



Structure and activity of human TMPRSS2 protease implicated in SARS-CoV-2 activation

Bryan J. Fraser^{1,2,8}, Serap Beldar^{3,8}, Almagul Seitova³, Ashley Hutchinson³, Dhiraj Mannar⁴, Yanjun Li³, Daniel Kwon^{1,2}, Ruiyan Tan¹, Ryan P. Wilson¹, Karoline Leopold⁴, Sriram Subramaniam⁴, Levon Halabelian^{3,5}✉, Cheryl H. Arrowsmith^{3,6,7}✉ and François Bénard^{1,2}✉

Transmembrane protease, serine 2 (TMPRSS2) has been identified as key host cell factor for viral entry and pathogenesis of SARS-CoV-2. Specifically, TMPRSS2 proteolytically processes the SARS-CoV-2 Spike (S) protein, enabling virus–host membrane fusion and infection of the airways. We present here a recombinant production strategy for enzymatically active TMPRSS2 and characterization of its matured proteolytic activity, as well as its 1.95 Å X-ray cocrystal structure with the synthetic protease inhibitor nafamostat. Our study provides a structural basis for the potent but nonspecific inhibition by nafamostat and identifies distinguishing features of the TMPRSS2 substrate binding pocket that explain specificity. TMPRSS2 cleaved SARS-CoV-2 S protein at multiple sites, including the canonical S1/S2 cleavage site. We ranked the potency of clinical protease inhibitors with half-maximal inhibitory concentrations ranging from 1.4 nM to 120 μM and determined inhibitor mechanisms of action, providing the groundwork for drug development efforts to selectively inhibit TMPRSS2.

Viruses hijack the biochemical activity of specific cell surface proteins to initially recognize the target cell and subsequently overcome the energetic barrier necessary for virus–host membrane fusion. SARS-CoV-2 uses its Spike (S) protein to recognize its primary cell surface receptor, ACE2 (refs. ^{1,2}), which is highly expressed in the aerodigestive tract, liver, kidneys and sex organs³. However, to bring virus and host membranes sufficiently close for fusion, the S protein requires proteolytic processing at the S1/S2 cleavage site to enable large conformational changes and expose subunit 2 (S2) harboring the matured membrane fusion machinery (Fig. 1a)^{4,5}. Transmembrane protease, serine 2 (TMPRSS2), which colocalizes with ACE2 at the cell membrane⁶, has been identified as the dominant proteolytic driver of S protein activation and SARS-CoV-2 infection of the aerodigestive tract^{1,7,8}.

TMPRSS2 belongs to the type 2 transmembrane serine protease (TTSP) family that comprises 19 surface-expressed trypsin-like serine proteases that normally participate in pericellular proteolytic cascades for degradative remodeling of the extracellular matrix⁹ and proteolytic activation of membrane proteins¹⁰, among other key epithelial homeostasis roles¹¹. TTSP family members experience complex posttranslational regulation. All TTSPs are glycoproteins expressed as single-chain proenzymes (or zymogens) that require proteolytic cleavage activation at a conserved (Arg/Lys)-(Ile/Val) peptide bond directly preceding their catalytic domains to mature their active site through interaction with the newly generated N terminus (Fig. 1a)¹². Some TTSPs can undergo autocleavage activation^{13,14} while others appear to require external activation by another protease^{15,16}. On activation, TTSPs can recruit endogenous, Kunitz-domain containing proteinaceous inhibitors that block substrate access and are themselves regulated to modulate active

TTSP activity^{17,18}. Dysregulated TTSP activity is a common feature that arises in cancers and results in increased tumor cell proliferation, invasiveness and metastasis^{19,20}, yet even basal activity levels of TMPRSS2, −4 (refs. ^{6,21,22}), −11d (refs. ^{22,23}) and −13 (refs. ^{22,23}) in normal tissues can be exploited by coronaviruses²⁴, influenza A^{25,26}, and influenza B²⁶ viruses for efficient activation and infection at the cell membrane. This family of proteases, and particularly TMPRSS2, therefore represent prime targets for therapeutic intervention to disable aggressive cancers and block initial viral infection. Antivirals that can effectively disable essential TTSP family members responsible for complete activation of the S protein^{8,27} may be drugs insensitive to new variant mutations owing to the conserved virus–host tropism with these proteases^{5,24,26}.

Despite this overwhelming motivation, TMPRSS2 and many other TTSPs have resisted therapeutic targeting campaigns in large part due to the difficulty associated with procuring active sources of recombinant enzyme conducive to rapid inhibitor screening and X-ray structural characterization of protein–inhibitor complexes. TTSPs are disulfide-rich, require site-specific proteolytic activation and pose inherent host cell toxicity challenges during recombinant production if proteolytic activity is uncontrolled during overexpression. Furthermore, TTSP inhibitors themselves pose unique pharmacological challenges. Since they are extracellular targets, the drugs targeting TTSPs may not need to cross the plasma membrane (if administered intravenously), but instead they must be exceedingly selective so as to avoid the concentrated pools of trypsin-like proteases in plasma responsible for coagulation cascades. TTSP drugs must also be biologically stable and have favorable biodistribution profiles to engage the target tissue with sufficient residence time to exert meaningful protective (or anti-tumor) effects.

¹Department of Molecular Oncology, British Columbia Cancer Research Institute, Vancouver, British Columbia, Canada. ²Department of Radiology, University of British Columbia, Vancouver, British Columbia, Canada. ³Structural Genomics Consortium, University of Toronto, Toronto, Ontario, Canada. ⁴Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, British Columbia, Canada. ⁵Department of Pharmacology and Toxicology, University of Toronto, Toronto, Ontario, Canada. ⁶Princess Margaret Cancer Centre, Toronto, Ontario, Canada. ⁷Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada. ⁸These authors contributed equally: Bryan J. Fraser, Serap Beldar. ✉e-mail: L.halabelian@utoronto.ca; Cheryl.arrowsmith@uhnresearch.ca; fbenard@bccrc.ca

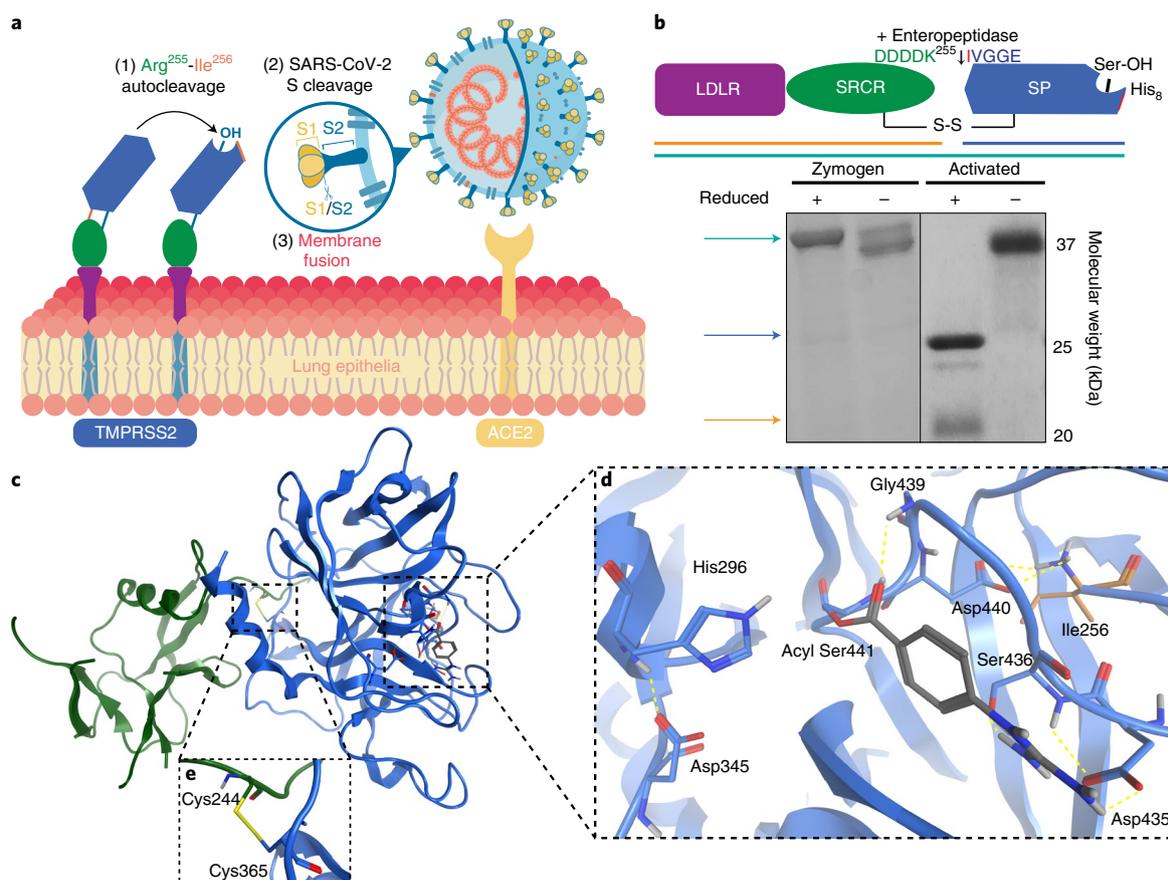


Fig. 1 | Engineered activation and structural characterization of stabilized TMPRSS2 ectodomain. **a**, Membrane-bound TMPRSS2 zymogen undergoes autocleavage activation at the Arg255-Ile256 peptide bond and the matured enzyme proteolytically processes SARS-CoV-2 Spike protein (magnified) docked to the ACE2 receptor (yellow) to drive membrane fusion. **b**, Engineered recombinant TMPRSS2 ectodomain (dasTMPRSS2) containing the LDLR-A domain, a Class A SRCR domain and a C-terminal trypsin-like S1 peptidase (SP) domain, features a DDDDK255 substitution to facilitate controlled zymogen activation. Purified, concentrated dasTMPRSS2 can self-activate. The noncatalytic (LDLR-A + SRCR) and catalytic (SP) chains are tethered by a disulfide bond and the activation status can be interrogated by SDS-PAGE under nonreducing and reducing (5% β -mercaptoethanol) conditions. Gel results are consistent with $n \geq 3$ independent biological experiments. **c**, X-ray crystal structure of dasTMPRSS2 pretreated with nafamostat, resulting in phenylguanidino acylation (gray sticks). **d**, Close-up view of the SP catalytic triad residues (His296, Asp345 and Ser441) and the postactivation Asp440:Ile256 salt bridge showing complete maturation of the protease. Polar contacts are shown as yellow dashed lines. **e**, The interdomain disulfide (Cys244-Cys365) maintains covalent attachment of the SRCR and SP domains.

As such, key structural insights are required to design TTSP-specific or pan-TTSP therapeutics.

Here we report a reliable method to produce highly active TMPRSS2, enabling the investigation of TMPRSS2-mediated SARS-CoV-2 S protein activation and associated inhibition by a panel of known clinical serine protease inhibitors. We determined the TMPRSS2 X-ray crystal structure, in a complex with nafamostat, a potent but nonselective trypsin-like serine protease inhibitor being investigated as a COVID-19 therapeutic. Our TMPRSS2 protein engineering and production strategy may be generally applicable for other TTSPs to enable their selective and rational targeting and uncover the molecular basis of their biological and pathological functions.

Results

Production and structure of activatable TMPRSS2 ectodomain.

TMPRSS2 is composed of an intracellular domain, a single-pass transmembrane domain and a biologically active ectodomain with three subdomains: a low-density lipoprotein receptor type-A (LDLR-A) domain, a Class A Scavenger Receptor Cysteine-Rich (SRCR) domain and a C-terminal trypsin-like serine peptidase (SP) domain

with a canonical Ser441-His296-Asp345 catalytic triad (Fig. 1a)²⁸. As TMPRSS2 is synthesized as a zymogen, it requires cleavage at a conserved Arg255-Ile256 peptide bond within its SRQSR255↓IVGGGE activation motif^{14,28}. We achieved this in high yield by replacing SRQSR255↓ with an enteropeptidase (TMPRSS15)-cleavable DDDDK255↓ sequence to prohibit auto-activation during overexpression, allowing purification of a secreted form of the full TMPRSS2 ectodomain zymogen from insect cells, analogous to a strategy used for the TTSP, matriptase (Methods)²⁹. Subsequent proteolytic activation with the addition of recombinant enteropeptidase afforded highly active, homogenous TMPRSS2 to milligram yields and was accordingly named the directed activation strategy TMPRSS2 (dasTMPRSS2, Fig. 1b and Extended Data Fig. 1a). Serendipitous efforts revealed that concentrated dasTMPRSS2 could self-activate without enteropeptidase addition across 6 hours despite this nonnatural sequence replacement, enabling simple size-exclusion purification of the activated species (Extended Data Fig. 1b,c). We determined the X-ray crystal structure of dasTMPRSS2 refined to 1.95 Å resolution after acylation of the catalytic Ser441 residue with nafamostat, a broad-spectrum synthetic serine protease inhibitor (released under Protein Data Bank (PDB) ID 7MEQ). We obtained clear electron

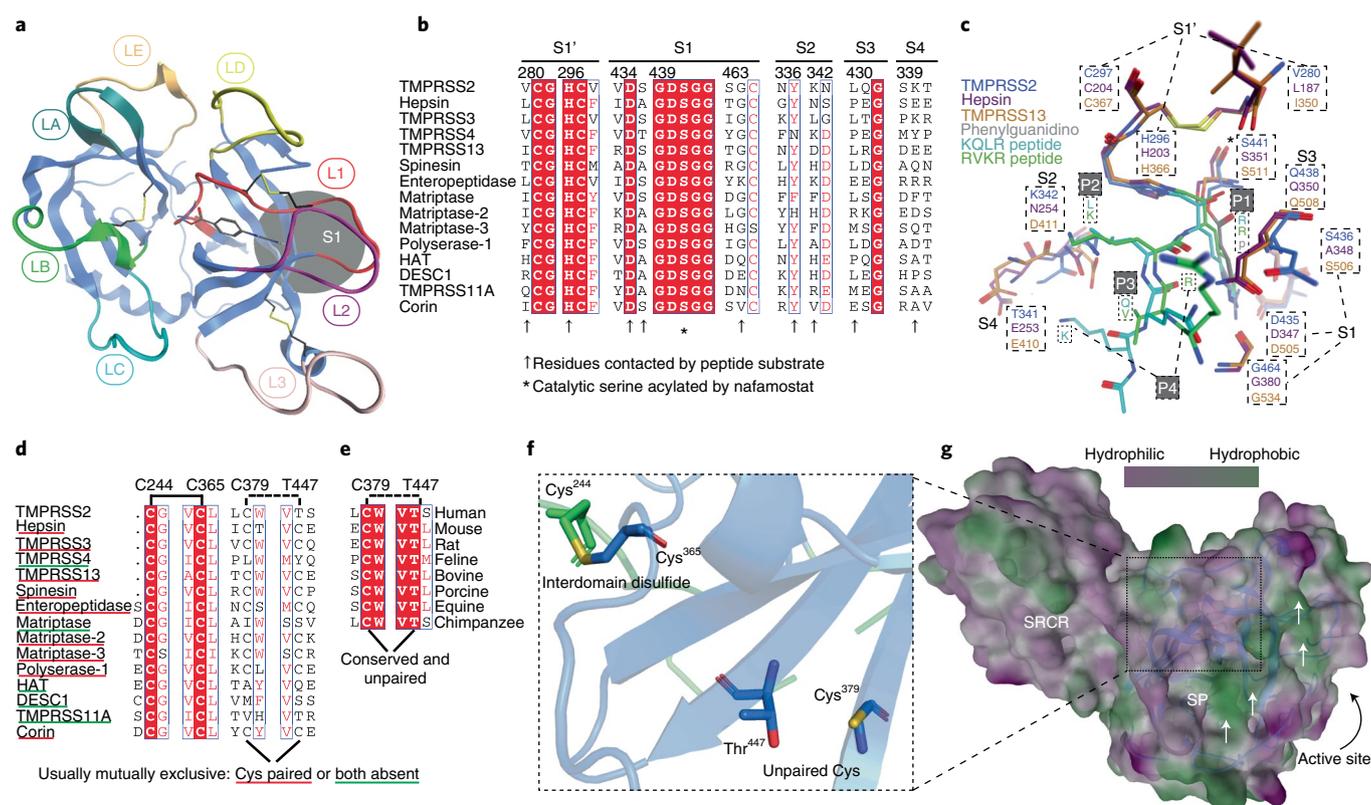


Fig. 2 | Molecular recognition of TMPRSS2 substrates is mediated through surface loops and potentially through an unpaired cysteine residue.

a, The substrate binding face of TMPRSS2 makes use of three disulfides (yellow sticks) and eight loops, Loops 1–3 (L1–L3) and LA–LE, to confer protein substrate specificity. Nafamostat (gray sticks) covalently bound to the catalytic Ser441 engages the S1 subsite of TMPRSS2 with residues from L1 and L2. **b**, Multiple sequence alignment of the TTSP family with protease subsites S1'–S4 residues highlighted. The catalytic serine residue acylated by nafamostat is denoted with an asterisk. The S2–S4 protease subsites show greater variability than S1' and S1 and confer divergent substrate specificity within the TTSP family. **c**, The subsites of TMPRSS2 (blue) superposed on the corresponding residues of hepsin (magenta, PDB 1Z8G) and TMPRSS13 (orange, PDB 6KD5) with their respective KQLR and decanoyl-RVKR covalent peptide ligands. **d,e**, A multiple sequence alignment of the human TTSP family members (**d**) shows complete conservation of the interdomain Cys244–Cys365 disulfide bond but identifies TMPRSS2 as uniquely possessing the only unpaired cysteine residue, Cys379, in the SP domain (**e**), which is conserved among mammalian TMPRSS2 orthologs. **f**, The highly conserved SRCR–SP interdomain disulfide is at the backside of the SP domain and is in close proximity to the unpaired Cys379 of TMPRSS2. **g**, MOE protein patch analysis of the TMPRSS2 ectodomain surface reveals a contiguous 360 Å² hydrophobic residue patch (green) highlighted with white arrows that is adjacent to the unpaired Cys379 at the backside of the SP domain.

density for residues 149–491 spanning the SRCR and SP domains but not residues 109–148 containing the flexible LDLR-A domain responsible for linking the protease to the plasma membrane, likely due to autoproteolytic cleavage of this domain before or during the crystallization process (Extended Data Fig. 1d). The engineered DDDDK255 activation motif was not resolved in the structure but rather terminated in an unstructured loop, consistent with matured TMPRSS13 (PDB 6KD5) and hepsin (PDB 1Z8G) structures containing their native activation motifs^{30,31}. The newly exposed N-terminal Ile256 of the SP domain formed a salt bridge with the side chain of Asp440 (Fig. 1d), confirming full maturation of the activation pocket and taken together with the Cys244–Cys365 interdomain disulfide, confirms that this structure represents a bioactive, stabilized form of the protease (Fig. 1e).

TMRSS2 has an accommodating substrate binding cleft. The TMPRSS2 SP domain is highly conserved with all TTSPs and conforms to the canonical chymotrypsin/trypsin fold with two six-stranded beta barrels converging to a central active site cleft harboring the catalytic triad (Fig. 2a)³². Divergent protein substrate specificity of these closely related proteases is conferred through highly variable, surface-exposed loops, denoted Loops

A–E and Loops 1–3 (Fig. 2a)³². Unique subsites formed on the face of the SP domain, S4–S3–S2–S1–S1'–S2'–S3'–S4' recognize substrate P4–P3–P2–P1↓P1'–P2'–P3'–P4' amino acid positions spanning the scissile bond (Fig. 2a,b). To rationally assign these subsites for TMPRSS2, we superposed the peptide-bound hepsin and TMPRSS13 SP domains (40.1 and 41.4% sequence identity of their SP domains, respectively) belonging to the same hepsin/TMPRSS subfamily as TMPRSS2. The S1 position of TMPRSS2 is occupied by the phenylguanidino moiety of nafamostat, forming salt bridges with the highly conserved Asp435, Ser436 and Gly464 residues in the same binding mode as the guanidino of P1 Arg residues observed in hepsin and TMPRSS13 (Figs. 1d and 2b,c and Supplementary Fig. 1a,b). The TMPRSS2 S2 subsite has a distinguishing Lys342 residue that likely confers a preference for small and/or electro-negative P2 substrates, similar to the S2 Lys in enteropeptidase that prefers P2 Asp residues³³. The S3 and S4 subsites appear open to accommodate various P3 and P4 amino acids and may make favorable receptor contacts with the respective Gln438 and Thr341 positions (Fig. 2b,c). N-terminal to the scissile bond, the buried S1' site appears to accept small, hydrophobic P1' residues. Overall, the TMPRSS2 active site appears capable of binding various substrate sequences with the strictest preference for the P1 and P2 positions.

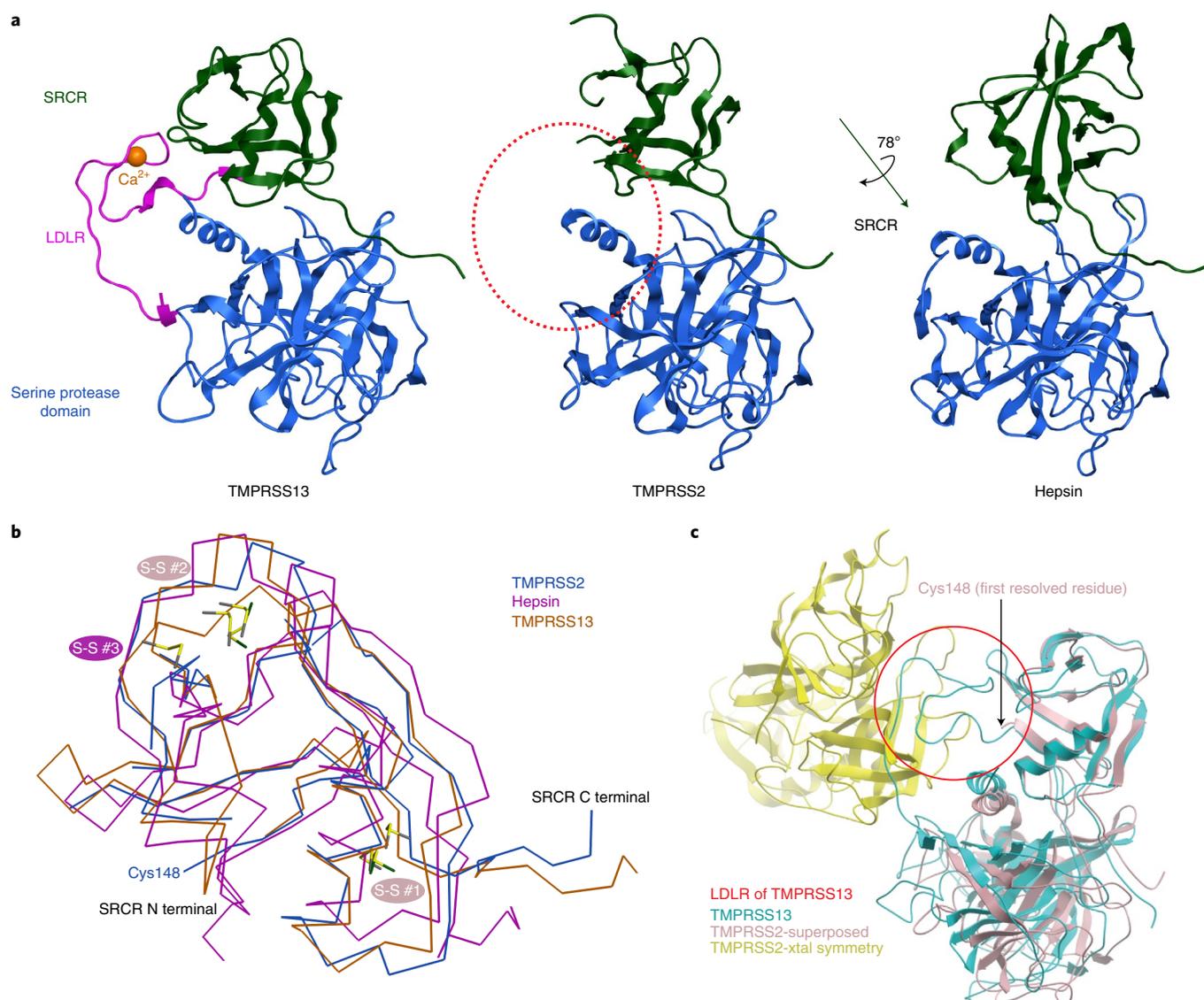


Fig. 3 | The stem domain of TMPRSS2 is structurally similar to TMPRSS13. a, Superposed serine protease domains (blue), and SRCR (green) of TMPRSS13, TMPRSS2 and hepsin. The LDLR domain of TMPRSS13 is shown in magenta with bound calcium in orange and the absence of the LDLR in TMPRSS2 is highlighted in red. **b**, Ribbon representation of the superimposed SRCR domains of TMPRSS2 (blue), Hepsin (magenta, PDB 1Z8G) and TMPRSS13 (orange, PDB 6KD5). **c**, Superposed models of TMPRSS13 in cyan and TMPRSS2 in salmon and its symmetry-related molecule in yellow. The LDLR domain that is present in TMPRSS13 structure would clash with the symmetry-related molecule in the TMPRSS2 crystal lattice, suggesting that the LDLR domain is not present in the crystal lattice.

Among TTSPs, the SP domain of TMPRSS2 uniquely possesses three disulfides and a single unpaired cysteine residue, Cys379 (Fig. 2d). In all other human TTSPs, this position forms a disulfide bond with an additional Cys at the residue equivalent of Thr447, or both cysteines are absent, resulting in three or four intradomain disulfides within their serine protease domains, respectively. This unpaired cysteine is conserved in mouse, rat, feline, bovine, porcine, equine and chimpanzee TMPRSS2 orthologs (Fig. 2e). Furthermore, the unpaired Cys379 is bordered by an expansive 360 Å² patch of exposed hydrophobic surface area in our structure that may serve as an interaction hub for TMPRSS2 binding partners (Fig. 2f,g).

The TMPRSS2 stem domain likely orients the SP domain. The SRCR domain is found enriched in proteins expressed at the surface of immune cells as well as in secreted proteins, and are thought to participate in protein–protein interactions and substrate

recruitment by TTSPs³⁴. Of the TTSPs, only hepsin and TMPRSS13 structures have been solved containing their complete ectodomains, motivating us to use these as templates for comparison to our structure. The Class A SRCR domain of TMPRSS2 is located on the backside of the SP domain away from the active site and is structurally similar to that of TMPRSS13 despite sharing only 19% sequence identity and both have two intradomain disulfides (Fig. 3a,b). These two SRCR domains adopt a compact, globular fold with similar orientations relative to their SP domains. The SRCR domain of hepsin (7.5% sequence identity) diverges substantially from TMPRSS2/13 with three intradomain disulfides and a tighter SRCR–SP association dominated by complementary electrostatic patches and buried surface area (Fig. 3a,b)³¹. Hepsin lacks an LDLR-A domain, evidently using the SRCR domain to connect the SP domain to the plasma membrane, potentially explaining its large 78° rotation relative to TMPRSS2/13. When hepsin/TMPRSS2/TMPRSS13 SRCR

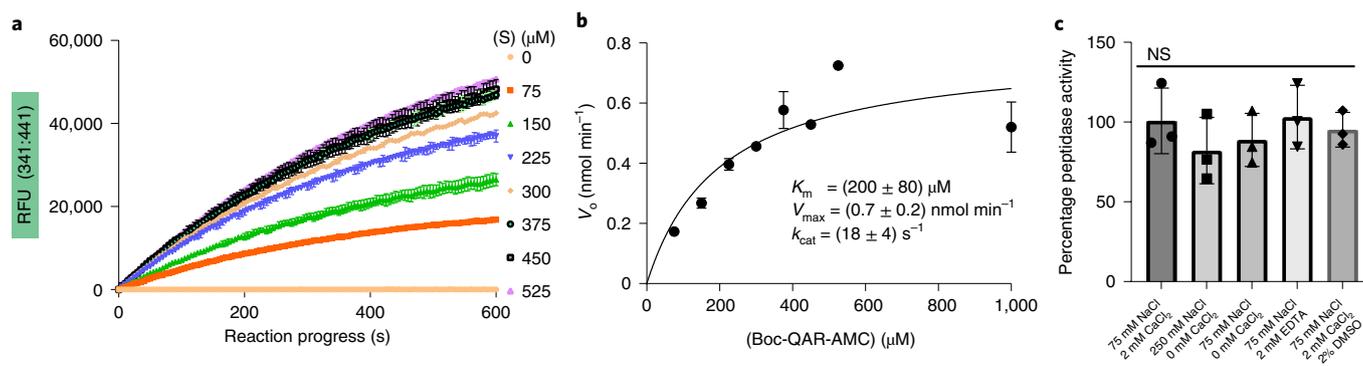


Fig. 4 | dasTMPRSS2 displays robust in vitro peptidase activity. **a**, The generic Boc-Gln-Ala-Arg-AMC fluorogenic peptide substrate is efficiently cleaved by 3.4 nM dasTMPRSS2 in Assay Buffer pH 8.0. Representative progress curves are shown in technical duplicate ($n=2$) with datapoints shown as mean values \pm s.d. **b**, Michaelis-Menten plot of initial reaction velocities for kinetic parameter estimation after curve fitting in GraphPad. Each reaction velocity was tabulated in technical quadruplet ($n=4$) and datapoints are shown as mean values \pm s.d. **c**, Relative peptidase activity (normalized to Assay Buffer; 25 mM Tris pH 8.0, 75 mM NaCl, 2 mM CaCl_2) calculated from initial reaction velocities under Tris pH 8.0 buffering conditions with varying concentrations of NaCl, CaCl_2 , EDTA and DMSO. Activity was measured in technical triplicate ($n=3$) and data are shown as mean values \pm s.d. No statistically significant differences from mean peptidase activity in Assay buffer were found (denoted as NS) in the indicated buffering conditions as determined by one-way analysis of variance for $n=3$ biologically independent samples examined over one independent experiment.

domains are superposed, they show high degrees of structural similarity owing to the conserved placement of disulfide bonds (Fig. 3b). These overall interdomain conformational differences may therefore play a role in orienting the SP domain relative to the plasma membrane as well as modulating activity through recognition or recruitment of partner proteins.

We observed no electron density for the LDLR-A domain of TMPRSS2, despite a similar ectodomain construct to that which afforded the TMPRSS13 crystal structure³¹. However, inspection of our X-ray crystal structure superposed on TMPRSS13 indicates that the LDLR domain that is present in the TMPRSS13 structure would clash with the symmetry-related molecule in the TMPRSS2 crystal lattice, suggesting that our LDLR domain is likely cleaved before or during crystallization (Fig. 3c and Extended Data Fig. 1d).

TMPRSS2 displays robust in vitro peptidase activity. To evaluate TMPRSS2 inhibitors and provide groundwork for future structure-activity relationship studies, we established in vitro proteolytic activity and inhibition assays. The generic TTSP fluorogenic peptide substrate Boc-Gln-Ala-Arg-7-aminomethylcoumarine (AMC) was rapidly cleaved by dasTMPRSS2, C terminal to Arg, thereby releasing AMC product and enabling initial reaction velocities (V_0) measurement within 60 seconds of enzyme addition (Fig. 4a and Methods). In Assay Buffer, dasTMPRSS2 had a K_m of $(200 \pm 80) \mu\text{M}$, V_{max} of $(0.7 \pm 0.2) \text{ nmol min}^{-1}$, k_{cat} of $(18 \pm 4) \text{ s}^{-1}$, k_{cat}/K_m of $(5.4 \pm 0.2) \mu\text{M}^{-1} \text{ min}^{-1}$ and specific activity at $(0.22 \pm 0.03) \mu\text{mol min}^{-1} \text{ mg}^{-1}$ enzyme purified to apparent homogeneity (Fig. 4b). This level of activity is high compared to other described recombinant TMPRSS2 enzyme^{16,22,35,36}, and enzyme activity was unaffected by the presence of Ca^{2+} , NaCl concentrations ranging 75–250 mM, EDTA and tolerant of 2% (v/v) dimethylsulfoxide (DMSO) that is encouraging for use in high-throughput inhibitor screening campaigns for new antiviral discovery (Fig. 4c).

Nafamostat rapidly acylates TMPRSS2 and slowly hydrolyzes. Nafamostat and camostat are serine protease inhibitors under investigation as anti-TMPRSS2 COVID-19 therapeutics (Clinical Trial.gov identifiers [NCT04583592](#) and [NCT04625114](#)). Both are reactive esters that form the same slowly reversible phenylguanidino covalent complex with the catalytic serine residue of trypsin-like serine proteases (Fig. 5a,b), as evidenced in the enteropeptidase-camostat costructure (Fig. 5c). Nafamostat and camostat dramatically

increased the apparent melting temperature ($T_{m,a}$) of dasTMPRSS2 by (25.5 ± 0.1) and $(24.8 \pm 0.3) ^\circ\text{C}$, respectively, as measured by differential scanning fluorimetry (DSF)³⁷ (Extended Data Fig. 2a) and was a key stabilizing feature to enable protein crystallization (Methods). Nafamostat demonstrated enhanced potency over camostat with half-maximum inhibitory concentration (IC_{50}) values of (1.4 ± 0.2) and $(9 \pm 4) \text{ nM}$, respectively, with 5 min of preincubation before the assay (Fig. 5d). However, IC_{50} values were time dependent and required further kinetic interrogation to assess their divergent potencies. Near simultaneous coaddition of these ester inhibitors with substrate allowed us to monitor the real-time acylation and plateau of enzymatic activity, analogous to kinetic investigations of camostat's acylation of enteropeptidase (Fig. 5e,f)³⁸. Nafamostat was 40-fold more potent than camostat with respective k_{inact}/K_i values of (1.44 ± 0.4) and $(0.035 \pm 0.002) \text{ nM}^{-1} \text{ min}^{-1}$ (Extended Data Fig. 2b). These results emphasize that single time-point IC_{50} values are insufficient for ranking mechanism-based, covalent inhibitors of this highly active protease in structure-activity relationship studies. As previously identified for matriptase, the nafamostat leaving group, 6-amidino-2-naphthol, fluoresces and can be used as a sensitive burst titrant to calculate the concentration of active protease by quantifying its production (Extended Data Fig. 2c,d and Methods)³⁶. The half-life of the phenylguanidino acyl-enzyme complex was (14.7 ± 0.4) hours as measured by the gradual rescue of dasTMPRSS2 peptidase activity after stoichiometric acylation with nafamostat (Extended Data Fig. 2e,f).

Noncovalent trypsin-like serine protease inhibitors benzamidine and sunflower trypsin inhibitor-1 (SFTI-I) were less potent with respective IC_{50} values of (120 ± 20) and $(0.4 \pm 0.2) \mu\text{M}$ (Fig. 4d), and K_i values of (80 ± 10) and $(0.4 \pm 0.2) \mu\text{M}$ (Extended Data Fig. 3a,b). The nafamostat leaving group 6-amidino-2-naphthol also competitively disabled dasTMPRSS2 activity with an IC_{50} of (1.6 ± 0.5) and K_i of $(1.1 \pm 0.3) \mu\text{M}$ (Extended Data Fig. 3c).

TMPRSS2 cleaves SARS-CoV-2 S protein at multiple sites. Cells expressing TMPRSS2 have been shown to efficiently cleave the S protein of SARS-CoV-1 at the S1/S2 (SLLR667↓) cleavage site and additional peripheral K/R residues to induce the necessary conformational changes leading to virus-host fusion at the plasma membrane^{5,39} (Figs. 1a and 6a,b). TMPRSS2 can activate S protein in virus-producing cells (termed *cis*-cleavage) as well as naïve host target cells activating S protein at the cell surface (*trans*-cleavage)³⁹.

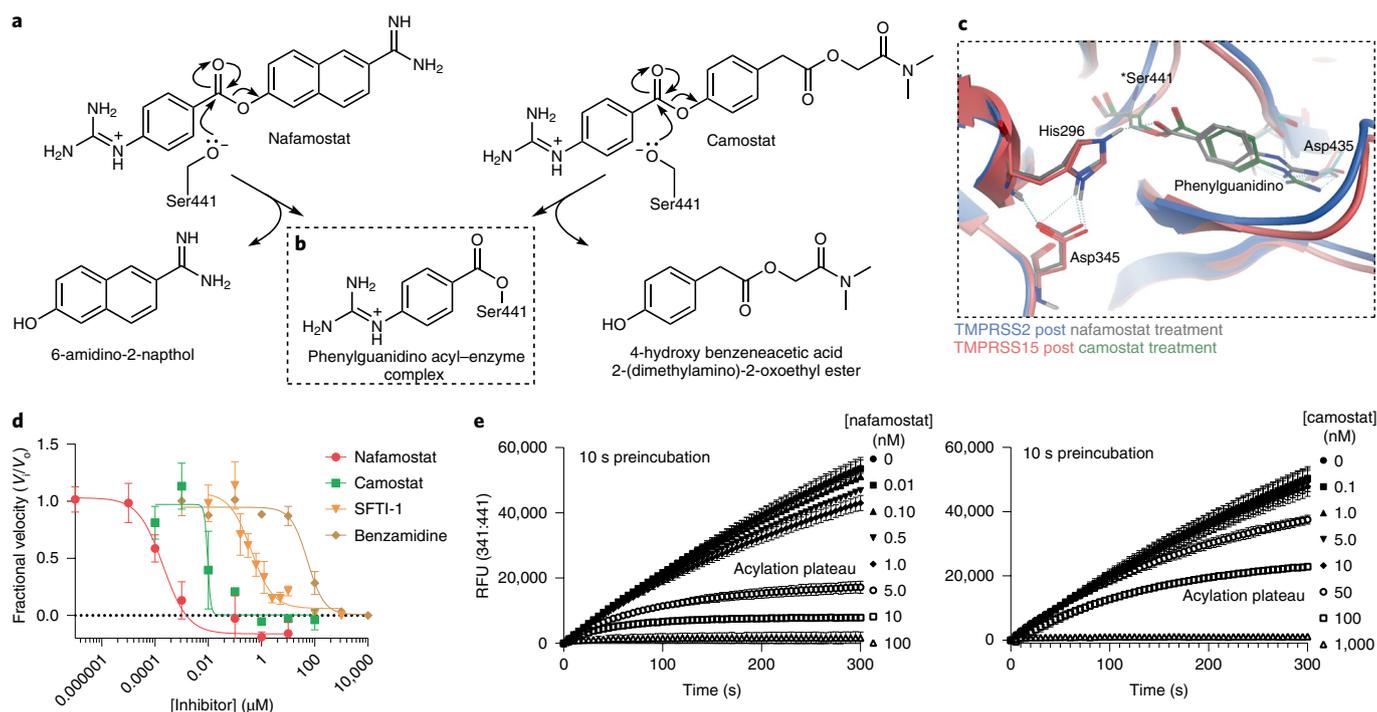


Fig. 5 | dasTMPRSS2 peptidase activity is blocked by clinical protease inhibitors. **a**, Nafamostat and camostat are attacked by the catalytic Ser441 of TMPRSS2, with respective leaving groups 6-amidino-2-naphthol and 4-Hydroxy benzeneacetic acid 2-(dimethylamino)-2-oxoethyl ester. **b**, A common phenylguanidino acyl-enzyme complex is formed after nafamostat or camostat treatment that specifically engages the S1 protease subsite. **c**, TMPRSS2 (blue) and TMPRSS15 (enteropeptidase, salmon; PDB ID 6ZOV) after nafamostat and camostat treatment, respectively. Catalytic triad residues are highlighted in bold and the conserved S1 subsite residue Asp435 shown. **d**, Clinical protease inhibitors preincubated for 5 min with dasTMPRSS2 block peptidase activity against Boc-QAR-AMC fluorogenic substrate with varying inhibitory potencies. Data are shown as mean \pm s.d. and were performed in technical and biological duplicate (total $n = 4$) and curve fit for absolute IC_{50} in GraphPad. **e**, Reaction progress curves of residual dasTMPRSS2 peptidase activity after 10 s preincubation with the indicated concentrations of nafamostat and camostat inhibitors in the presence of 100 μ M Boc-QAR-AMC substrate. Plateaus within progress curves demonstrate time-dependent acylation resulting in complete inhibition of dasTMPRSS2 peptidase activity. Data are shown as mean \pm s.d. in technical duplicate ($n = 2$) and results were consistently obtained across $n = 3$ biological experiments.

However, the unavailability of validated recombinant TMPRSS2 enzyme has led to challenges in unambiguously mapping TMPRSS2 cleavage sites in SARS-CoV-1 and -2, as the typical workflow with cell-based systems and S protein western blotting produces complex fragment patterns and do not apply unbiased, discovery-based mass spectrometry approaches to localize residue cleavages. Notably, the 'S2' site' (PTKR797↓ for SARS-CoV-1 and predicted PSKR815↓ for SARS-CoV-2), which was identified as a trypsin cleavage site for SARS-CoV-1 (ref. 3), has been proposed as an important TMPRSS2 cleavage site for SARS-CoV-2 by analogy, but no biochemical evidence is currently available to support cleavage at this residue.

The most significant cleavage site for SARS-CoV-2 infection is the multibasic S1/S2 sequence (RRAR685↓, Fig. 6a,b), which was initially hypothesized to confer preferential processing by intracellular furin protease²⁷, possibly explaining the severe and systemic COVID-19 disease symptoms. This hypothesis was supported by cell infection studies showing that multibasic, peptide-based furin inhibitors prevented S1/S2 cleavage and attenuated viral entry⁷. However, subsequent studies showed that these inhibitors are promiscuous and disable multiple cell surface-expressed proteases that process multibasic substrates in addition to furin, and more selective furin inhibitors cannot fully abrogate S activation⁴. Furthermore, furin-deficient cells can still generate S1/S2 cleaved virus⁸. Propagation of live SARS-CoV-2 virus in TMPRSS2-deficient but not furin-deficient cell lines results in a loss of the multibasic S1/S2 site⁸, attenuating viral infectivity toward TMPRSS2⁺ cells. We sought to characterize TMPRSS2's proteolytic activity toward

S1/S2 by incubating recombinant dasTMPRSS2 and/or furin protease with stabilized SARS-CoV-2 S protein with S1/S2 knocked out (RRAR⁶⁸⁵ → GSAS⁶⁸⁵; HexaPro construct) or with S1/S2 intact (denoted HexaFurin, Fig. 6b). As expected from previous studies using recombinant, S1/S2 intact S protein, HexaFurin sustained partial S1/S2 cleavage during production in human embryonic kidney 293 cells presumably due to endogenously expressed furin⁴⁰ (Fig. 6c). Incubation with recombinant furin for 16 hours converted the remaining intact HexaFurin to a product yielding S1 and S2 fragments by SDS-PAGE (Fig. 6c) and was unable to cleave HexaPro due to the S1/S2 KO (Extended Data Fig. 4a).

In contrast, using both the HexaFurin and HexaPro constructs, we observed that dasTMPRSS2 could cleave the S protein at three distinct sites with variable efficiency (Fig. 6d–g). HexaFurin was cleaved to only the S1 and S2 fragments within 5 min of dasTMPRSS2 addition (Fig. 6d), demonstrating the S1/S2 site was best recognized by TMPRSS2 across a minimal incubation. This is the first biochemical evidence that TMPRSS2 recognizes the SARS-CoV-2 S1/S2 site with high efficiency. HexaPro, lacking the S1/S2 cleavage site, was cleaved across 30 min to produce a larger 150 kDa fragment X and 70 kDa fragment Y when analyzed under nonreducing conditions, motivating us to name this putative site as an 'X/Y' cleavage site (Fig. 6e). Running the same 30 min digest sample under reducing SDS-PAGE revealed an additional putative cleavage site that is spanned by two cysteine residues participating in a disulfide bond, splitting the observed fragment X into daughter fragments Q at 120 kDa and N + receptor binding domain (RBD) at 35 kDa

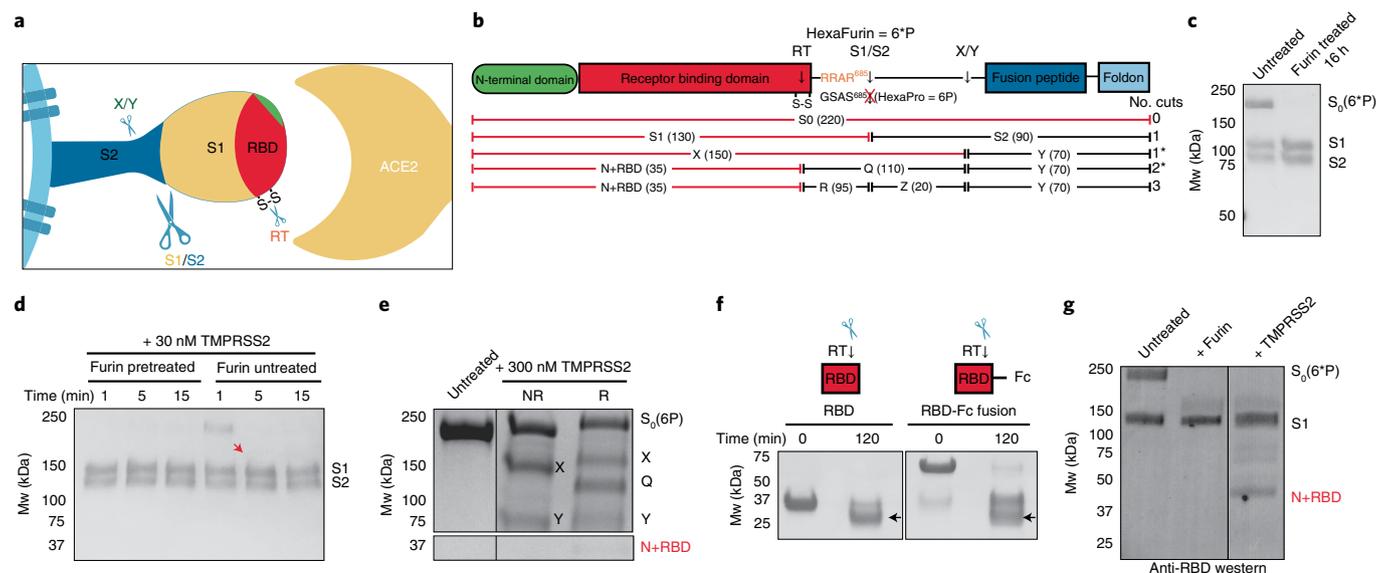


Fig. 6 | dasTMPRSS2 efficiently cleaves recombinant SARS-CoV-2 Spike protein constructs at multiple sites. **a**, The trimeric Spike protein (depicted here in its monomer form) is presented at the surface of the viral membrane and binds the transmembrane human ACE2 receptor through its RBD. Three distinct TMPRSS2 protease cleavage sites are predicted (indicated by scissors). **b**, Schematic map of known (S1 and S2) and inferred protein fragments after TMPRSS2 treatment, as derived from protein bands (**c–g**) with the indicated approximate molecular weights (in parentheses) produced as a result of 0–3 protease cleavage events. Recombinant S protein is cleaved at the S1/S2 site, at an unknown site within S2 termed the X/Y cleavage site and within RBD termed the RBD–TMPRSS2 (RT) site that is spanned by a disulfide bond. Fragments containing RBD are highlighted in red, and cut patterns denoted with an asterisk are only detected using S1/S2 KO S protein, HexaPro. **c**, HexaFurin S protein is converted to S1 and S2 band fragments after 16 h incubation with 10 U of recombinant furin protease (NEB). Mw, molecular weight. **d**, dasTMPRSS2 fully converts untreated HexaFurin to S1 and S2 fragments within 5 min. **e**, HexaPro S protein treated with 300 nM TMPRSS2 for 30 min is cleaved at both the X/Y and RBD–TMPRSS2 cleavage sites, with the latter cleavage event apparent only under reducing (R) SDS–PAGE conditions rather than nonreducing (NR). **f**, Recombinant S protein constructs containing only the RBD and an RBD–Fc fusion are cleaved at the RBD–TMPRSS2 cleavage site after 120 min of TMPRSS2 incubation. **g**, Western blot analysis of HexaFurin S protein digestions with an RBD-specific primary antibody shows the expected S1 band fragment, the N + RBD fragment and several intermediate sized fragments also containing RBD. All gels and blots are consistent with $n = 3$ independent biological experiments.

(Fig. 6e). Notably, S protein band fragments ran at a slightly higher apparent molecular weight under reducing conditions, likely reflecting increased flexibility and reduced gel mobility. We localized this cleavage site hidden by a disulfide bond to near the C-terminal end of the RBD using two S protein constructs containing only RBD. TMPRSS2 treatment produced a common cleavage product (highlighted with an arrow, Fig. 6f), motivating us to name this putative cleavage site the RBD–TMPRSS2 cleavage site.

To visualize all three cleavage sites simultaneously, we treated HexaFurin for 30 min with dasTMPRSS2 and analyzed digests with a western blot using an antibody directed toward the S protein RBD (Fig. 6g). At least seven bands were observed on reducing SDS–PAGE that are consistent with the three proposed cleavage sites, dominated by rapid, initial S1/S2 cleavage (Extended Data Fig. 4b). The western banding patterns observed (S1/S2, X/Y and RBD–TMPRSS2 cleavages) are consistent with cellular studies monitoring SARS-CoV-1 S protein processing by TMPRSS2 that produced many fragments and enabled shedding of the S1 fragment³⁹, which acts as an immune decoy to compromise the ability of neutralizing S protein antibodies.

Stoichiometric amounts of nafamostat completely blocked dasTMPRSS2-mediated HexaPro activation over 2 h (Extended Data Fig. 4c) and are consistent with this drug's ability to potently block SARS-CoV-2 pseudovirus entry to TMPRSS2 + Calu-3 (ref. 41) and Caco-2 (refs. 1,22) lung and colorectal cancer cell lines. Bromhexine hydrochloride, another agent previously under clinical investigation for anti-TMPRSS2 COVID-19 therapy (NCT04273763)⁴², showed no inhibition in either the peptidase or HexaPro cleavage assay formats (Extended Data Fig. 4d,e), corroborating reports of its

ineffectiveness in blocking SARS-CoV-2 pseudovirus entry⁴³ and further underscores the need for new, selective TMPRSS2 inhibitors.

Discussion

We have produced and characterized a source of TMPRSS2 enzyme that will enable future inhibitor development as antivirals and thorough molecular interrogation of coronavirus and influenza virus activation, owing to TMPRSS2's critical and widespread role in viral tropism. Although nafamostat potently neutralizes TMPRSS2 activity, it is nonselective and disables trypsin-like serine proteases involved in coagulation such as plasmin, FXa and FXIIa, as well as other TTSPs through its generic arginine-like engagement with the S1 subsite^{44,45}. In line with this promiscuity and nonspecific acylation, nafamostat is biologically unstable with a half-life of 8 minutes (NCT04418128 and NCT04473053). Camostat operates similarly to nafamostat but with poorer TMPRSS2 inhibitory potency and is hydrolyzed to two additional products in vivo, each with further compromised potencies²². These selectivity and consequential stability issues challenge the abilities of these drugs to adequately engage active TMPRSS2 at pharmacological concentrations in vivo, despite their favorable in vitro antiviral potency^{41,43}. Clearly, selective and biologically stable drugs for TMPRSS2 must be further explored and may be achieved through inhibitors engaging the more TMPRSS2-specific S2, S3 and S4 protease subsites identified in our crystal structure.

Our demonstration that TMPRSS2 can efficiently cleave the multibasic S1/S2 site of the S protein supports the notion that instead of conferring furin dependence, the virulent properties of this site on the S protein may derive from recognition and cleavage

by airway-expressed TTSPs, which is supported by the demonstrated roles that TMPRSS4 (refs. ^{6,21}), TMPRSS11d (refs. ^{13,22,23}) and TMPRSS13 (refs. ^{22,23}), which colocalize with ACE2 (ref. ⁶), play in enabling SARS-CoV-2 infection across various tissues.

The X/Y and RBD-TMPRSS2 cleavage sites proposed here may have significant roles in TMPRSS2-mediated SARS-CoV-2 activation and infection of cells. In future studies, we will identify the precise residue at which cleavage occurs, enabling enzymatic and cell-based KO experiments to interrogate their biological function.

Unfortunately, many cellular studies of TTSPs^{1,22,23,46,47}, with some exceptions^{13,48,49}, have not interrogated the zymogen activation status to ensure their activity. This limits what conclusions can be drawn on the biological significance of each protease target in models of viral infectivity, as the (typically transfected) cell lines in use may simply not provide the requisite activators, protein partners or physiological conditions conducive to a matured and active TTSP. This variable can be removed by quantifying the relative abundance of the catalytic chain (under reducing conditions) through SDS-PAGE or western blotting and/or developing activity-based, TTSP-selective chemical probes to accurately dissect the biological activity of each protease.

Our characterization and X-ray crystal structure of dasTMPRSS2 motivates interrogation of a mechanism by which the native, membrane-bound enzyme could be (auto)proteolytically processed peripheral to the activation motif and thereby shed as a soluble enzyme into the extracellular space, analogous to matriptase³⁰. A study using TMPRSS2-specific antibodies has reported detection of a secreted enzyme product in the airways that was proposed to play a functional role in pericellular proteolytic activation¹⁴. Due to the disulfide-linked nature of activated TMPRSS2 outlined here (Fig. 1b), former studies may have mischaracterized the catalytic (or noncatalytic depending on the epitope recognized by the antibody) subunit as a shed SP domain when it would instead resolve to the intact species under nonreducing conditions (Fig. 1b). Taken together, a biochemical characterization of these secreted species is required to clarify their activation status and subunit organization, as a shed form of TMPRSS2 localized to the extracellular milieu would have profound pathobiological and therapeutic targeting implications.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41589-022-01059-7>.

Received: 24 August 2021; Accepted: 9 May 2022;

Published online: 8 June 2022

References

- Hoffmann, M. et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **181**, 271–280 (2020).
- Hörmich, B. F. et al. SARS-CoV-2 and SARS-CoV spike-mediated cell-cell fusion differ in their requirements for receptor expression and proteolytic activation. *J. Virol.* **95**, e0002-21 (2021).
- Liu, J. et al. SARS-CoV-2 cell tropism and multiorgan infection. *Cell Discov.* **7**, 2–5 (2021).
- Tang, T. et al. Proteolytic activation of SARS-CoV-2 spike at the S1/S2 boundary: potential role of proteases beyond furin. *ACS Infect. Dis.* **7**, 264–272 (2021).
- Reinke, L. M. et al. Different residues in the SARS-CoV spike protein determine cleavage and activation by the host cell protease TMPRSS2. *PLoS ONE* <https://doi.org/10.1371/journal.pone.0179177> (2017).
- Wruck, W. & Adjaye, J. SARS-CoV-2 receptor ACE2 is co-expressed with genes related to transmembrane serine proteases, viral entry, immunity and cellular stress. *Sci. Rep.* **10**, 21415 (2020).
- Bestle, D. et al. TMPRSS2 and furin are both essential for proteolytic activation of SARS-CoV-2 in human airway cells. *Life. Sci. All.* **3**, e202000786 (2020).
- Sasaki, M. et al. SARS-CoV-2 variants with mutations at the S1 / S2 cleavage site are generated in vitro during propagation in TMPRSS2-deficient cells. *PLoS Pathog.* **17**, e1009233 (2021).
- Mukai, S. et al. Matriptase and MET are prominently expressed at the site of bone metastasis in renal cell carcinoma: immunohistochemical analysis. *Hum. Cell* **28**, 44–50 (2015).
- Heurich, A. et al. TMPRSS2 and ADAM17 cleave ACE2 differentially and only proteolysis by TMPRSS2 augments entry driven by the severe acute respiratory syndrome coronavirus spike protein. *J. Virol.* **88**, 1293–1307 (2014).
- List, K. et al. Epithelial integrity is maintained by a matriptase-dependent proteolytic pathway. *Am. J. Pathol.* **175**, 1453–1463 (2009).
- Tseng, C. et al. Matriptase shedding is closely coupled with matriptase zymogen activation and requires de novo proteolytic cleavage likely involving its own activity. *PLoS ONE* <https://doi.org/10.1371/journal.pone.0183507> (2017).
- Zhang, C. et al. Intracellular autoactivation of TMPRSS11A, an airway epithelial transmembrane serine protease. *J. Biol. Chem.* **295**, 12686–12696 (2020).
- Afar, D. E. H. et al. Catalytic cleavage of the androgen-regulated TMPRSS2 protease results in its secretion by prostate and prostate cancer epithelia. *Cancer Res.* **61**, 1686–1692 (2001).
- Lu, D., Yuan, X., Zheng, X. & Sadler, J. E. Bovine proenteropeptidase is activated by trypsin, and the specificity of enteropeptidase depends on the heavy chain. *J. Biol. Chem.* **272**, 31293–31300 (1997).
- Lucas, J. M. et al. The androgen-regulated protease TMPRSS2 activates a proteolytic cascade involving components of the tumor microenvironment and promotes prostate cancer metastasis. *Cancer Discov.* **4**, 1310–1325 (2014).
- Lee, M. S. et al. Autoactivation of matriptase in vitro: Requirement for biomembrane and LDL receptor domain. *Am. J. Physiol. - Cell Physiol.* **293**, 95–105 (2007).
- Tsai, C. H. et al. HAI-2 suppresses the invasive growth and metastasis of prostate cancer through regulation of matriptase. *Oncogene* **33**, 4643–4652 (2014).
- Larzabal, L. et al. Overexpression of TMPRSS4 in non-small cell lung cancer is associated with poor prognosis in patients with squamous histology. *Br. J. Cancer* **105**, 1608–1614 (2011).
- Ko, C. J. et al. Androgen-induced TMPRSS2 activates matriptase and promotes extracellular matrix degradation, prostate cancer cell invasion, tumor growth, and metastasis. *Cancer Res.* **75**, 2949–2960 (2015).
- Zang, R. et al. TMPRSS2 and TMPRSS4 promote SARS-CoV-2 infection of human small intestinal enterocytes. *Sci. Immunol.* **5**, eabc3582 (2020).
- Hoffmann, M. et al. Camostat mesylate inhibits SARS-CoV-2 activation by TMPRSS2-related proteases and its metabolite GBPA exerts antiviral activity. *E. Bio. Medicine* **65**, 103255 (2021).
- Kishimoto, M. et al. TMPRSS11d and TMPRSS13 activate the SARS-CoV-2 spike protein. *Viruses* **13**, 384 (2021).
- Matsuyama, S. et al. Efficient activation of the severe acute respiratory syndrome coronavirus spike protein by the transmembrane protease TMPRSS2. *J. Virol.* **84**, 12658–12664 (2010).
- Limburg, H. et al. TMPRSS2 is the major activating protease of Influenza A virus in primary human airway cells and influenza B virus in human type II pneumocytes. *J. Virol.* **93**, e00649-19 (2019).
- Botzcher, E. et al. Proteolytic activation of influenza viruses by serine proteases TMPRSS2 and HAT from human airway epithelium. *J. Virol.* **80**, 9896–9898 (2006).
- Hoffmann, M., Kleine-Weber, H. & Pohlmann, S. A multibasic cleavage site in the spike protein of SARS-CoV-2 is essential for infection of human lung cells. *Mol. Cell* **78**, 779–784 (2020).
- Chen, Y. et al. TMPRSS2, a serine protease expressed in the prostate on the apical surface of luminal epithelial cells and released into semen in prostasomes, is misregulated in prostate cancer cells. *Am. J. Pathol.* **176**, 2986–2996 (2010).
- Yamasaki, Y., Satomi, S., Murai, N., Tsuzuki, S. & Fushiki, T. Inhibition of membrane-type serine protease 1/matriptase by natural and synthetic protease inhibitors. *J. Nutr. Sci. Vitaminol.* **49**, 27–32 (2003).
- Herter, S. et al. Hepatocyte growth factor is a preferred in vitro substrate for human hepsin, a membrane-anchored serine protease implicated in prostate and ovarian cancers. *Biochem. J.* **136**, 125–136 (2005).
- Ohno, A. et al. Crystal structure of inhibitor-bound human MSPL that can activate high pathogenic avian influenza. *Life Sci. Alliance* **4**, 1–11 (2021).
- Perona, J. J. & Craik, C. S. Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold. *J. Biol. Chem.* **272**, 29987–29990 (1997).
- Boulware, K. T. & Daugherty, P. S. Protease specificity determination by using cellular libraries of peptide substrates (CLiPS). *Proc. Natl Acad. Sci. USA* **103**, 7583–7588 (2006).

34. Yap, N. V. L., Whelan, F. J., Bowdish, D. M. E. & Golding, G. B. The evolution of the scavenger receptor cysteine-rich domain of the class A scavenger receptors. *Front. Immunol.* **6**, 1–9 (2015).
35. Meyer, D., Sielaff, F., Hammami, M. & Otcher-Friebertsh, E. B. Identification of the first synthetic inhibitors of the type II transmembrane serine protease TMPRSS2 suitable for inhibition of influenza virus activation. *Biochem. J.* **343**, 331–343 (2013).
36. Nimishakavi, S., Raymond, W. W., Gruenert, D. C. & Caughey, G. H. Divergent inhibitor susceptibility among airway lumen-accessible tryptic proteases. *PLoS ONE* <https://doi.org/10.1371/journal.pone.0141169> (2015).
37. Niesen, F. H., Berglund, H. & Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* **2**, 2212–2221 (2007).
38. Sun, W. et al. Targeting enteropeptidase with reversible covalent inhibitors to achieve metabolic benefits. *J. Pharmacol. Exp. Ther.* **379**, 510–521 (2020).
39. Glowacka, I. et al. Evidence that TMPRSS2 activates the severe acute respiratory syndrome coronavirus spike protein for membrane fusion and reduces viral control by the humoral immune response. *J. Virol.* **85**, 4122–4134 (2011).
40. Gobeil, S. M. et al. D614G mutation alters SARS-CoV-2 spike conformation and enhances protease cleavage at the S1 / S2 junction. *Cell Rep.* **34**, 108630 (2021).
41. Hoffmann, M., Schroeder, S., Kleine-weber, H., Müller, M. A. & Drosten, C. Nafamostat mesylate blocks activation of SARS-CoV-2: new treatment option for COVID-19. *Antimicrob. Agents Chemother.* **64**, 19–21 (2020).
42. Li, T. et al. Bromhexine hydrochloride tablets for the treatment of moderate COVID-19: an open-label randomized controlled pilot study. *Clin. Transl. Sci.* **13**, 1096–1102 (2020).
43. Shrimp, J. H. 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.* **3**, 997–1007 (2020).
44. Hitomi, Y., Ikari, N. & Fujii, S. Inhibitory effect of a new synthetic protease inhibitor (FUT-175) on the coagulation system. *Haemostasis* **15**, 164–168 (1985).
45. Muto, S., Imai, M. & Asano, Y. Mechanisms of hyperkalemia caused by nafamostat mesilate. *Gen. Pharmacol.* **26**, 1627–1632 (1995).
46. Chen, Y. et al. A high-throughput screen for TMPRSS2 expression identifies FDA-approved compounds that can limit SARS-CoV-2 entry. *Nat. Commun.* **12**, 3907 (2021).
47. Ko, C. et al. Inhibition of TMPRSS2 by HAI-2 reduces prostate cancer cell invasion and metastasis. *Oncogene* **39**, 5950–5963 (2020).
48. Wang, H., Li, S., Wang, J., Chen, S. & Sun, X. N-glycosylation in the protease domain of trypsin-like serine proteases mediates calnexin-assisted protein folding. *eLife* <https://doi.org/10.7554/eLife.35672> (2018).
49. Wang, H. et al. Distinct roles of N-glycosylation at different sites of corin in cell membrane targeting and ectodomain shedding. *J. Biol. Chem.* **290**, 1654–1663 (2015).
50. Tanabe, L. M. & List, K. The role of type II transmembrane serine protease-mediated signaling in cancer. *FEBS J.* **284**, 1421–1436 (2017).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2022

Methods

Construct design and cloning. A construct encoding residues 109–492 comprising soluble TMPRSS2 ectodomain was amplified by two PCR fragments (Addgene plasmid no. 53887) and subcloned into the pFHMS-PLIC C donor plasmid by the ligation independent cloning (LIC) method. The final construct contained a N-terminal honeybee melittin signal sequence peptide and C-terminal His₆-tag (Fig. 1b). Mutations targeting the activation sequence SSQSR255↓IVGGG (arrow indicates the cleavage site) were implemented to replace the SRQSR255 residues with an enteropeptidase-cleavable DDDDK255 graft with two sets of primer pairs (Supplementary Table 1) generating mutations for S251D/R252D/Q253D/S254D/R255K. Our engineered dasTMPRSS2 protein expression construct is available on Addgene (plasmid no. 176412). Plasmid transfer vector containing the TMPRSS2 gene was transformed into *Escherichia coli* DH10Bac cells (Thermo Fisher, catalog no. 10361012) to generate recombinant viral Bacmid DNA. Sf9 cells were transfected with Bacmid DNA using JetPrime transfection reagents (PolyPlus Transfection Inc., catalog no. 114-01) according to the manufacturer's instructions, and recombinant baculovirus particles were obtained and amplified from P1 to P2 viral stocks. Recombinant P2 viruses were used to generate suspension culture of baculovirus infected insect cells for scaled-up production of TMPRSS2.

The SARS-CoV-2 Spike ectodomain HexaPro construct was a gift from J. McLellan⁵¹, and the S1/S2 site was restored (GSAS⁶⁸⁵ → RRAR) through site-directed mutagenesis with primers in Supplementary Table 1 (HexaFurin construct).

Baculovirus mediated dasTMPRSS2 protein production in Sf9 insect cells.

Sf9 cells were grown in I-Max Insect Medium (Wisent Biocenter, catalog no. 301-045-LL) to a density of 4×10^6 cells per ml and infected with 20 ml^{-1} of suspension culture of baculovirus infected insect cells before incubation on an orbital shaker (145 r.p.m., 26 °C).

dasTMPRSS2 protein purification. Cell culture medium containing the final secreted protein product AA-(TMPRSS2(109–492))-EFVHHHHHHHH was collected by centrifugation (20 min, 10 °C, 6,000g) 4–5 d postinfection when cell viability dropped to 55–60%. Media was adjusted to pH 7.4 by addition of concentrated PBS stock, then supplemented with 15 ml^{-1} settled Ni-NTA resin (Qiagen) at a scale of 12l. Three batch Ni-NTA purifications were used to capture protein, with each round requiring shaking in 2l flasks for 2 h at 16 °C (110 r.p.m.), collection by centrifugation (5 min 1,000g) and then transferring to a gravity flow column. Beads were washed with 3 column volumes of ice-cold PBS before elution with PBS supplemented with 500 mM imidazole. Elution samples were concentrated to 4.5 mg ml^{-1} using 30 kDa molecular weight cutoff (MWCO) Amicon filters and zymogen activation was achieved by dialyzing protein 1:1,000 against Assay Buffer (25 mM Tris pH 8.0, 75 mM NaCl, 2 mM CaCl₂) at room temperature across 6 h (Extended Data Fig. 1a). Activated samples were exchanged to size-exclusion chromatography buffer (50 mM Tris pH 7.5, 250 mM NaCl), spun down at 17,000g, then loaded to a Superdex 75 gel filtration column. Fractions spanning the dominant peak eluting at 80 ml (Extended Data Fig. 1b,c) were evaluated for appropriate banding on reducing SDS–PAGE before pooling and concentrating. For dasTMPRSS2 enzyme samples, 2 μl aliquots of 10,000× enzyme assay stocks (34 μM) were prepared by concentrating protein to 1.6 mg ml^{-1} in Storage Buffer (50 mM Tris pH 7.5, 250 mM NaCl, 25% glycerol), then flash-frozen in liquid nitrogen and stored at –80 °C until thawed immediately before use for each enzyme assay to minimize autolysis and maintain reproducible enzyme concentrations.

HexaPro and HexaFurin production and purification. Expi293F cells (Life Technologies catalog no. A1435102) were transiently transfected with expression plasmid encoding HexaPro/Furin using FectoPro transfection reagent (Polyplus transfection SA, catalog no. 116-010) with 5 mM sodium butyrate being added at the time of transfection (Sigma, catalog no. 303410). After 4–5 days posttransfection time, the cell culture was collected, supernatant cleared by centrifugation and the pH was adjusted by adding 10× Buffer (50 mM Tris pH 8.0, 150 mM NaCl). Secreted protein was captured by two rounds of batch absorption with 4 ml^{-1} of pre-equilibrated Ni Sepharose beads (GE Healthcare, catalog no. 17-5318-01). The bound beads were transferred to gravity flow column and sequentially washed with 30 column volumes of Wash Buffer (50 mM HEPES 7.5, 300 mM NaCl, 5% glycerol), followed by Wash Buffer supplemented with 25 mM imidazole. Protein was eluted in Elution Buffer (Wash Buffer with 250 mM imidazole) and concentrated using Amicon Ultra Centrifugal Filter Units, 15 ml, 100 kDa (Millipore Sigma catalog no. UFC910024) before size-exclusion chromatography purification using a Superose 6 Increase 10/300 GL (GE Healthcare catalog no. 29-0915-96), in a buffer composed of 20 mM HEPES pH 7.5 and 200 mM NaCl.

Protein crystallization and data collection. After size-exclusion purification of activated dasTMPRSS2, samples were pooled and concentrated to 2 mg ml^{-1} . Protein was treated with 3:1 nafamostat:dasTMPRSS2 for 10 min at room temperature and exchanged into Assay Buffer supplemented with 3:1 nafamostat using four spin cycles in 30 kDa Amicon MWCO filters (14,000 r.p.m., 15 min, 4 °C) to remove low molecular weight autolytic fragments from the 42 kDa enzyme. Acylated enzyme was then concentrated to 8 mg ml^{-1} and centrifuged

(14,000 r.p.m., 10 min, 4 °C) before automated screening at 18 °C in 96-well Intelliplates (Art Robin) using the Phoenix protein crystallization dispenser (Art Robbins). Protein was dispensed as $0.3 \mu\text{l}$ of sitting drops and mixed 1:1 with precipitant. The RedWing and SGC precipitant screens were tested and amorphous, nondiffracting crystals were consistently produced when grown over 30% Jeffamine ED-2001 (Hampton Research) with 100 mM HEPES pH 7.0. To acquire a diffraction quality crystal, acylated dasTMPRSS2 was treated with 50 U of PNGase F (NEB, 37 °C for 45 min) to trim N-glycan branches, then centrifuged (14,000 r.p.m., 4 °C, 10 min) before setting $2 \mu\text{l}$ of hanging drops with 1:1 protein:precipitant and grown for 10 days. Crystals were then cryo-protected using reservoir solution supplemented with roughly 5% (v/v) ethylene glycol and cryo-cooled in liquid nitrogen. X-ray diffraction data were collected on the beamline 24-ID-E at the Advanced Photon Source.

Solving the dasTMPRSS2:nafamostat crystal structure. X-ray diffraction data were processed with XDS⁵². Initial phases were obtained by molecular replacement in Phaser MR⁵³, using (PDB 1Z8G) as a starting model. Model building was performed in COOT⁵⁴ and refined with Buster⁵⁵. Structure validation was performed in Molprobit⁵⁶. Data collection and refinement statistics are summarized in Supplementary Table 2, and a phenylguanidino ligand electron density omit map as well as the detailed TMPRSS2 residue interactions are available in Supplementary Fig. 1a,b.

Gel electrophoresis and western blotting. SDS–PAGE was carried out with $15 \mu\text{l}$ of Mini-Protean (BioRad) or $60 \mu\text{l}$ of Novex Wedgewell (Invitrogen) 4–20% Tris–Glycine gels for 30 min under constant voltage at 200 V. Protein samples were mixed with 4× Laemmli buffer (BioRad) and subjected to differential reducing ($\pm 5 \text{ mM}$ β-mercaptoethanol, Gibco), then boiling at 95 °C for 5 min to probe the covalent nature of protein complexes and subunits. The Precision Plus Protein marker (BioRad) was used as a standard.

For SARS-CoV-2 RBD western blotting, SDS–PAGE was carried out as described, followed by wet transfer in Transfer Buffer (25 mM Tris pH 8.3, 192 mM glycine, 20% MeOH (v/v)) to a polyvinylidene difluoride membrane (80 V, 53 min, 4 °C). Membranes were incubated in Blocking Buffer (5% skim milk in tris-buffered saline with Tween (TBST)) for 1 h at room temperature, washed 5× with TBST, then probed overnight with 1/3,000 mouse anti-RBD primary mAb (Abcam ab277628) solution at 4 °C. Membranes were then washed 5× with TBST and probed with 1/5,000 FITC-labeled goat anti-mouse IgG secondary pAb (Abcam ab6785) and imaged for fluorescence on the Typhoon FLA7000 biomolecular imager (GE healthcare).

Enzyme peptidase and inhibition assays. Peptidase assays with fluorogenic Boc-Gln-Ala-Arg-AMC substrate (Bachem catalog no. 4017019.0025) were performed in 96-well plates (Greiner Fluotrak) at $200 \mu\text{l}$ of reaction volumes in a FlexStation microplate reader (Molecular Devices) using the SoftMax Pro software (v.5.4.6) at 24 °C. Fluorescence was monitored with the fastest kinetic read settings across 5 min at 341:441 nm excitation:emission and converted to a product AMC concentration using standard curves at each substrate concentration to correct for the inner-filter effect. All inhibition assays contained 2% (v/v) DMSO and initial reaction velocities were tabulated over the linear portion of the first 60 s of progress curves.

To determine Michaelis–Menten kinetic parameters, $50 \mu\text{l}$ of 4× enzyme stock (12.8 nM) was added through automated addition 1:3 to microplates containing $150 \mu\text{l}$ of substrate (0.5–1000 μM) in triplicate and initial reaction velocities were plotted against substrate concentration and curve fit using GraphPad Prism (v.9.1.1).

IC₅₀ potencies of nafamostat mesylate (MedChemExpress catalog no. HY-B0190A), camostat mesylate (MedChemExpress catalog no. HY-13512), benzamidine HCl (Sigma catalog no. 434760-25G), bromhexine HCl (SelleckChem catalog no. S2060) and SFTI-1 (Methods) were initially determined by preincubating dasTMPRSS2 with inhibitor at concentrations ranging from 0.1 nM–100 μM for 5 min, before enzyme–inhibitor mixes were added to substrate through automated addition. Then, seven inhibitor concentrations spanning three orders of magnitude across the IC₅₀ value were used and inhibitor reaction velocities were normalized to uninhibited enzyme and plotted as one-site dose response curves in GraphPad. The apparent K_i (K_i^{*}) of classical competitive trypsin-like serine protease inhibitors benzamidine and SFTI-1 were determined using equation (1),

$$K_i^* \approx \frac{IC_{50}}{1 + \frac{[S]}{K_M}} \quad (1)$$

where [S] is the concentration of substrate Boc-QAR-AMC used in the assay and K_M is the Michaelis constant (200 μM).

Time-dependent IC₅₀ measurement and k_{i,max}/K_i determination. Camostat IC₅₀ curves were generated using seven concentrations of inhibitor in the range of 0.1–1,000 nM inhibitor and nafamostat between 0.01 and 100 nM with a DMSO control as described. The time dependence of inhibitor potencies was measured by using Flexstation Flex kinetic reads that automatically transferred dasTMPRSS2–inhibitor

mixes to substrate wells at the indicated preincubation timepoints (Extended Data Fig. 2b). For the 10 s timepoint, a kinetic read was performed after manual addition of enzyme, followed by substrate, using a multichannel pipette to capture the fast acylation (Fig. 5e). Kinetic parameters K_i^* and k_{inact} were determined with the simplified equations (2) and (3), respectively, assuming a one-step kinetic inhibition mechanism



$$K_i^* \approx \frac{t_{50}^{(2)} - t_{50}^{(1)}}{\frac{I_{50}^{(2)}}{I_{50}^{(1)}} - \frac{t_{50}^{(1)}}{t_{50}^{(2)}}} \quad (3)$$

$$k_{\text{inact}} \approx \frac{1}{t_{50}^{(2)}} \exp \left[\ln \left(\frac{K_i^*}{I_{50}^{(2)}} - 1 \right) + b \right], \quad (4)$$

where the conversion factor $b=0.558$ is applied for concentrations in μM and time in s.

Active site quantification of dasTMPRSS2. The acylation of dasTMPRSS2 by nafamostat and concomitant production of the fluorogenic 6-amidino-2-naphthol leaving group was measured by incubating serial dasTMPRSS2 enzyme dilutions from 6.4 to 0.8 nM with excess (10 μM) nafamostat, similar to previous efforts with matriptase³⁶ (Extended Data Fig. 2c,d). Microplate reading at 320 and 490 nm excitation and emission, respectively, were used to calculate the number of dasTMPRSS2 active site residues and calibrate peptidase activity and inhibition assays at 3.4 nM enzyme.

Nafamostat inhibition half-life. The half-life of the phenylguanidino acyl-enzyme complex after nafamostat treatment was measured for dasTMPRSS2 using methods established for camostat with enteropeptidase³⁸. Briefly, dasTMPRSS2 (3.4 μM) was mixed with slight excess nafamostat (5 μM or DMSO control) and incubated at room temperature for 20 min. After incubation, unbound nafamostat was removed by passage and three washes in a 3 kDa MWCO Amicon filter centrifuged at maximum speed. Acylated or untreated dasTMPRSS2 samples were then transferred in quadruplet to a microplate containing either 125 or 250 μM Boc-QAR-AMC substrate (final concentration of 3.4 nM enzyme). Fluorescent reads were carried out immediately, analogous to IC_{50} assays, but across a period of 8 h to monitor for rescued dasTMPRSS2 activity after deacylation. The acylated traces were fit to a one-phase association exponential in GraphPad to derive the half-life for activity recovery, normalized to the uninhibited initial reaction velocity.

SFTI-1 synthesis and purification. Reagents and solvents were purchased from commercial sources and used without further purification, unless otherwise stated. High-performance liquid chromatography (HPLC) was performed on (1) an Agilent 1260 infinity system equipped with a model no. 1200 quaternary pump and a model no. 1200 ultraviolet absorbance detector or (2) an Agilent 1260 Infinity II Preparative System equipped with a model 1260 Infinity II preparative binary pump, a model 1260 Infinity variable wavelength detector (set at 220 nm) and a 1290 Infinity II preparative open-bed fraction collector. The HPLC column used for analysis was a Phenomenex Luna C18 semipreparative column (5 μm , 45,250 \times 10 mm). The HPLC column used for synthesis was a preparative column (Gemini, NX-C18, 46 μm , 110 \AA , 50 \times 30 mm) purchased from Phenomenex. Mass analyses were performed using a Waters 2695 Separation module and Waters-Micromass ZQ mass spectrometer system.

Peptides were synthesized on a Liberty Blue automated microwave peptide synthesis (CEM Corporations). Fmoc-Gly-OH was loaded on a 2-chlorotriethyl resin (AdvancedChemTech) in CH_2Cl_2 and 2,4,6-trimethylpyridine at room temperature. The resin was capped with a solution of 10/5/85 MeOH/*N,N*-diisopropylethylamine (DIEA)/ CH_2Cl_2 for 10 min. Fmoc groups were removed using 20% piperidine in dimethylformamide (DMF) at 50 $^\circ\text{C}$ for 10 min. Using HATU/DIEA chemistry (5 eq. each in DMF) at 50 $^\circ\text{C}$ for 10 min per cycle, Fmoc-Asp(OBno)-OH, Fmoc-Pro-OH, Fmoc-Phe-OH, Fmoc-Cys(ACM)-OH, Fmoc-Ile-OH, Fmoc-Pro-OH, Fmoc-Pro-OH, Fmoc-Ile-OH, Fmoc-Ser(tBu)-OH, Fmoc-Lys(Boc)-OH, Fmoc-Thr(tBu)-OH, Fmoc-Cys(ACM)-OH and Fmoc-Arg(Pbf)-OH were loaded sequentially. Fmoc-Arg(Pbf)-OH and Fmoc-amino acids coupled after Pro were coupled using two cycles. The peptide was cleaved from the resin using 20% HFIP in CH_2Cl_2 with no subsequent purification. The peptide was cyclized in solution using 3 eq. of PyAOP reagent and 6 eq. of DIEA in DMF using microwave heating at a concentration of 1 mg ml^{-1} at 90 $^\circ\text{C}$ for 15 min. The reaction was concentrated in vacuo. Next, the cyclized peptide was treated with 95/2.5/2.5 TFA/TIS/ H_2O and stirred for 3 h. The crude peptide mixture was concentrated and added to a solution of cold diethyl ether. The suspension was centrifuged at 2,500 r.p.m. for 7 min, the supernatant diethyl ether was discarded, and the solids were diluted into water, frozen and lyophilized to yield a white powder. The reaction was purified using preparative HPLC, eluted

with 16–36% acetonitrile in water with 0.1% TFA over 20 min at a flow rate of 30 ml min^{-1} . The retention time was 11.23 min and the collected fractions were frozen and lyophilized. Next, the collected white powder was dissolved in 300 μl of 50% AcOH in H_2O and 30 μl of 1 M HCl. Next, 10 μl of 0.1 M I_2 in AcOH was added to the reaction mixture and stirred for 5 h at room temperature, while being monitored using semipreparative HPLC. After confirmation of starting material consumption, the reaction was added to a 5 ml solution of 0.1 M NaOAc and purified using preparative HPLC, eluted with 21–41% acetonitrile in water with 0.1% TFA over 20 min at a flow rate of 30 ml min^{-1} . The retention time was 10.13 min and the collected fractions were frozen and lyophilized to yield a white powder. Electrospray ionization–mass spectrometry calculated $[\text{M}+2\text{H}]^{2+}$ for $\text{C}_{67}\text{H}_{106}\text{N}_{18}\text{O}_{18}\text{S}_2$ 757.37 and found 757.94.

DSF. Apparent melting temperature ($T_{M,a}$) shifts were measured for various dasTMPRSS2-inhibitor incubations using SYPRO Orange dye (Life Technologies, catalog no. S-6650) and monitoring fluorescence at 470 and 510 nm excitation and emission, respectively, using the Light Cycler 480 II (Roche Applied Science). Samples were prepared in quadruplet in 384-well plates (Axygen; catalog nos. PCR-384-C; UC500) at a final volume of 20 μl containing 0.05 mg ml^{-1} dasTMPRSS2, 1 μM compound or vehicle control and 5 \times SYPRO Orange. Thermal melt curves were generated between 25 and 95 $^\circ\text{C}$ at a gradient of 1 $^\circ\text{C min}^{-1}$ and plots prepared with the online DSFworld application³⁷ for $T_{M,a}$ determination.

SARS-CoV-2 S protein activation and inhibition. Recombinant SARS-CoV-2 S protein constructs HexaFurin and HexaPro were concentrated to 0.5 mg ml^{-1} in Assay Buffer and incubated with the indicated concentrations of furin protease (NEB) or dasTMPRSS2. Digestions took place over 16 h for furin and from 5–120 min for dasTMPRSS2. Furin digestions were terminated by the addition of 4 mM EDTA whereas dasTMPRSS2 digestions were terminated with 5 μM nafamostat (as TMPRSS2 activity is insensitive to EDTA, Fig. 4c), then SDS-PAGE samples were immediately prepared with the addition of 4 \times SDS-PAGE loading buffer and boiled for 5 min at 95 $^\circ\text{C}$. Next, 4 μg of S protein were loaded per well under each condition and gels were visualized by Coomassie blue staining. For anti-RBD western blotting, 2 μg of S protein were loaded per well.

For cleavage inhibition assays (Extended Data Fig. 4b,d), dasTMPRSS2 diluted to 320 nM in Assay Buffer was preincubated 15 min with inhibitor (final 1% DMSO (v/v)) or DMSO control, then assays were started by transfer of enzyme-inhibitor mixes to S protein. S protein–dasTMPRSS2 mixtures were incubated at room temperature for 2 h with nafamostat and 30 min for bromhexine.

Multiple sequence alignments. Multiple sequence alignments were prepared to compare the human TTSP family members as well as TMPRSS2 mammalian orthologs. Human TTSP FASTA sequences (isoform 1) were accessed from UniProt and mammalian TMPRSS2 orthologs identified with UniProt BLAST. Sequences were aligned with Clustal Omega³⁸ and annotated with ESPrnt v.3.0 (ref. 39).

Protein visualization and property calculation. The structure of dasTMPRSS2 was inspected and compared to other TTSPs using PyMol v.2.4.1 (Schrodinger) and the Molecular Operating Environment (MOE) (Chemical Computing Group) 2019 software suite. The exposed hydrophobic patches of TMPRSS2 were calculated using the MOE Protein Patch Analyzer tool⁴⁰.

Data exclusions and statistics. The two-sided Grubbs' test was used to determine and exclude single outliers present in sample data performed in triplicate or quadruplet. Single datapoint outliers were identified in Extended Data Fig. 2b (triplicate samples, 5 and 10 nM nafamostat), Extended Data Fig. 2c (triplicate samples; 0 nM 6-amidino-2-naphthol) Extended Data Fig. 2d (triplicate samples; twofold enzyme concentration) and Extended Data Fig. 2e (quadruplet samples; nafamostat 125 μM substrate + nafamostat; 250 μM substrate + DMSO). Within the supplied source data documents, these exclusions are denoted with blue text and asterisks.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The coordinates and structure of the phenylguanidino TMPRSS2 acyl-enzyme complex have been deposited in the PDB with accession number 7MEQ on April 7, 2021, and released on April 21, 2021. Our engineered dasTMPRSS2 protein expression construct is available on Addgene (Plasmid no. 176412). Hepsin, TMPRSS13 and TMPRSS15 protein structures were accessed using PDB ID 1Z8G, 6KD5 and 6ZOV, respectively. All other relevant manuscript data are available in the source data provided with this paper.

References

- Hsieh, C. et al. Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* **369**, 1501–1505 (2020).
- Kabsch, W. XDS. *Acta Crystallogr.* **66**, 125–132 (2010).

53. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
54. Emsley, P. & Lohkamp, B. Features and development of Coot. *Acta Crystallogr.* **D66**, 486–501 (2010).
55. Bricogne G. et al. BUSTER version 2.10.3. Cambridge, United Kingdom: Global Phasing Ltd. (2017).
56. Williams, C. J. et al. MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2017).
57. Wu, T. et al. Three essential resources to improve differential scanning fluorimetry (DSF) experiments. Preprint at *bioRxiv*, <https://doi.org/10.1101/2020.03.22.002543> (2020).
58. Madeira, F. et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641 (2019).
59. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–W324 (2014).
60. Jetha, A. et al. Homology modeling and structure-based design improve hydrophobic interaction chromatography behavior of integrin binding antibodies. *MAbs* **10**, 890–900 (2018).

Acknowledgements

We thank J. McLellan for generously providing the SARS-CoV-2 S protein construct plasmids, I. Chau for assistance with DSF and S.-T. Tseng for preparation of Fig. 1 graphic art. This work was supported by the BC Leadership Chair in Functional Cancer Imaging to F.B., the Canadian Institute of Health Research grant (no. FDN154328) to C.H.A., the Canada Excellence Research Chair to S.S. and a Mitacs Accelerate Internship to B.F. This work is based on research conducted at the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from the National Institutes of Health (grant no. P30 GM124165). The Eiger 16M detector on 24-ID-E beamline is funded by a NIH-ORIP HEI grant (no. S10OD021527). This research used resources of the Advanced Photon Source, a US Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under contract no. DE-AC02-06CH11357. The Structural Genomics Consortium is a registered charity (no. 1097737) that receives funds from Bayer AG, Boehringer Ingelheim, Bristol Myers Squibb, Genentech, Genome Canada through Ontario Genomics Institute (grant no. OGI-196),

the EU and EFPIA through the Innovative Medicines Initiative 2 Joint Undertaking (EUbOPEN grant no. 875510), Janssen, Merck KGaA (also known as EMD in Canada and the United States), Pfizer and Takeda.

Author contributions

F.B., C.H.A., L.H. and S.S. provided project supervision. B.J.F., S.B., A.S., C.H.A. and F.B. conceived the project. B.J.F., S.B., A.S., C.H.A. and L.H. designed the experiments. Y.L. cloned TMPRSS2 protein constructs for expression. Y.L. and D.M. cloned SARS-CoV-2 S protein constructs for expression. A.S. and A.H. produced TMPRSS2 protein in insect cells. A.S., A.H., D.M. and K.L. produced SARS-CoV-2 S protein in human embryonic kidney cells. B.J.F., S.B., A.S., A.H., D.M. and K.L. purified recombinant proteins. B.J.F., S.B. and L.H. crystallized TMPRSS2. L.H., S.B. and B.J.F. collected diffraction data. L.H. solved the crystal structure. B.J.F., S.B. and L.H. performed bioinformatic and structural analyses. B.J.F., D.K., R.P.W. and R.T. performed fluorogenic peptidase activity and inhibitor potency assays and B.J.F. analyzed kinetics. D.K. synthesized, purified and characterized SFTI-1 peptide. B.J.F., D.K. and S.B. managed inhibitor compound libraries. B.J.F., D.M. and R.T. performed gel-based S protein digestion assays and S.B. and B.J.F. performed DSF assays. B.J.F., S.B., L.H., R.T. and D.M. prepared the figures. B.J.F., S.B., L.H., C.H.A., F.B., A.S., A.H. and Y.L. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

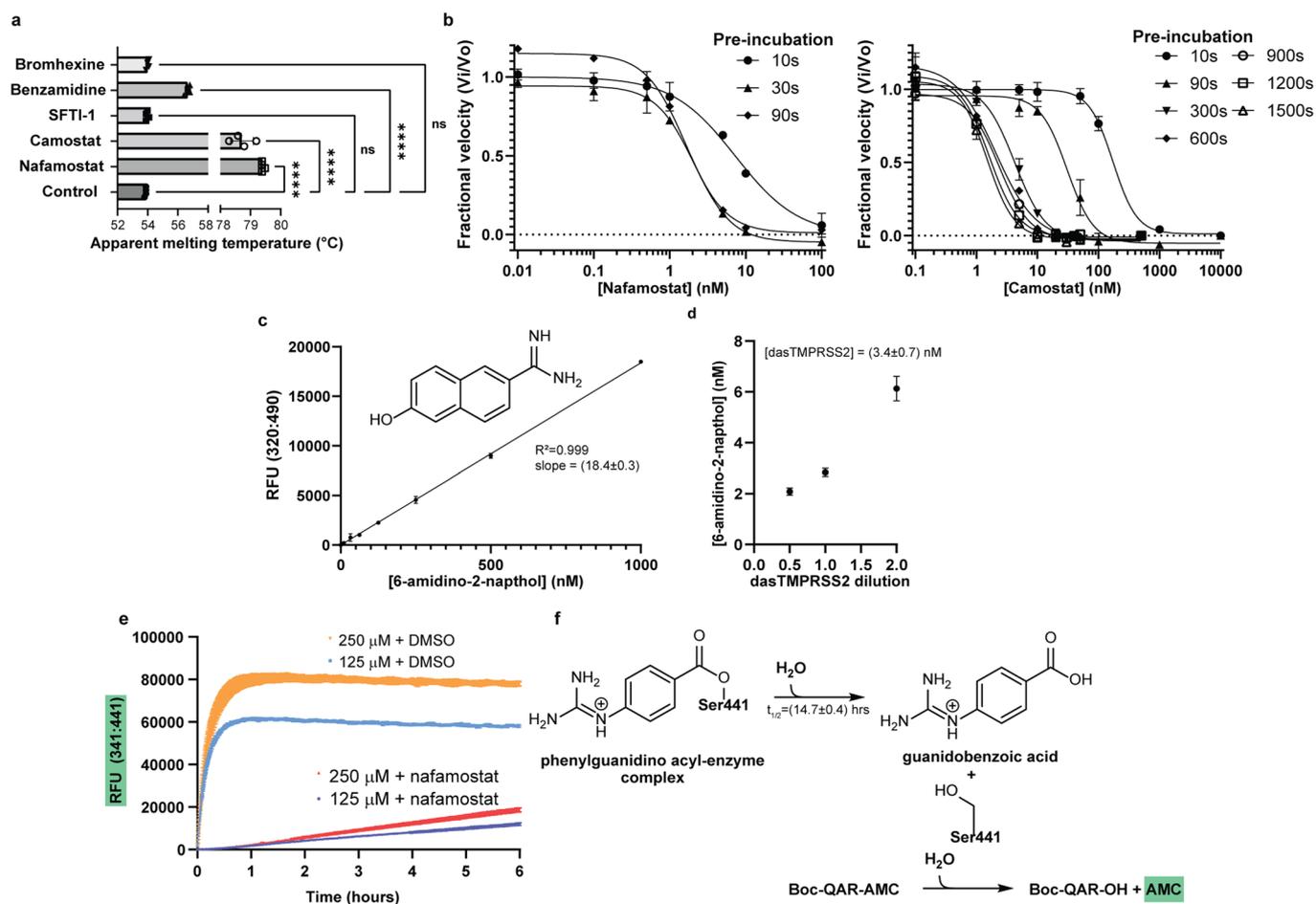
Extended data are available for this paper at <https://doi.org/10.1038/s41589-022-01059-7>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41589-022-01059-7>.

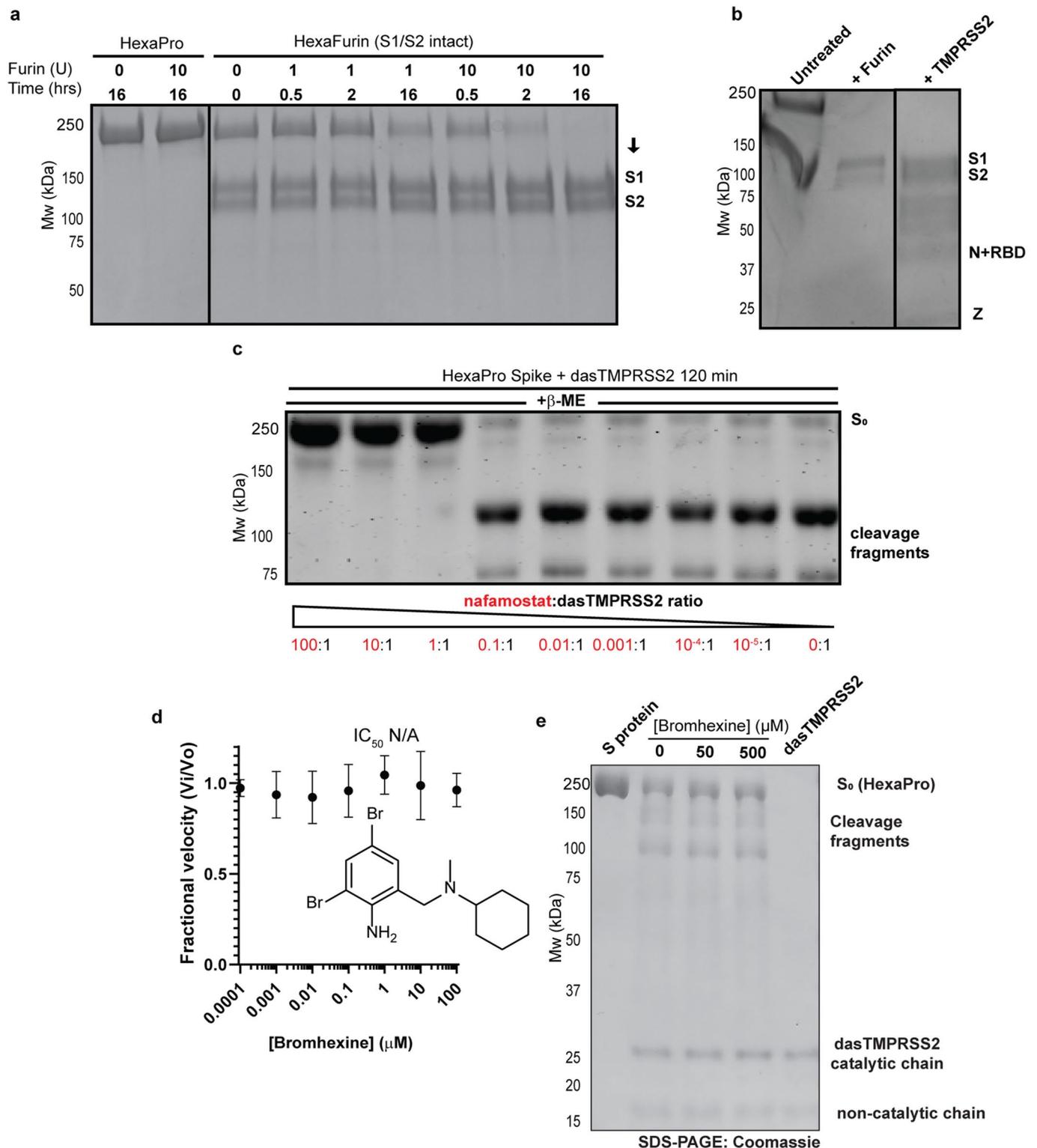
Correspondence and requests for materials should be addressed to Levon Halabelian, Cheryl H. Arrowsmith or François Bénard.

Peer review information *Nature Chemical Biology* thanks Matthew Hall, Andrey Kovalevsky and Nobuo Maita for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.



Extended Data Fig. 2 | Rapid nafamostat acylation of the dasTMPRSS2 catalytic serine is thermally stabilizing and slowly deacylates. **a** Apparent dasTMPRSS2 melting temperatures (as determined by differential scanning fluorimetry; Methods) are increased for benzamidine, camostat, and nafamostat at 1 μ M concentration but are not increased by SFTI-1 or bromhexine relative to control incubation with DMSO. Datapoints are shown as means \pm SD. Samples were run in technical quadruplet and statistically significant differences ($p < 0.0001$; denoted as ****) from control were determined by one-way ANOVA for $n=4$ biologically independent samples (or $n=3$ samples for SFTI-1 and benzamidine) examined over 1 independent experiment. **b** IC₅₀ plots of camostat (left) and nafamostat (right) across increasing lengths of inhibitor pre-incubation time with dasTMPRSS2 enzyme, enabling calculation of k_{inact} and K_i via Equations 2 and 3 (Methods). Datapoints are shown as means \pm SD. Samples were run in technical duplicate ($n=2$) and curve fit for absolute IC₅₀ in GraphPad. Consistent plots were obtained for $n=3$ independent biological experiments. **c** Fluorescence standard curve for the quantification of the nafamostat leaving group, 6-amidino-2-naphthol. Samples were run in technical triplicate ($n=3$) and fit to a linear curve in GraphPad producing the indicated slope and R-squared values. **d** dasTMPRSS2 enzyme quantification after 5 min incubation with 1 μ M nafamostat and monitoring 6-amidino-2-naphthol production as endpoint fluorescence (320:490). The indicated dilutions estimate a 3.2 nM dasTMPRSS2 concentration with samples run in technical duplicate ($n=2$). **e** Residual dasTMPRSS2 peptidase activity monitored by turnover of Boc-QAR-AMC substrate following stoichiometric enzyme acylation by nafamostat. Enzyme co-incubations with nafamostat or DMSO were transferred to wells containing 250 μ M or 125 μ M substrate (final enzyme concentration of 3.2 nM) and kinetic traces of nafamostat-treated samples were curve fit in GraphPad to calculate the half-life of activity recovery (Methods). **f** Chemical mechanism of hydrolysis of the phenylguanidino acyl-enzyme complex and recovery of dasTMPRSS2 peptidase activity observed kinetically in (e).



Extended Data Fig. 4 | Recombinant S protein is differentially cleaved by dasTMPRSS2 and furin and is a useful substrate for dasTMPRSS2 inhibitor characterization. **a** Recombinant furin (NEB enzyme units indicated) cannot cut SARS-CoV-2 S protein with a KO at the S1/S2 cleavage site (HexaPro), but can cut S1/S2 intact S protein at S1/S2 over a period of 16 hrs at room temperature. 4 μg protein were loaded per well and visualized with Coomassie blue and gel results were consistent across $n=3$ independent biological experiments. **b** HexaPro S protein digestion across 120 min with 300 nM dasTMPRSS2 following 15 min pre-incubation with 100:1 to 0.00001:1 nafamostat:dasTMPRSS2. 4 μg protein were loaded per well and visualized with Coomassie blue and gel results were consistent across $n=3$ independent biological experiments. **c** Attempted IC₅₀ plot of bromhexine (with its chemical structure shown) of residual dasTMPRSS2 peptidase activity in the presence of 10 μM Boc-QAR-AMC substrate showing no detectable inhibition. Data are shown as mean ± SD for $n=3$ technical replicates. **d** Pre-incubation of bromhexine with 300 nM TMPRSS2 for 15 min results in no inhibition of proteolytic activity towards SARS-CoV-2 S protein construct HexaPro digested over 30 minutes. 4 μg protein were loaded per well and visualized with Coomassie blue and gel results were consistent across $n=3$ independent biological experiments.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection X-ray diffraction data were collected on the beamline 24-ID-E at the Advanced Photon Source (APS). Enzyme kinetics were monitored using a FlexStation 3 Multi-Mode Microplate Reader (Molecular Devices) using the SoftMax Pro (Version 5.4.6) software.

Data analysis X-ray diffraction data were processed with XDS and merged with Aimless (version 0.7.4). Phaser MR (CCP4Interface 7.1.008) was used for molecular replacement. Models were built with winCOOT (version 0.8.9.2 EL), and refined with Buster (Version: 2.10.3). The SoftMax Pro (Version 5.4.6) was used for enzyme progress curve collection, and kinetic curve fitting of enzyme rate data was performed using GraphPad Prism (Version 9.1.1). The online DSFWorld tool was used to analyze fluorescent data from differential scanning fluorimetry assays for apparent melting temperature determination. Tmprss2 mammalian ortholog sequences were identified using UniProt BLAST and aligned and annotated with Clustal Omega and Espript 3.0, respectively. Structural analyses were performed using PyMOL v2.4.1 and the Molecular Operating Environment 2019 Software suite.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The coordinates and structure of the phenylguanidino TMPRSS2 acyl-enzyme complex have been deposited in the PDB with accession number 7MEQ on April 7, 2021, and released on April 21, 2021. Our engineered dasTMPRSS2 protein expression construct is available on Addgene (Plasmid #176412). Hepsin, TMPRSS13, and TMPRSS15 protein structures were accessed using PDB ID 1Z8G, 6KD5, and 6ZOV, respectively. All other relevant manuscript data is available in Source Data.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The number of technical replicates were sufficient to provide a degree of uncertainty of 20% or less and therefore judged to allow unambiguous conclusions to be drawn. Exact samples sizes and replicates are described in Methods and figure legends.
Data exclusions	Single datapoint outliers have been specifically called out in the Methods section of the manuscript, Data exclusions and statistics, and also highlighted in Source Data Excel sheets. Statistical tests for single datapoint outlier exclusions are outlined in the manuscript.
Replication	All kinetic measurements were reproducible within the uncertainties reported. In most cases the experiments were successfully repeated on separate dates with n=3 independent biological experiments, with specific replicate details available in the manuscript.
Randomization	Kinetic measurements were performed as a series of independent experiments. Randomization has no meaning to the present work
Blinding	Kinetic measurements were performed as a series of independent experiments, and, as repetitions of identical procedures, blinding has no meaning or relevance to the work.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used	Mouse anti-SARS-CoV-2 Spike Glycoprotein RBD primary antibody (Abcam ab277628; 1/3000); Goat anti-mouse IgG H&L chain-FITC (Abcam ab6785; 1/5000)
Validation	ab277628 was validated to detect HEK-293 cells transfected with pCI-Neo-Mod containing the SARS-CoV2-bd cDNA with additional info at (https://www.abcam.com/sars-cov-2-spike-glycoprotein-rbd-antibody-5g8-ab277628.html). ab6785 was validated to detect several primary mouse monoclonal antibodies with specific info at (https://www.abcam.com/goat-mouse-igg-hl-fitc-ab6785.html)

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)

Expi293F™ Cells, Catalog number: A14527 Life Technologies Corporation ; Gibco™ Spodoptera frugiperda (Sf9) cells, Catalog number: 12659017 Thermo Fisher

Authentication

All cell lines were tested for authentication by STR profiling

Mycoplasma contamination

All cell lines were mycoplasma negative, as determined by MycoAlert® Mycoplasma Detection Kit (Lonza; Cat # LT07-318)

Commonly misidentified lines
(See [ICLAC](#) register)

No commonly misidentified cell lines were used according to ICLAC database (<https://iclac.org/databases/cross-contaminations/>)