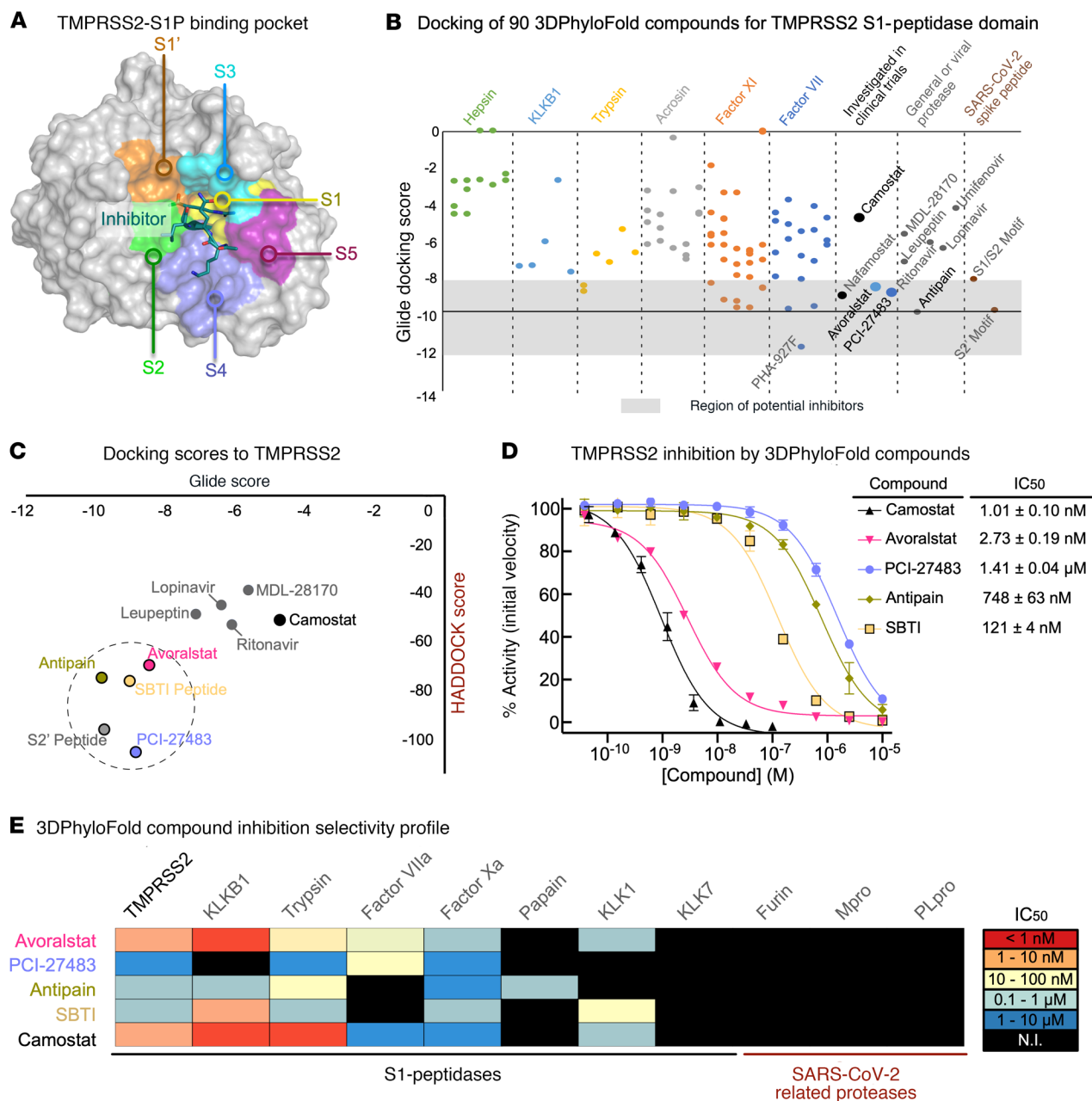
27483 ($IC_{50} = 1.41 \pm 0.04 \mu M$; Figure 3D). Inhibition by avoralstat was as potent as camostat ($IC_{50} = 1.01 \pm 0.10$ nM), which targets TMRPSS2 and is currently under clinical investigation for SARS-CoV-2 treatment (7). We further explored the selectivity profile of the 4 3DPhyloFold inhibitors and camostat as a positive control by testing them against 6 S1-peptidases identified in 3DPhyloFold (i.e., KLKB1, trypsin, factor VIIa, factor Xa, KLK1, and KLK7), 3 other proteases involved in SARS-CoV-2 infection (i.e., furin, Mpro, and PLpro), and a negative control papain. As expected, each 3DPhyloFold compound displayed potent inhibition of its

Figure 3. Compounds derived from 3DPhyloFold in silico analysis inhibit TMPRSS2 activity in vitro. (**A**) TMPRSS2-S1P structural model ligand-binding subpockets (S1′, S1, S2, S3, S4, S5). (**B**) Docking scores of compounds curated from 3DPhyloFold and (**C**) correlation between algorithms. Potential inhibitors clustered around the natural S2′-peptide motif. (**D**) TMPRSS2 inhibition by 3DPhyloFold compounds. IC_{50} data represent mean ± SEM; $n = 3$; calculated from the Hill equation. (**E**) IC_{50} value selectivity profile against S1- and SARS-CoV-2–related proteases. IC_{50} data was calculated from the Hill equation (mean ± SEM; $n = 3$).

pseudovirions bearing the SARS-CoV-2 spike glycoprotein and a luciferase reporter system. Human Calu-3 2B4 airway cells were incubated with camostat, avoralstat, PCI-27483, antipain, and SBTI. Pseudovirions harboring the pantropic VSV glycoprotein (VSV-G) served as controls since they transduce cells independent of TMPRSS2 (16). Indeed, no compounds prevented VSV-G pseudovirus entry, given that the luciferase signal remained constant (Figure 5C). Camostat inhibited SARS-CoV-2 pseudovirus entry ($EC_{50} = 0.7 ± 0.2$ μM), and avoralstat displayed similar inhibition ($EC_{50} = 2.8 ± 0.7$ μM). PCI-27483, antipain, and SBTI dis-

played modest inhibition but were too weak to determine reliable EC_{50} values (Figure 5D). Next, inhibition of authentic SARS-CoV-2 was tested in Calu-3 2B4 cells by measuring viral genomes. Camostat, avoralstat, and antipain significantly reduced SARS-CoV-2 replication (the amount of nucleocapsid gene [viral RNA] compared with vehicle; $P < 0.0001$). PCI-27483 and SBTI showed less inhibition (Figure 5E). A dose-response of camostat and avoralstat displayed significant reduction in SARS-CoV-2 infection beginning at 100 nM. Camostat and avoralstat caused more than a 10-fold decrease in viral RNA signal with a 1 μM dose (Figure 5F).
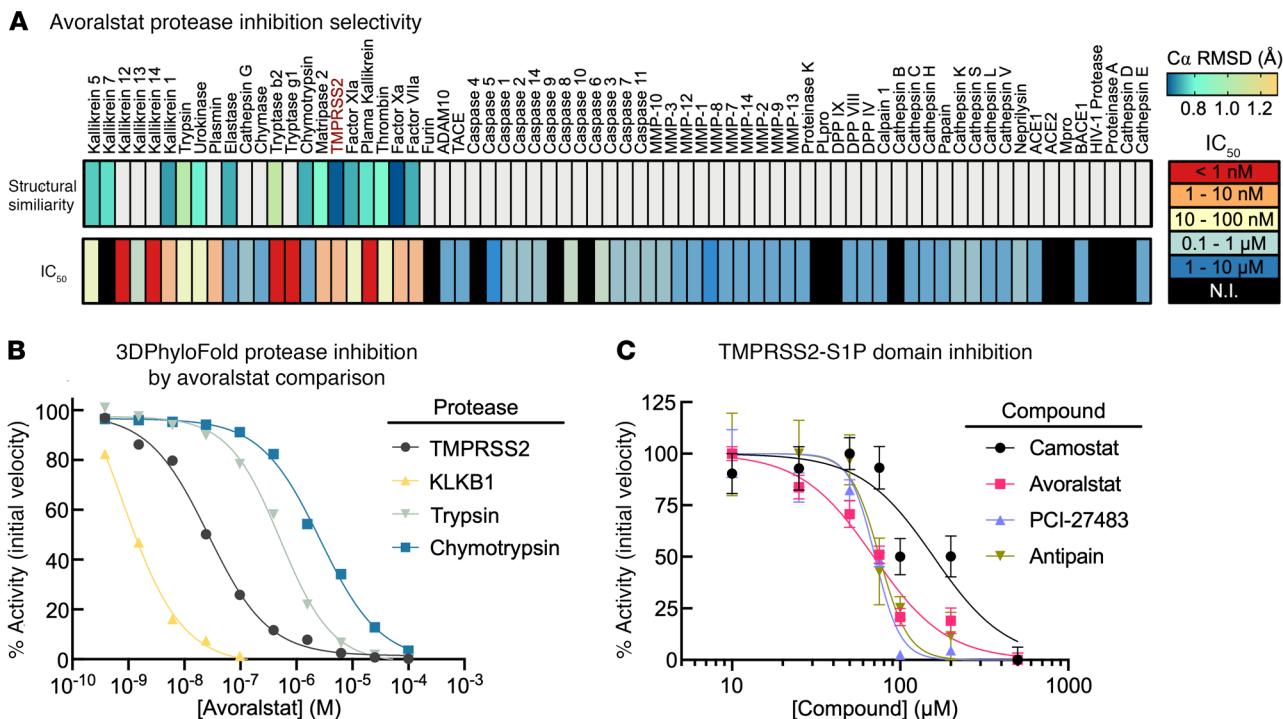
Figure 4. Avoralstat inhibits TMPRSS2 activity in vitro. (A) Avoralstat selectivity against 70 proteases. Top row – Cα-RMSD of pairwise structural alignments to TMPRSS2. **(B)** Protease inhibition by avoralstat correlates with 3DPhyloFold prediction. **(C)** TMPRSS2-S1P domain inhibition. IC$_{50}$ data represent mean ± SEM; n = 3; calculated from the Hill equation.

SARS-CoV-2 showed more sensitivity to avoralstat and camostat than MERS-CoV, another coronavirus that also uses TMPRSS2 to facilitate entry (Supplemental Figure 7) (6).

Avoralstat inhibits SARS-CoV-2 entry in vivo. No therapy targeting host-sensitizing proteases has been validated in an in vivo model of COVID-19. There is no known viral infection dose or animal model that fully recapitulates human disease, so the critical in vivo measure for testing prophylactic efficacy is the reduction of viral load. Using a mouse model of SARS-CoV-2 lung infection (Ad5-hACE2 transduced WT BALB/c mice) (17), we compared the efficacy of avoralstat and camostat in modifying SARS-CoV-2 infection. Cohorts of mice were infected (i.n.) with either 3×10^3 or 1×10^5 PFU of SARS-CoV-2. Mice were treated with avoralstat, camostat (30 mg/kg i.p. injection), or vehicle (DMSO). Lungs were harvested 1 day after infection and viral titers measured by plaque assay. Both avoralstat and camostat significantly reduced the lung tissue titers in both cohorts (Figure 6, A and B). In a third cohort of mice, avoralstat or camostat was administered 4 hours prior and 4 hours after a 1×10^5 PFU of SARS-CoV-2 i.n. challenge. Mice were given twice daily drug doses for 3 days after infection. Lungs harvested at 5 days after infection demonstrated significant reduction in viral titers by both drugs. Strikingly, the lung tissue virus titers were below the limit of detection in 3 of 4 avoralstat-treated mice (Figure 6C). Changes in weight, indicating the severity of illness, were monitored. Beginning at 4 days after infection, there was significant weight loss in the vehicle- and camostat-treated mice, whereas the weight of the avoralstat-treated mice remained relatively constant, suggesting a significant protective effect (Figure 6D). Although there was significant weight loss in the camostat

group, an avoralstat therapeutic effect was observed later at 7 days after infection compared with the vehicle-treated groups (Figure 6D). In a fourth cohort of mice, a biological dose-response was strongly supported after we further increased the SARS-CoV-2 challenge dose to 1×10^6 PFU. Avoralstat or camostat was administered 4 hours prior and 4 hours after a SARS-CoV-2 i.n. challenge. Mice were then given 2 drug doses daily for 3 days after infection. At the higher challenge dose, an early viral titer reduction was not observed as seen in lower titers (i.e., 3×10^3 or 1×10^5 PFU). Yet, a significant decrease of viral titer was observed at 4 days after infection for both avoralstat- and camostat-treated groups (Figure 6E). Moreover, avoralstat still showed a significant weight rescue effect beginning from 7 days after infection, whereas camostat did not show any rescue effect compared with the vehicle-treated group (Figure 6F). Thus, the inhibitory effect of avoralstat observed in biochemical and cell assays extended to prophylactic treatment of mice infected with escalating doses of SARS-CoV-2.

Discussion

Drug repositioning is an important strategy to address human disease at a faster pace than conventional drug development, especially in the setting of a global viral pandemic. Avoralstat, a clinically tested oral KLKB1 inhibitor evaluated for the treatment of hereditary angioedema, successfully inhibited SARS-CoV-2 infection and illness in mice. Avoralstat is orally bioavailable, which could facilitate prophylactic administration to people at high risk for COVID-19, particularly where specialized vaccine transport, cold storage, and medically skilled delivery staff are not available. Avoralstat possess a favorable plasma half-life of
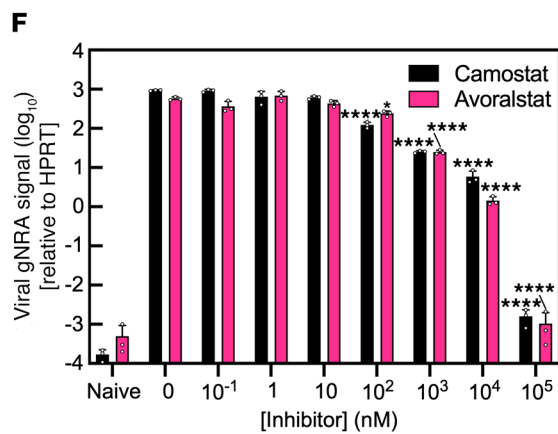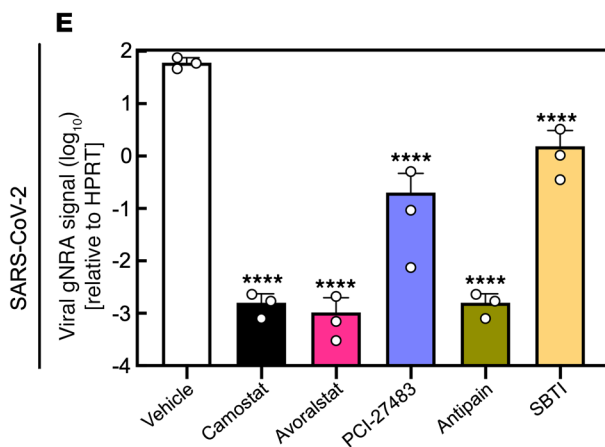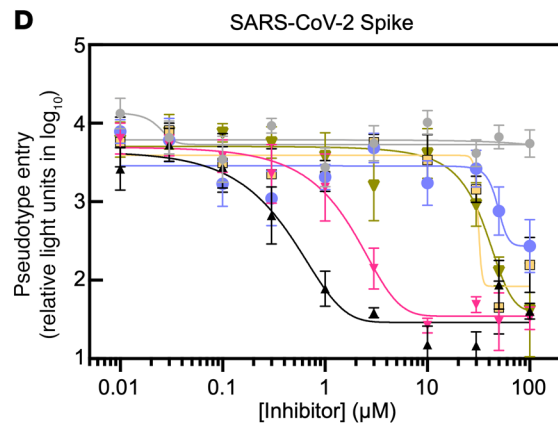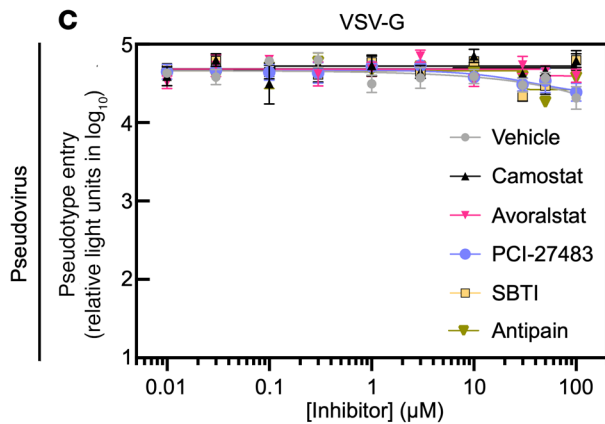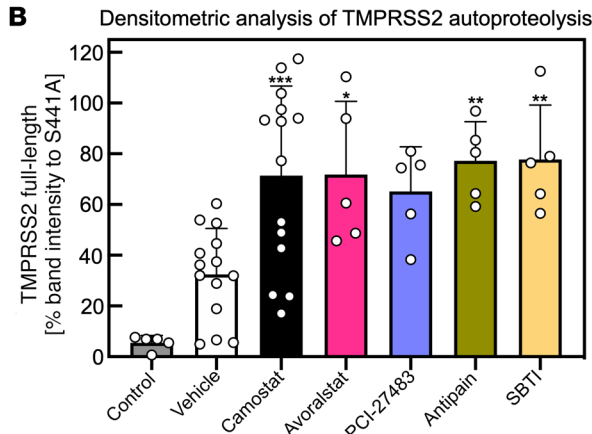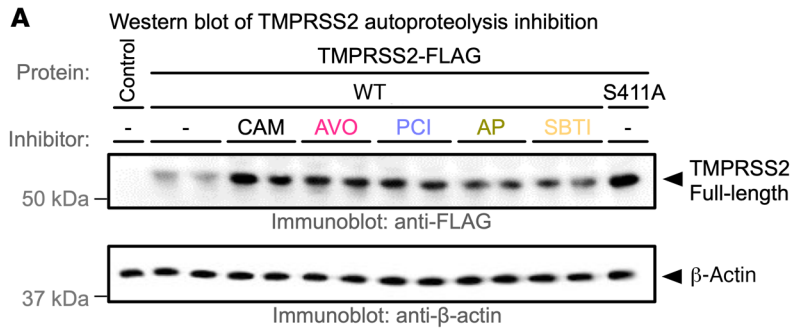
Figure 5. Avoralstat inhibits viral entry directed by SARS-CoV-2 spike proteins. (**A** and **B**) HEK cells treated with camostat, avoralstat, PCI-27483, antipain, or SBTI 2 hours before TMPRSS2 transfection. TMPRSS2 reduced autoproteolysis (increased TMPRSS2-FL signal; 1-way ANOVA followed by Dunnett's multiple-comparison test; *$P < 0.0332$, **$P < 0.0021$, ***$P < 0.0002$ compared with vehicle; $n = 5$ for each group except vehicle and camostat [$n = 14$]). Calu-3-cells were treated with compounds and inoculated with pseudovirions harboring (**C**) VSV-G or (**D**) SARS-CoV-2-spike-protein ($n = 6$; fit to the Hill equation). Calu-3 cells were treated with (**E**) 100 μM or (**F**) indicated concentrations of each compound, then incubated with SARS-CoV-2. Viral gRNA was measured after 24 hours. Data represent mean ± SEM; $n = 3$. (**E**) Compounds reduced viral signal at 100 μM (1-way ANOVA followed by Tukey's multiple-comparison test; mean ± SEM; $n = 3$; ****$P < 0.0001$ compared with vehicle). (**F**) Viral signal was reduced beginning from 100 nM (2-way ANOVA followed by Dunnett's multiple-comparison test; mean ± SEM; $n = 3$; *$P < 0.0332$, ****$P < 0.0001$ compared with vehicle).

12–31 hours, compared with the short half-life of camostat, due to an easily cleavable ester bond, giving it a terminal half-life of roughly 1 hour (18–21). It is possible that the observed efficacy of avoralstat compared with camostat in our in vivo study is due to its longer plasma half-life. If avoralstat i.p. administered to mice was 100% available in plasma, a 30 mg/kg of dose is roughly 80 nM in the water volume of a mouse (average volume percentile of water in mice is 73%). This molar concentration is about 29-fold higher compared with the in vitro IC_{50}. Although bioavailability was not directly measured, the mouse avoralstat dose of 30 mg/kg translates to a maximum recommended starting dose of 0.224 mg/kg in humans according to FDA guidelines, which is markedly lower than the dose previously administered in clinical trials (up to 1500 mg per day). Our biochemical assay data for avoralstat's protease selectivity showed that there are 10 proteases that have similar or lower avoralstat IC_{50} values than for TMPRSS2 (Figure 4A), yet avoralstat showed relatively minor and manageable side effects in humans. No grade 3 adverse events were observed in phase 1 through phase 3 clinical trials for avoralstat, and serious adverse events were no more prevalent in treatment compared with placebo groups (20, 21). It is possible that the observed efficacy of avoralstat compared with camostat in our in vivo study is due to its longer plasma half-life. Avoralstat has relatively minor and manageable side effects. The doses we tested in mice through i.p. administration correlate with a substantially lower dose than previous oral administration to humans in clinical trials (22), suggesting an appropriate dose of avoralstat for treating COVID-19 may be achievable with reasonable safety.

The application of a targeted structure-based phylogeny approach allowed us to identify and rationally prioritize several candidate TMPRSS2 inhibitors not considered by other drug repositioning strategies: 3DPhyloFold pointed to closely related proteins missed by primary-sequence comparisons, supporting a mechanism-based, hypothesis-driven selection of curated inhibitor candidates. Many of the small molecules tested in this study could be further developed for alternative routes of administration and for more potency and selectivity against TMPRSS2. The SBTI protein might serve as a cheap and natural source TMPRSS2 inhibitor, given it has been widely used in biomedical research (23). Interestingly, avoralstat and PCI-27483 were both represented in HTSs but may have been missed because of the lack of sufficient testing conditions (e.g., dosage or cell line) (24).

Our in vivo studies underscore that targeting TMPRSS2 is a tenable strategy for COVID-19 treatment. A reduction of viral load achieved by an alternative mechanism to that of vaccination could act synergistically to reduce illness and transmission. In addition, TMPRSS2 is implicated in the cleavage of the envelope-glycoproteins of many other viruses, including SARS-CoV, MERS-CoV, HCoV-229E, HCoV-OC43, HCoV-HKU1, and HCoV-NL63; influenza viruses; parainfluenza viruses; and human metapneumovirus (8). Thus, targeting this host machinery could be applied as a long-term strategy for future zoonotic coronaviruses and other respiratory viruses. This may be especially important if targeting viral proteins is only partially effective, natural infection does not confer long-lasting immunity, and combination therapies are needed to reduce the likelihood of resistance (25).

Methods

Experimental model and subject details: mice, virus, and cells

Specific pathogen–free 6-week-old male and female BALB/c mice were purchased from Envigo and maintained in the animal care facilities at the University of Iowa. The human serotype 5 adenoviral vector expressing human ACE2 under the control of the CMV promoter was previously described (VVC-McCray-7580; University of Iowa Viral Vector Core) (17). The SARS-CoV-2 strains (SARS-Related Coronavirus 2 Isolate USA-WA1/2020) were obtained from BEI Resources (NR-52281) and Calu-3 2B4 cells were obtained from the Perlman Laboratory, University of Iowa. HEK-293T cells (ATCC, CRL-3216) were obtained from the Viral Vector Core Facility at the University of Iowa. pVSV-ΔG-Luc was previously described (16). Calu-3 2B4 cells were grown in MEM (GIBCO) supplemented with 20% FBS.

Database search and sequence alignment

We first searched the UniProt database for reviewed entries denoted as transmembrane serine proteases (containing an S1-peptidase domain). This initial search yielded 9 manually curated sequences. A seed multiple-sequence alignment of S1-peptidase domains was then constructed using MAFFT v7 (alignment strategy: FFT-NS-1) (26). Using HMMER-3.1 and the seed alignment, we produced an HMM profile and used it to broaden the search against the UniProt database (search restricted to reviewed sequences) (27). This search yielded a total of 828 S1-peptidase sequences. We discarded fragmented sequences (<200 amino acids) that appeared too short to truly represent the S1-peptidase fold, and redundant proteins were further filtered using CD-HIT v4 (100% threshold) (28). This resulted in a pool of 742 proteins that were aligned using MAFFT v7 (alignment strategy FFT-NS-2) (26). Sequences producing many gaps in the alignment were removed using MaxAlign, resulting in 600 S1-peptidase sequences (29).

Phylogenetic tree reconstruction

We used the IQ-TREE-1.6.2 algorithm to generate a maximum likelihood tree of the 600 S1-peptidase sequences (30). The IQ-TREE model finder tool was used to determine the best substitution model to fit the data. The Whelan and Goldman (WAG) substitution model was determined to be the best fit to the data. Bootstrap analysis was performed using the "ultra-fast" method in IQ-TREE-1.6.2 with 1000 replicas.

Structural modeling of TMPRSS2-S1P

Briefly, a BLAST search of human TMPRSS2-S1P against the Protein Data Bank (PDB) returned the structure of human hepsin (PDB 1Z8G) as the top hit. Other close matches were KLKB1 (PDB 6ESO), plasminogen (PDB 4DUR), and prostatin (PDB 3E16). A TMPRSS2-S1P model was generated with the hepsin template (41% sequence identity) using Phyre2, MODELLER, and SWISS-Model. The models were in agreement and aligned well with minor variations in surface-exposed loop regions. The TMPRSS2-S1P model was then analyzed by ConSurf as previously described (31). The 600 sequences from our sequence-based phylogenetic analysis underwent multiple-sequence alignment using MAFFT and conservation scores were calculated using the Bayesian method option in ConSurf. The TMPRSS2-S1P binding pocket was inferred by comparison to the structure of hepsin bound to a pep-
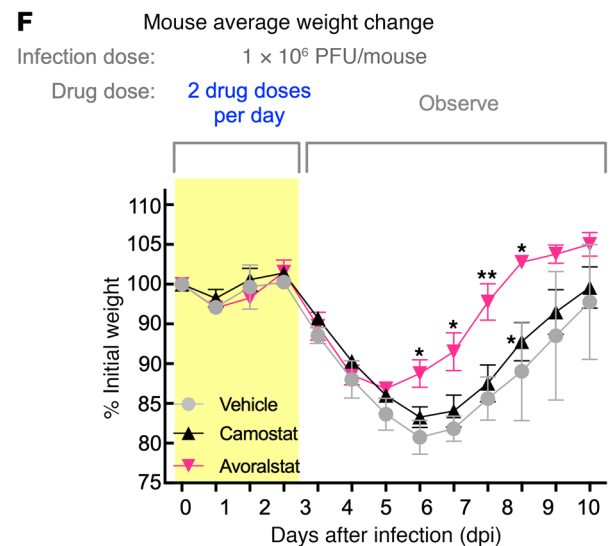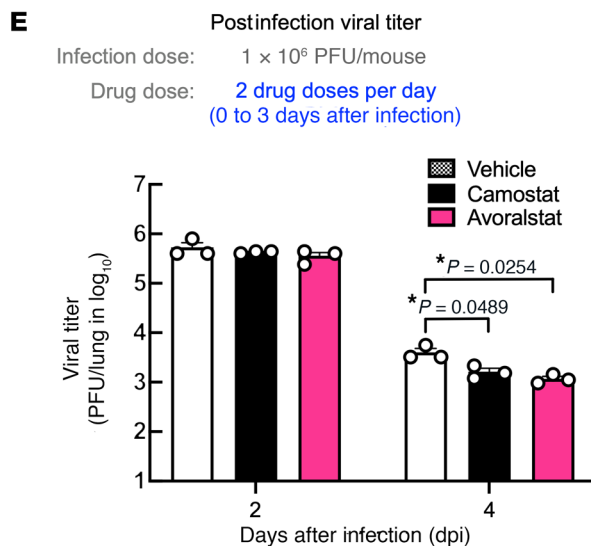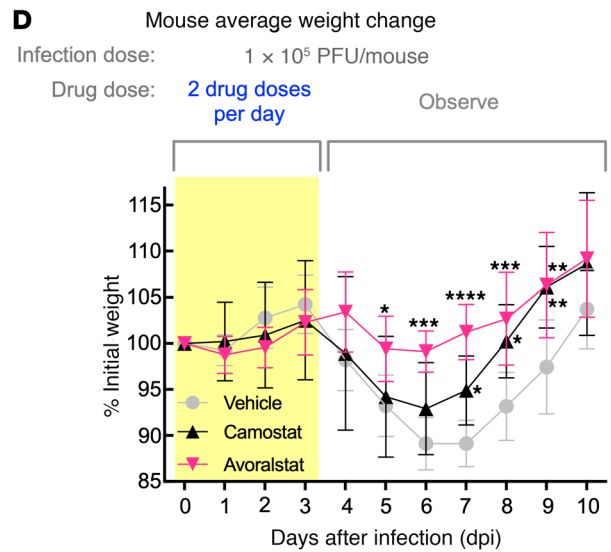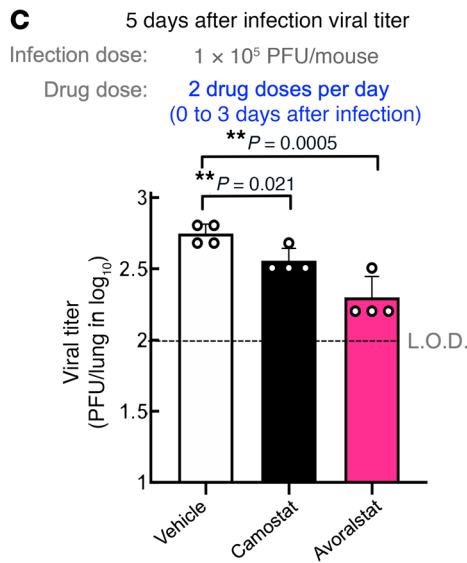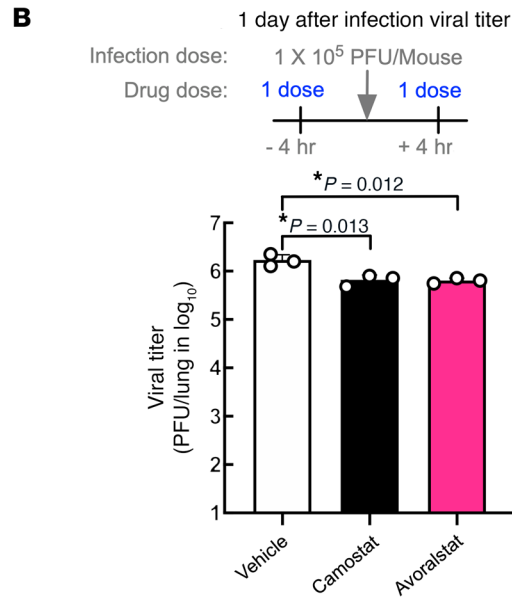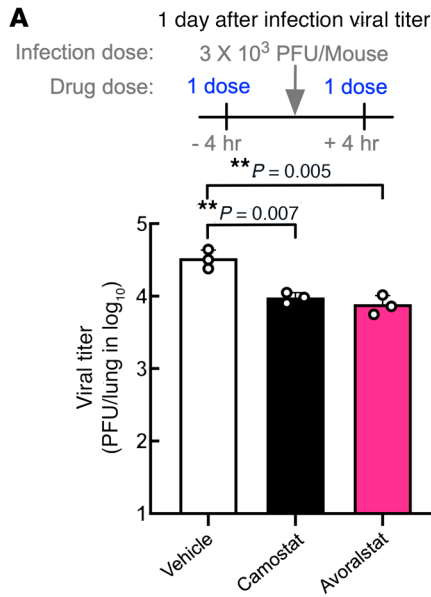
Figure 6. Avoralstat reduces SARS-CoV-2 infection in mice. (**A–C**) Ad5-hACE2-transduced BALB/c mice treated with avoralstat, camostat, or DMSO 4 hours before and after SARS-CoV-2 viral challenge showed significantly reduced lung viral titer (1-way ANOVA followed by Tukey's multiple-comparison test; *n* = 3 or 4; LOD, limit of detection). (**D**) Avoralstat protected mice from weight loss better than camostat, *n* = 6 for each group (2-way ANOVA followed by Dunnett's multiple-comparison test; *P < 0.0332, **P < 0.0021, ***P < 0.0002, ****P < 0.0001 compared with vehicle; mean ± SEM). (**E**) Avoralstat significantly reduced lung viral titer, *n* = 3 for each group (1-way ANOVA followed by Tukey's multiple-comparison test). (**F**) At the highest dose of SARS-CoV-2, mouse weight loss was better mitigated by avoralstat compared with vehicle and camostat, *n* = 6 for each group (2-way ANOVA followed by Dunnett's multiple-comparison test; *P < 0.0332, **P < 0.0021 compared with vehicle; mean ± SEM).

tidomimetic inhibitor (PDB 1Z8G) in PyMOL (the PyMOL Molecular Graphics System, v1.8 Schrödinger, LLC.).

Structure-based phylogenetic analysis

There are over 2000 structures of S1-peptidase domains represented in the PDB. We therefore searched the Pfam database for structures of mammalian peptidases and selected 74 representative structures (representing the WT protein) with an atomic resolution of 3.2 Å or better (Supplemental Table 2) (32). One structure per unique protein, fitting the above criteria, was selected. Structures (with reflection data deposited in the PDB) were evaluated by their reported global validation metrics in PDB-REDO (33). Rerefined structural models were used for further analysis. Structures were superimposed using PyMOL to calculate the pairwise root mean square deviation (RMSD) between protein alpha carbon atoms (Cα). A structural dissimilarity matrix (SDM) was constructed using the Cα RMSD values in order to generate a phylogenetic tree as previously described (31). To expedite the pairwise alignment process, we developed a Python-based script (named 3DPhyloFold) to perform the pairwise alignment of protein structures and generate an SDM. The phylogenetic tree was constructed using the UPGMA (unweighted pair group method with arithmetic mean) method in MEGAX software as previously described (31). For comparison, the sequences from the corresponding structures were also analyzed by sequence-based phylogeny. The 75 S1-peptidase sequences were aligned with MAFFT v7 (26) and analyzed in IQ-TREE-1.6.2. (30). The Jones-Taylor-Thornton (JTT) substitution model was determined to be the best fit to the data. Bootstrap analysis was performed in IQ-TREE-1.6.2 (1000 replicas). A TMPRSS2-S1P structure-similarity score for each analyzed protease was calculated by dividing the pairwise sequence identity (to TMPRSS2-S1P) by the Cα RMSD of the pairwise alignment (Supplemental Figure 2).

Database search for S1-peptidase inhibitors

We first searched for inhibitors designed for the related proteins. We filtered for cases where a strong structure-activity-relationship (SAR) between the ligand and protein was studied, where applicable. We discarded studies where the inhibitors displayed low potency, focusing on groups of inhibitors that displayed submicromolar inhibition for their intended protein. We focused on cases where the inhibitors studied contained a guanidine or guanidine-like functional group to interact with the S1 specificity pocket. Inhibitors were then prepared and docked against our TMPRSS2-S1P model.

In silico docking calculations

Published crystal structures of inhibitor-bound trypsin-3 (PDB 1H4W), KLKB1 (PDB 6O1S), and factor VII (PDB 1W7X) were loaded into Maestro software (Schrödinger release 2019-3). The TMPRSS2-S1P model described above was used. The protein preparation wizard was used to prepare the proteins for docking and simulations. The default parameters were used for the optimization of hydrogen-bond assignment (sampling of water orientations and use of pH 7.0). Water molecules beyond 3 Å of heteroatoms or with fewer than 3 hydrogen bonds to non–water molecules were removed. Restrained energy minimization was applied using the OPLS3e force field. Prepared protein systems were further checked by Ramachandran plots, ensuring there were no steric clashes. To generate receptor grids for small molecule docking, the cocrystalized ligand was selected as the grid-defining ligand for each system. For peptides, the grid size was made suitable for peptides to be docked. Default van der Waals radius scaling parameters were used (scaling factor of 1, partial charge cutoff of 0.25). For docking of the ligands into the various prepared proteins, the 3D structure was loaded into Maestro. Ligprep was used to prepare the ligands (by generating possible states at pH 7.0 ± 2.0 and retaining the specified stereochemical properties). The prepared small molecule ligands and peptide fragments were then docked using the most stringent docking mode (extra precision, "XP") of Glide. Parameters and output files for the Glide runs can be found in Mendeley Data (https://data.mendeley.com/) under the dataset identifier (DOI) 10.17632/h3pmycddwc.1.

Docking SBTI to TMPRSS2

The HADDOCK 2.4 online docking tool was used to generate an TMPRSS2-S1P/SBTI complex structure model (34). The TMPRSS2-S1P homology model and the SBTI structure (PDB 1AVW) were used for docking. To define the potential interaction surface between TMPRSS2 and SBTI, the TMPRSS2-S1P homology model was superimposed to the wild boar trypsin structure in complex with SBTI (PDB 1AVW) using PyMOL. The following residues of SBTI were designated as active residues: 501–502, 510, 512–514, 560–572, and 616–617. The overall Cα RMSD between the 2 models was 0.54 Å. SBTI was also docked to porcine trypsin (PDB 1AVW), human factor VII (PDB 1W7X), and human KLKB1 (PDB 6O1S). The HADDOCK scores represent the average score of the best cluster. The parameters and output files for the HADDOCK run can be found in Mendeley Data (https://data.mendeley.com/) under the DOI 10.17632/h3pmycddwc.1.

Protease activity array

Avoralstat, SBTI, PCI-27483, and antipain were assessed for inhibition against TMPRSS2 and a panel of recombinant proteases by commercial services from Reaction Biology. The Reaction Biology profile tested in a 10-dose IC_{50} with, in triplicate, a 4-fold serial dilution starting at 10 μM against 11 proteases in Figure 3E and a 3-fold serial dilution starting at 10 μM against 70 proteases in Figure 4A. Compounds exhibited no fluorescent background that could interfere with the assay. The protease activities were monitored as a time-course measurement of the increase in fluorescence signal from fluorescently labeled peptide substrate, and initial linear portion of slope (signal/min) was analyzed.

TMPRSS2-S1P expression and purification

The human TMPRSS2-S1P sequence (residues 252–489) was cloned into a pET28a vector with an N-terminal 6x-His tag. Plas-

mids were amplified and isolated from DH5α cells and transformed into *E. coli* BL21 (DE3). BL21 cells expressing TMPRSS2-S1P were induced with 0.5 mM Isopropyl β-d-1-thiogalactopyranoside (IPTG). Cell pellets were resuspended in 35 to 50 mL of lysis buffer (50 mM Tris, 150 mM NaCl, 20 mM Imidazole pH 8.0, 1 tablet of EDTA-free protease inhibitor [Roche, COEDTAF-RO], DNaseI [Roche, 11284932001]) and lysed and centrifuged for 30 minutes at 18,000 g at 4°C. Pellets were denatured (50 mM Tris, 150 mM NaCl, 6 M guanidinium chloride, 1 M L-arginine, 2 mM DTT pH 8.0), resuspended, and filtered with 0.22 μm filter. Refolding buffer-1 (50 mM Tris, 150 mM NaCl, 2 M guanidinium chloride, 1 M L-arginine pH 8.0) was applied to SnakeSkin dialysis tubing (10,000 MWCO; Thermo Fisher Scientific) and underwent refolding by dialyzing in 2 L of refolding buffer-1 at 4°C. After the overnight refolding, the sample was filtered with 0.22 μm filter to remove aggregates and went through another step of dialysis in 2 L of refolding buffer-2 (50 mM Tris, 150 mM NaCl, 250 mM L-arginine pH 8.0) for 1.5 hours at room temperature. Sample was concentrated with a 10 kDa nominal molecular weight limit (NMWL; pore size) spin concentrator and passed over a HiLoad 16/600 Superdex 200 pg (GE Healthcare, 28-9893-35) size-exclusion column (SEC) connected to an ÄTKA pure fast protein liquid chromatography (FPLC) system (GE Healthcare). The column was equilibrated with SEC buffer (50 mM Tris, 150 mM NaCl, pH 8.0). The final purity of recombinant TMPRSS2-S1P used for in vitro assays was more than 95% (Supplemental Figure 5A).

Measurement of TMPRSS2-S1P activity

TMPRSS2-S1P proteolytic activity was confirmed by hydrolysis of the synthetic urokinase substrate, Cbz-GGR-AMC (Echelon Biosciences, 869-25). An enzyme titration in the presence of 50 μM Cbz-GGR-AMC revealed that maximal TMPRSS2-S1P activity occurred at high nanomolar (250–500 nM) protein concentrations (data not shown). The remaining assays were performed as follows. Briefly, 250 nM of purified TMPRSS2-S1P was added to a reaction buffer containing 50 mM Tris-HCl (pH 8.0) and 150 mM NaCl in black-bottom 96-well plates (100 μL per reaction). Inhibition experiments were carried out in the presence of 50 μM Cbz-GGR-AMC in the presence 10 to 500 μM compound: camostat (Sigma-Aldrich, SML0057), avoralstat (MedChemExpress, HY-16735), PCI-27483 (Cayman Chemical, 21334), antipain (Sigma-Aldrich, A6191), leupeptin (Sigma-Aldrich, L2884), MDL-28170 (Sigma-Aldrich, M6690), ritonavir (Sigma-Aldrich, SML0491), or 5% DMSO (as a negative control). DMSO caused SBTI (Roche, 10109886001) to precipitate out of solution (unpublished observation). Inhibition experiments with SBTI (2 to 150 μM) were therefore performed in the absence of DMSO. Reactions were run at 37°C for 30 minutes on a fluorometric plate reader (Tecan Spark). Proteolytic activity was measured as change in raw fluorescence units (ΔRFU; λ_{exc} = 373 nm, λ_{em} = 455 nm) at 30-second intervals. All experiments were performed in triplicate. The initial velocity (RFU/sec) of the reaction was measured by calculating the slope of the fluorescence data from the first 3 minutes. Kinetic parameters were then calculated by direct fitting to the Michaelis-Menten or Hill equation in GraphPad Prism 8. There was no activity, as expected with the cysteine protease substrate sLY-AMC (Bachem, 4002047; negative control; Supplemental Figure 5).

TMPRSS2 autoproteolysis assay

HEK-293T cells (ATCC, CRL-3216) were obtained from the Viral Vector Core Facility at the University of Iowa. Cells were grown in DMEM supplemented with 5% FBS (Gibco) and penicillin and streptomycin (Gibco, WT15140-122) and were maintained in a humidified atmosphere of 5% CO_2 at 37°C. Plasmid pEGFPN1 was obtained from Clontech. TMPRSS2-FL cDNA (pcDNA3.1-SARS-2-S-C9; obtained from the Gallagher Laboratory, Loyola University Medical Center, Illinois). Briefly, TMPRSS2-FL cDNA, containing a C-terminal anti-FLAG epitope tag, was amplified with PCR using pCMV-Sport6-TM-PRSS2 template. The amplificates were cloned into pCAGGS.MCS via SacI and XhoI sites. The enzymatically inactive pCAGGS-TMPRSS2(-S441A)FLAG mutant cDNA was generated using QuickChange Site-Directed Mutagenesis Kit per the manufacturer's instructions (Agilent Technologies). Transient transfections of HEK-293T cells were performed using PolyFect transfection reagent per the manufacturer's instructions (QIAGEN). For transfection, 2 μg of each plasmid (GFP [served as negative control], TMPRSS2 WT, and S411A mutant) were dissolved in serum-free media. PolyFect (20 μL) was added to the DNA solution followed by 10-minute incubation at room temperature. Growth media (0.6 mL) was then added to the reaction tubes and the transfection mix was immediately added onto the cells. Next, 24 hours after transfection, cell lysates were prepared using HNB buffer (50 mM HEPES [pH 7.4], 100 mM NaCl, 0.01% bovine serum albumin) containing 0.1% protease inhibitor (Sigma-Aldrich, P2714), incubated on ice for 20 minutes and centrifuged at 2000 g for 10 minutes. Supernatants were collected and protein concentration determined by DC protein assay reagent kit (Bio-Rad). After separation by SDS-PAGE (4% to 12% Bis-Tris gradient gel), proteins were transferred to a PVDF membrane and blocked for 1 hour at room temperature using 5% nonfat dry milk in TBST. Membranes were probed with mouse monoclonal anti-Flag antibody (1:1000; Sigma-Aldrich, F3165) for 16 hours at 4°C. Blots were then washed 3 times with TBST (10 minutes/wash) and subsequently incubated with IgG labelled with HRP-conjugated secondary anti-mouse antibody (1:5000; Thermo Fisher Scientific, 31432). Proteins were visualized by SuperSignal West Pico PLUS chemiluminescence reagent on a MyECL imager (Thermo Fisher Scientific). Membranes were reprobed with β-actin (1:5000; Sigma-AldrichA2228) as a loading control. The TMPRSS2-FL band intensity of each lane was normalized using the band intensity of corresponding β-actin loading control. Then, the normalized intensity of each lane was converted to the relative band intensity by comparison to the normalized band intensity of TMPRSS2-S441A in the same gel.

Pseudovirus transduction assay

HEK-293T cells were transfected to express either the SARS-CoV-2 spike protein (with the cytoplasmic tail removed; residues 1–1255) or the full-length VSV-G protein. Then, these cells were transduced with a VSV vector expressing luciferase (VSV-ΔG-Luc) and pseudotyped with SARS-CoV-2 spike protein or VSV-G. After 2 hours at 37°C, the cells were washed 3 times to remove residual virus. Supernatant containing pseudovirus was harvested 3 times at 24-hour intervals and centrifuged to remove cellular debris. Pseudovirus from the 3 collections was pooled and ultracentrifuged through a 20% sucrose cushion for purification and concentration (100×). For the transduction assays, Calu-3 2B4 cells were grown in 96-well plates until confluent. Cells were incubated with the respective compounds for 1 hour at 37°C. After

1 hour, cells were transduced with pseudovirus, maintaining the same concentration of compounds, and incubated overnight. Transduction efficiency was assessed by quantifying luciferase activity in cell lysates using a commercial kit (Luciferase Assay System, Promega, E1500) and a plate-reading luminometer (SpectraMax i3x, Molecular Devices).

Infectious SARS-CoV-2 neutralization assay

The 2019n-CoV/USA-WA1/2019 strain of SARS-CoV-2 (Accession number MT985325.1) used in these studies was passaged on Calu-3 2B4 cells and sequence verified. Calu-3 2B4 cells were plated in 48-well plates. Cells were incubated with medium containing indicated compounds or vehicle for 1 hour at 37°C. The medium was removed and SARS-CoV-2 (MOI = 0.1) in medium containing indicated compounds was added into each well. The cells were incubated with viruses for 1 hour at 37°C. Next, the viruses were removed, and cells were rinsed with PBS once to remove remaining viruses. After that, cells were incubated with medium containing indicated compounds overnight. The following day, total cellular RNA was isolated using Directzol RNA MiniPrep kit (Zymo Research, R2052) from TRIzol (Invitrogen, 15596018). A DNase treatment step was included. Total RNA (500 ng) was used for cDNA syntheses by High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, 4368814). Real-time PCR was applied to quantify viral genomic RNA and hypoxanthine phosphoribosyltransferase 1 (HPRT) mRNA levels (SARS-2-N1-F primer: GACCCCAAAATCAGC-GAAAT; SARS-2-N1-R primer: TCTGGTTACTGCCAGTTGAATCTG; human HPRT-F primer: AGGATTTGGAAAGGGTGTTTATTC; human HPRT-R primer: CAGAGGGCTACAATGTGATGG; Integrated DNA Technologies). The relative abundance of viral genomic RNA normalized to HPRT was calculated and presented as $2^{-\Delta CT}$.

Transduction and infection of mice

Mice were anesthetized with ketamine and xylazine (87.5 mg/kg ketamine and 12.5 mg/kg xylazine) and transduced (i.n.) with 2.5×10^8 FFU of Ad5-ACE2 in 75 mL DMEM. Five days after transduction, mice were infected (i.n.) with SARS-CoV-2 (3×10^3 or 1×10^5 PFU) in a total volume of 50 mL DMEM. Infected mice were treated with avoralstat, camostat (30 mg/kg i.p. injection), or vehicle (DMSO; negative control) either 4 hours before and after being challenged by virus or 2 doses per day (8 to 9 hours apart) for 3 days after infection. Virus titers were measured in harvested lungs by plaque assay 1, 2, 4, or 5 days after infection. The weight was monitored for 10 or 12 days after infection.

SARS-CoV-2 plaque assay

Lung homogenate supernatants were serially diluted in DMEM. Vero E6 cells in 12-well plates were inoculated at 37°C in 5% CO_2 for 1 hour with gentle rocking every 15 minutes. After removing the inocula, plates were overlaid with 1.2% agarose containing 10% FBS. After further incubation for 3 days, overlays were removed, and plaques were visualized by staining with 0.1% crystal violet. Viral titers were calculated as PFU per lung. All work with SARS-CoV-2 was conducted in the biosafety level 3 laboratories of the University of Iowa.

Statistics

Kinetic parameters of biochemical protease assays. Data are displayed as mean ± SEM (*n* = 3; number of reactions) and fit to the Hill equation (Figure 3D and Figure 4, B and C) or to the Michaelis-Menten equation (Supplemental Figure 5C) in GraphPad Prism 8.

TMPRSS2 autoproteolysis assay (Figure 5B). For densitometry analysis, the analyzed densitometry data from the total of 5 gel runs in Figure 5A and Supplemental Figure 6 were combined (*n* = 14 for vehicle and camostat; *n* = 5 for each group of GFP, avoralstat, PCI-27483, antipain, and SBTI). Data represent the mean ± SEM from 2 independent experiments analyzed by 1-way ANOVA followed by Dunnett's multiple-comparison test using GraphPad Prism 8.0. Differences of *P* < 0.0332 were considered statistically significant (*P* < 0.0332, **P* < 0.0021, ***P* < 0.0002).

Pseudovirus transduction assay (Figure 5, C and D). Data represent the mean ± SEM (*n* = 6; number of technical replicates) and are fit to the Hill equation (Figure 5C: camostat R^2 = 0.71; avoralstat R^2 = 0.74; PCI-27483 R^2 = 0.20; SBTI R^2 = 0.49; antipain R^2 = 0.48).

Infectious SARS-CoV-2 neutralization assay (Figure 5E). SARS-CoV-2 viral gRNA in the presence of 100 μM of DMSO (vehicle; negative control) or inhibitor. Data represent the mean ± SEM (*n* = 3; number of technical replicates) and were analyzed by 1-way ANOVA followed by Tukey's multiple-comparison test (****P* < 0.0001 compared with vehicle).

Infectious SARS-CoV-2 neutralization assay (Figure 5F and Supplemental Figure 7). SARS-CoV-2 or MERS-CoV viral gRNA as a function of camostat or avoralstat concentration. Data represent the mean ± SEM (*n* = 3; number of technical replicates) and were analyzed by 2-way ANOVA followed by Dunnett's multiple-comparison test (*P* < 0.0332, ****P* < 0.0001 compared with vehicle).

Viral titers in transduction and infection of mice (Figure 6, A, B, and E). Data are represented as mean ± SEM (*n* = 3; number of mice) and were analyzed by 1-way ANOVA followed by Tukey's multiple-comparison test (*P* < 0.0332; **P* < 0.0021, ***P* < 0.0002, ****P* < 0.0001 compared with vehicle).

Viral titers in transduction and infection of mice (Figure 6C). Data are represented as mean ± SEM (*n* = 4 for each group) and were analyzed by 1-way ANOVA followed by Tukey's multiple-comparison test (*P* < 0.0332; **P* < 0.0021, ***P* < 0.0002, ****P* < 0.0001 compared with vehicle).

Weights in transduction and infection of mice (Figure 6, D and F). Data are represented as mean ± SEM (*n* = 6; number of mice) and were analyzed by 2-way ANOVA followed by Dunnett's multiple-comparison test (*P* < 0.0332; **P* < 0.0021, ***P* < 0.0002, ****P* < 0.0001 compared with vehicle). *P* values of less than 0.0332 were considered significant.

Study approval

All animal studies were approved by the IACUC of the University of Iowa.

Data and materials availability. Reagents are available with a Materials Transfer Agreement. 3DPhyloFold is open source and available at Mendeley Data (https://data.mendeley.com/) under DOI 10.17632/kk3gkzdsbf.2. The implementation notes, code, and description of methodology are available on the site. The raw docking data and parameters have been deposited to Mendeley Data with the DOI 10.17632/h3pmycddwc.1.

Author contributions

AGB and VBM were responsible for the study concept and design. YJS, GV, DP, KL, MO, and SS performed data acquisi-

Address correspondence to: Vinit B. Mahajan, 1651 Page Mill Rd., Room #2113, Byers Eye Institute, Department of Ophthalmology, Stanford University, Palo Alto, California 94304, USA. Phone: 650.723.6995; Email: vinit.mahajan@stanford.edu. Or to: Alexander G. Bassuk, 25 S. Grand Ave, Med Labs Building, Room 2040a, Department of Pediatrics, Neurology, and the Iowa Neuroscience Institute, University of Iowa, Iowa City, Iowa 52242, USA. Phone: 773.744.6338; Email: alexander-bassuk@uiowa.edu. Or to: Paul B. McCray Jr., 6320 PBDB, 169 Newton Road, Department of Pediatrics, University of Iowa, Iowa City, Iowa 52242, USA. Phone: 319.335.6844; Email: paul-mccray@uiowa.edu.

1. Kawase M, et al. Simultaneous treatment of human bronchial epithelial cells with serine and cysteine protease inhibitors prevents severe acute respiratory syndrome coronavirus entry. *J Virol*. 2012;86(12):6537–6545.

2. Iwata-Yoshikawa N, et al. TMPRSS2 contributes to virus spread and immunopathology in the airways of murine models after coronavirus infection. *J Virol*. 2019;93(6):e01815-18.

3. Zhou P, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020;579(7798):270–273.

4. Hoffmann M, et al. Correction: Wang, M., et al. HSP70-eIF4G interaction promotes protein synthesis and cell proliferation in hepatocellular carcinoma. *Cancers (Basel)*. 2020;12(11):3410.

5. Kim TS, et al. Phenotypic analysis of mice lacking the Tmprss2-encoded protease. *Mol Cell Biol*. 2006;26(3):965–975.

6. Shirato K, et al. Middle East respiratory syndrome coronavirus infection mediated by the transmembrane serine protease TMPRSS2. *J Virol*. 2013;87(23):12552–12561.

7. Shrimp JH, et al. An enzymatic TMPRSS2 assay for assessment of clinical candidates and discovery of inhibitors as potential treatment of COVID-19. *ACS Pharmacol Transl Sci*. 2020;3(5):997–1007.

8. Shen LW, et al. TMPRSS2: A potential target for treatment of influenza virus and coronavirus infections. *Biochimie*. 2017;142:1–10.

9. Midgley I, et al. Metabolic fate of 14C-camostat mesylate in man, rat and dog after intravenous administration. *Xenobiotica*. 1994;24(1):79–92.

10. Talevi A, Bellera CL. Challenges and opportunities with drug repurposing: finding strategies to find alternative uses of therapeutics. *Expert Opin Drug Discov*. 2020;15(4):397–401.

11. Schweinitz A, et al. Design of novel and selective inhibitors of urokinase-type plasminogen activator with improved pharmacokinetic properties for use as antimetastatic agents. *J Biol Chem*. 2004;279(32):33613–33622.

12. Jin Z, et al. Structure of M(pro) from SARS-CoV-2 and discovery of its inhibitors. *Nature*. 2020;582(7811):289–293.

13. Shin D, et al. Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. *Nature*. 2020;587(7835):657–662.

14. Johnson BA, et al. Furin cleavage site is key to SARS-CoV-2 pathogenesis [preprint]. https://doi.org/10.1101/2020.08.26.268854. Posted on bioRxiv August 26, 2020.

15. Shulla A, et al. A transmembrane serine protease is linked to the severe acute respiratory syndrome coronavirus receptor and activates virus entry. *J Virol*. 2011;85(2):873–882.

16. Whitt MA. Generation of VSV pseudotypes using recombinant ΔG-VSV for studies on virus entry, identification of entry inhibitors, and immune responses to vaccines. *J Virol Methods*. 2010;169(2):365–374.

17. Sun J, et al. Generation of a broadly useful model for COVID-19 pathogenesis, vaccination, and treatment. *Cell*. 2020;182(3):734–743.

18. Schneider CA, et al. [Assessing myocardial viability in chronic myocardial infarct with 18F-fluoro-D-glucose positron emission tomography and 99mTc-MIBI SPECT]. *Z Kardiol*. 1994;83(2):124–131.

19. Choi JY, et al. Nafamostat mesilate as an anticoagulant during continuous renal replacement therapy in patients with high bleeding risk: a randomized clinical trial. *Medicine (Baltimore)*. 2015;94(52):e2392.

20. Cornpropst M, et al. Safety, pharmacokinetics, and pharmacodynamics of avoralstat, an oral plasma kallikrein inhibitor: phase 1 study. *Allergy*. 2016;71(12):1676–1683.

21. Riedl MA, et al. Evaluation of avoralstat, an oral kallikrein inhibitor, in a Phase 3 hereditary angioedema prophylaxis trial: the OPuS-2 study. *Allergy*. 2018;73(9):1871–1880.

22. FDA. Estimating the Maximum Safe Starting Dose in Initial Clinical Trials for Therapeutics in Adult Healthy Volunteers https://www.fda.gov/regulatory-information/search-fda-guidance-documents/estimating-maximum-safe-starting-dose-initial-clinical-trials-therapeutics-adult-healthy-volunteers. Updated August 24, 2018. Accessed April 8, 2021.

23. Song HK, Suh SW. Kunitz-type soybean trypsin inhibitor revisited: refined structure of its complex with porcine trypsin reveals an insight into the interaction between a homologous inhibitor from Erythrina caffra and tissue-type plasminogen activator. *J Mol Biol*. 1998;275(2):347–363.

24. Bakowski MA, et al. Oral drug repositioning candidates and synergistic remdesivir combinations for the prophylaxis and treatment of COVID-19 [preprint]. https://doi.org/10.1101/2020.06.16.153403. Published on bioRxiv June 16, 2020.

25. Fragkou PC, et al. Corrigendum to "Review of trials currently testing treatment and prevention

of COVID-19" [Clin Microbiol Infect 26.8 (2020) 988-998]. *Clin Microbiol Infect*. 2020;27(3):499.

26. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 2013;30(4):772–780.

27. Finn RD, et al. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res*. 2011;39(Web Server issue):W29–W37.

28. Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*. 2006;22(13):1658–1659.

29. Gouveia-Oliveira R, et al. MaxAlign: maximizing usable data in an alignment. *BMC Bioinformatics*. 2007;8:312.

30. Nguyen LT, et al. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*. 2015;32(1):268–274.

31. Velez G, et al. Structural insights into the unique activation mechanisms of a non-classical calpain and its disease-causing variants. *Cell Rep*. 2020;30(3):881–892.

32. Finn RD, et al. Pfam: the protein families database. *Nucleic Acids Res*. 2014;42(Database issue):D222–D230.

33. Joosten RP, et al. PDB_REDO: automated re-refinement of X-ray structure models in the PDB. *J Appl Crystallogr*. 2009;42(pt 3):376–384.

34. de Vries SJ, et al. The HADDOCK web server for data-driven biomolecular docking. *Nat Protoc*. 2010;5(5):883–897.