Structural basis of TMPRSS2 zymogen activation and recognition by the HKU1 seasonal coronavirus

Ignacio Fernandez, Nell Saunders, Stéphane Duquerroy, William H Bolland, Atousa Arbabian, Eduard Baquero, Catherine Blanc, Pierre Lafaye, Ahmed Haouz, Julian Buchrieser, et al.

Structural basis of TMPRSS2 zymogen activation and recognition by the HKU1 seasonal coronavirus

Ignacio Fernández[1,8], Nell Saunders[2,8], Stéphane Duquerroy[1,3], William H. Bolland[2], Atousa Arbabian[1], Eduard Baquero[4], Catherine Blanc[5], Pierre Lafaye[6], Ahmed Haouz[7], Julian Buchrieser[2], Olivier Schwartz[2,*], Félix A. Rey[1,9,*]

[1] Institut Pasteur, Université de Paris Cité, CNRS UMR 3569, Structural Virology Unit, 75015, Paris, France.

[2] Institut Pasteur, Université de Paris Cité, CNRS UMR 3569, Virus & Immunity Unit, 75015, Paris, France and Vaccine Research Institute, Créteil, France

[3] Université Paris-Saclay, Faculté des Sciences, Orsay, France.

[4] Institut Pasteur, Université de Paris Cité, INSERM U1222, Nanoimaging core, 75015, Paris, France.

[5] Institut Pasteur, Université de Paris Cité, Pasteur-TheraVectys Joint Lab, Paris, France

[6] Institut Pasteur, Université Paris Cité, CNRS UMR 3528, Antibody Engineering Facility-C2RT, 75015, Paris, France.

[7] Institut Pasteur, Université Paris Cité, CNRS UMR 3528, Crystalogenesis Facility-C2RT, 75015, Paris, France.

[8]Equal contribution
[9]Lead contact

*Correspondence: olivier.schwartz@pasteur.fr, felix.rey@pasteur.fr

ABSTRACT

The human seasonal coronavirus HKU1-CoV, which causes common colds worldwide, relies on the sequential binding to surface glycans and transmembrane serine protease 2 (TMPRSS2) for entry into target cells. TMPRSS2 is synthesized as a zymogen that undergoes autolytic activation to process its substrates. Several respiratory viruses, in particular coronaviruses, use TMPRSS2 for proteolytic priming of their surface spike protein to drive membrane fusion upon receptor binding. We describe the crystal structure of the HKU1-CoV receptor binding domain in complex with TMPRSS2, showing that it recognizes residues lining the catalytic groove. Combined mutagenesis of interface residues and comparison across species highlight positions 417 and 469 as determinants of HKU1-CoV host tropism. The structure of a receptor-blocking nanobody in complex with zymogen or activated TMPRSS2 further provides the structural basis of TMPRSS2 activating conformational change, which alters loops recognized by HKU1-CoV and dramatically increases binding affinity.

INTRODUCTION

The cell surface transmembrane serine protease 2 (TMPRSS2) proteolytically primes the spike protein of multiple coronaviruses for entry into cells of the respiratory tract and was recently identified as the entry receptor for the human HKU1 coronavirus (HKU1-CoV)[1]. TMPRSS2 belongs to the "type 2 transmembrane serine proteases" (TTSP) involved in proteolytic remodeling of the extracellular matrix. Dysregulation of TMPRSS2 and other TTSPs has been observed in malignancies and is associated to tumor proliferation and invasiveness [2-5]. Moreover, TMPRSS2 is an androgen regulated protease that activates prostate cancer metastatic cascades [6].

HKU1-CoV is a seasonal beta-coronavirus that was initially isolated in 2005 at the University of Hong Kong[7], hence its name. It causes common colds worldwide and can develop complications in young children, the elderly and immunocompromised individuals[8]. Entry of HKU1-CoV into cells relies on its trimeric spike surface glycoprotein, which catalyzes the fusion of the viral envelope with the target cell membrane. The spike is synthetized as a precursor cleaved into two subunits, S1 and S2. S1 harbors an N-terminal domain (NTD), a receptor binding domain (RBD), and subdomains SD1 and SD2. The RBD adopts conformations known as 'up' and 'down', which determine different forms on the spike (closed, 1-RBD-up, 2-RBD-up, 3-RBD-up or "open"). The S2 subunit is the membrane fusion effector and requires further cleavage at a second site - the S2' site, to be functional [9]. As an important entry factor, one of the roles of TMPRSS2 is to cleave HKU1-CoV S2 at the S2' site [1], priming it for driving membrane fusion.

The HKU1-CoV spike NTD binds α2,8-linked 9-O-acetylated disialosides (9-O-Ac-Sia(α2,8)Sia) [10-12] causing a conformational change that exposes the RBD and opens the spike [13]. The transition into open states has been associated to the activation of the protein to trigger membrane fusion[14,15]. The surface sialo glycan is thus a primary receptor of HKU1-CoV. The activity of a hemagglutinin esterase (HE) anchored in the viral envelope[10,16] allows dissemination of the virus to relevant tissues by acting as a receptor-destroying enzyme[17]. We have shown that TMPRSS2 recognizes the HKU1-CoV RBD in a so far uncharacterized mode[1]. Even though sialic acid is important to trigger efficient HKU1-CoV entry into TMPRSS2 expressing cells[1], no experimental evidence has demonstrated the interplay between the two receptors. We also developed nanobodies against TMPRSS2, some of which bind without affecting recognition of the RBD (i.e., nanobody A01), while others block the interaction, reducing infection of HKU1-CoV susceptible cells, and inhibit TMPRSS2 proteolytic activity (i.e., nanobody A07)[1].

TMPRSS2 is a type II single-pass transmembrane protein with an N-terminal cytosolic tail and an ectodomain containing a low-density lipoprotein receptor type-A (LDLR-A) domain followed by a class A scavenger receptor cysteine-rich (SRCR) domain and a C-terminal trypsin-like serine peptidase (SP) domain[18]. It is synthesized as a proenzyme (zymogen) that undergoes auto cleavage to reach its mature, active conformation[19,20]. The active site contains the catalytic triad H296-D345-S441, which cleaves after arginine 255 of the zymogen for autolytic activation, at the peptide sequence ($RQSR_{255}\downarrow IVGG$), where the arrow denotes the scissile bond. The substrate sequence, designated P4-P3-P2-P1↓P1'-P2'-P3'-P4', is recognized through specific TMPRSS2 residues that form the corresponding S4-S3-S2-S1-S1'-S2'-S3'-S4'

sites along the catalytic groove[21]. TMPRSS2 substrates include proteins of the tumor microenvironment[6], such as the hepatocyte growth factor (cleaved sequence $KQLR_{494}\downarrow VVNG$), tissue plasminogen activator ($PQFR_{310}\downarrow IKGG$) or human glandular kallikrein 2 ($IQSR_{24}\downarrow IVGG$) as well as the S2' site of the coronavirus spike[22-26], at the sequence motifs $SSSR_{900}\downarrow SLLE$ (HKU1-CoV isolate N5) or $PSKR_{815}\downarrow SFIE$ (SARS-CoV-2 Wuhan). The amino acid cleavage preferences at positions P1-P4 of the substrate have been determined by positional scanning of synthetic peptide libraries[6].

Here, we determined the X-ray structure of TMPRSS2 in its zymogen form as a ternary complex with the HKU1-CoV genotype B RBD (from here on, HKU1B RBD) and the non-competing nanobody A01. TMPRSS2 binds exclusively to RBDs in the 'up' conformation, making spike opening mandatory for binding. Sialo glycan binding to the spike at the NTD is a required first step that allows fusion with TMPRSS2 expressing cells. We further identified key TMPRSS2 residues determining host tropism through extensive functional validation of the RBD-TMPRSS2 interface. We also determined a high-resolution X-ray structure of TMPRSS2 in complex with nanobody A07, showing that it inserts its long CDR3 into the protease's active site, explaining its role in blocking HKU1-CoV entry and TMPRSS2 activity. The crystals of the TMPRSS2/A07 complex contained two forms, the zymogen and activated protease, illustrating the activation mechanism of TMPRSS2 and providing key information for the development of specific, zymogen-targeting drugs. The conformational change caused by the activating cleavage of TMPRSS2 involves loops that are recognized by the HKU1-CoV RBD, and protease activation is required for high-affinity binding.

RESULTS

The HKU1B RBD partially blocks the TMPRSS2 catalytic groove. We determined the X-ray structure of the ternary $TMPRSS2^{S441A}$/RBD/A01 complex (FigureS1) to 3.55 Å resolution, displayed in Figure 1A-E, with the crystallographic statistics listed in Table 1. We used the inactive $TMPRSS2^{S441A}$ mutant to increase protein yield and avoid protein degradation during purification and crystallization. As described previously, the SP domain of TMPRSS2 has a characteristic chymotrypsin like fold[21], featuring two β-barrels with the substrate binding groove in between[27]. This groove is surrounded by eight loops: loops 1, 2 and 3 in the C-terminal β-barrel controlling specificity for the P1 position of the substrate, and loops A, B, C and E, (N-terminal β-barrel) as well as loop D (C-terminal β-barrel), which affect specificity at more distal positions[28,29]. The structure of the complex (Figure 1B) shows that the elongated HKU1B RBD has a structure resembling a pincer plier at its distal end (in the "insertion" domain, Figure 1A), with two jaws (j1 and j2) that grab the TMPRSS2 SP domain at a surface involving loops 1-3 (L1-3) as well as loop C (LC). The RBD pincer plier's base is made by a distal β-hairpin of the RBD (residues 503-519, β10-β11), with the loop connecting the two β-strands forming one of the jaws (j1, aa 509-512) while the other is formed by the segment immediately downstream the β-hairpin (j2, residues 519-533) (Figure 1B). The base of the plier contacts loops 2 and 3 of the SP domain, while the j1 jaw inserts into the catalytic groove and contacts loops 3 and C and j2

interacts with loop 1 (Figure 1B). The RBD thus interacts with both β-barrels of the TMPRSS2 SP domain, at either side of the groove between them, thereby obstructing access of the substrate to the catalytic site. We performed an *in vitro* TMPRSS2 activity assay and showed that incubation with increasing RBD concentrations inhibits the protease activity (FigureS1C).

Nanobody A01 binds at the interface between the SRCR and SP domains, and contacts loops of the SP C-terminal β-barrel at the side opposite to the catalytic groove (FigureS1B), with a large, buried surface area (BSA) of 2500 Å^2 (1300 Å^2 on the nanobody and 1200 Å^2 on TMPRSS2) (Table 2).

Highly complementary interacting surfaces. The HKU1B RBD/TMPRSS2 interface buries a surface area of about 1600 Å^2 (~800 Å^2 on each side, Table 2) and has a shape complementarity (Sc) value of 0.72, which is substantially higher than the typical 0.64–0.68 range observed for antibody–antigen complexes[30]. The polar network at the interface involves salt bridges and several hydrogen bonds (listed in TableS1 and displayed in Figure 1C, left panel). The polar interactions surround a central hydrophobic patch in the surfaces of both RBD and TMPRSS2 (Figure 1D) which gives rise to a non-polar cluster at the center of the interface (Figure 1C, right panel). We refer from hereon to RBD or TMPRSS2 residues using suffix R or T, respectively. Of note, residues W515$_R$ and R517$_R$, which, as we showed previously[1], are crucial for the interaction with TMPRSS2, are located at the interface. The residues of the catalytic triad are not directly contacted by the RBD, which instead reaches residues of the substrate binding sites (K342$_T$ from the S1 site and T341$_T$ at S2, in loop C), or that are very close to them (i.e. N433$_T$ and S463$_T$ with respect to D435$_T$ of the catalytic triad, and S436$_T$ and G464$_T$ in the S1 site; Figure 1E). Comparison with the mode of binding of the RBD of other coronaviruses to their receptor (Figure 1F) showed that the HKU1B RBD shares a similar buried surface area at the interface, although in the other coronaviruses the RBDs lack the pincer plier structure observed in HKU1.

Only an RBD in the "up" conformation can bind TMPRSS2. To understand the interactions of the HKU1B RBD with TMPRSS2 in the context of the trimeric spike, we superimposed the X-ray structure of the complex on the RBD of the HKU1-CoV spike[13,31]. This exercise showed that TMPRSS2 cannot bind the RBD in the closed conformation because of clashes with the adjacent S1 protomers in the trimer (Figure 1G). In contrast, TMPRSS2 binding is unencumbered with the RBD in the "up" conformation, which is achieved only upon binding a disialoside glycan (9-O-Ac-Sia(α2,8)-Sia) in an allosteric pocket present in the NTD)[13]. Our RBD/TMPRSS2 structure therefore predicts that mutations that abolish binding of sialo glycans would result in a spike protein unable to bind TMPRSS2. We tested this prediction by using a split-GFP assay for syncytia formation in TMPRSS2-expressing cells[1]. We introduced the W89A mutation in the NTD, which prevents sialo glycan binding[11] and generates spike proteins that display the same closed conformation as the wild-type spike in the apo form, as determined by cryo-EM[13]. In line with these observations, the HKU1B W89A spike protein did not induce syncytia, despite normal surface expression (Figure 1H and FigureS2). These data confirmed that TMPRSS2 acts as entry receptor only after the allosteric conformational change induced on the spike by sialo glycan binding to the NTD.

146

147 **Functional analysis of the RBD-TMPRSS2 interface**

148 **Mutagenesis of the HKU1-CoV spike protein**. To establish the importance of the residues at the
149 interface of the HKU1B RBD/TMPRSS2 complex, we used a battery of assays (Figure 2). We introduced
150 mutations in plasmids coding for the HKU1B spike, generating 16 mutants that were properly expressed in
151 293T cells as assessed by flow cytometry (FigureS2A). We co-transfected the spike mutants with
152 TMPRSS2 to evaluate their fusogenic properties in the syncytia formation assay. Most of the tested spike
153 variants showed a significantly reduced ability to induce cell-cell fusion, especially when the mutations
154 involved residues that form salt bridges ($K487_R$ and $D505_R$) or are part of the hydrophobic cluster formed
155 at the interface ($L510_R$, $V512_R$, $L521_R$, $Y528_R$) (Figure 2A). The effects on cell-cell fusion were further
156 confirmed by introducing selected mutations on soluble RBDs and testing the binding to $TMPRSS2^{S441A}$ by
157 biolayer interferometry (BLI) (Figure 2B). Overall, the mutants showed diminished binding to the receptor
158 measured at the stationary state, with $L510R_R$, $Y528A_R$ and $K487A_R$ having the most significant effects.

159 **Design and analysis of TMPRSS2 mutants.** We assessed the effects of TMPRSS2 mutations
160 within the identified HKU1B interface on syncytia formation (Figure 2C), binding of soluble HKU1B spike
161 (Figure 2D) and HKU1B pseudovirus entry (Figure 2E). We generated 19 mutants that had similar or slightly
162 higher cell surface expression levels than the wild-type protein, as determined by flow cytometry
163 (FigureS2B). To assess their proteolytic activity, we used another seasonal coronavirus (229E) as a reporter
164 system. When co-expressed with its receptor, aminopeptidase N (APN)[32], the 229E spike is unable to
165 trigger cell-cell fusion in 293T (FigureS2B). Addition of wild-type TMPRSS2, but not catalytically inactive
166 $TMPRSS2^{S441A}$, triggers cell-cell fusion in 293T cells co-transfected with APN and 229E spike (FigureS2B).
167 All the TMPRSS2 mutants studied induced fusion in cells co-expressing APN and 229E spike, indicating
168 that they cleave the 229E spike and are catalytically active (FigureS2B). Co-transfecting the wild type
169 HKU1B spike with TMPRSS2 variants showed that essentially all mutations at the receptor binding interface
170 had a significantly reduced capacity to induce HKU1B-mediated cell-cell fusion (Figure 2C). As expected,
171 the control mutant $R316A_T$, located in loop E and not involved in receptor binding, induced normal cell-cell
172 fusion.

173 We also analyzed by flow cytometry the binding of soluble HKU1B spikes to 293T cells expressing the
174 TMPRSS2 variants. Most of the mutants impaired in cell-cell fusion showed decreased binding (Figure 2D).
175 The most affected mutants were those that abrogated salt-bridges ($D417A_T$ and $R470A_T$) or hydrogen
176 bonds ($Y469A_T$) or disrupted the hydrophobic cluster at the interface ($L410R_T$, $Y414A_T$, $L430R_T$, $W461A_T$).
177 The control $R316A_T$ showed normal binding levels. We also used TMPRSS2-expressing cells to determine
178 their sensitivity to lentiviral pseudotypes bearing the HKU1B spike (Figure 2E). We previously reported that
179 mutations abrogating TMPRSS2 protease activity (R225Q and S441A) mediate viral infection at similar
180 levels as the wild-type protein because the pseudoviruses can follow an endocytic route of entry[1]. While
181 the mutation $R316A_T$ did not affect pseudovirus infection, the TMPRSS2 variants that showed reduced

binding to the soluble spike were also inefficient for viral entry. The Y416A$_T$ mutant did not trigger pseudovirus entry (Figure 2E), nor induced cell-cell fusion (Figure 2C).

Overall, as summarized in the heatmaps (Figure 2F), our characterization of more than 30 point-mutants indicates that changes in residues at the interface of the HKU1B-RBD/TMPRSS2 complex impair binding, fusogenicity and infection, validating the crystal structure.

TMPRSS2 residues at position 417 and 469 are key to define host tropism.

We identified six positions in the TMPRSS2 SP domain - Y414$_T$, Y416$_T$, D417$_T$, L419$_T$, L430$_T$, W461$_T$, L469$_T$ and R470$_T$ - as the most relevant for the interaction with the HKU1B RBD. We then examined the pattern of conservation of these residues across different TMPRSS2 orthologs by aligning 201 mammalian TMPRRS2 sequences between amino acids 340 and 473 (numbering corresponding to human TMPRSS2), which span the interface with the HKU1B RBD. The resulting sequence logo (Figure 3A) indicated that among the residues listed above, Y416$_T$, L419$_T$, L430$_T$ and W461$_T$ are strictly conserved, while Y414$_T$ is conserved in more than 80% of the analyzed species. In contrast, the variability is high at positions 417 and 469, where D and Y, the respective residues in humans, are not frequently found. Many species have a non-charged residue in the place of human D417, while polar (N) or non-polar (L, F) residues are found in the place of human Y469. These observations suggest that residues at positions 417 and 469 may be determinants of TMPRSS2 function as HKU1 receptor in different mammalian species. We tested this hypothesis by assessing the capacity of TMPRSS2 from selected mammals (macaque, mouse, hamster, and ferret), to induce HKU1-CoV cell-cell fusion and pseudovirus infection. The respective TMPRSS2s display [D;N], [N;L], [N;L], and [N;N] instead of [D;Y] as in humans at positions [417;469] (Figure 3D). Despite similar expression levels (FigureS2C), only human TMPRSS2 induced syncytia in 293T cells co-expressing the HKU1B spike (Figure 3B, left panel). The macaque TMPRSS2 (which differs only at position 469), allowed reduced pseudovirus entry (Figure 3B, right panel). Introducing the N469Y mutation into macaque TMPRSS2 restored syncytia formation and increased pseudovirus infection by ~10-fold (Figure 3C). Changing Q467 on the macaque protein to the mostly conserved lysine did not show further significant effects on cell-cell fusion or pseudotype entry (Figure 3C). With mouse TMPRSS2, the double mutant engineered to change the [N;L] to the human [D;Y] motif restored syncytia formation and pseudovirus entry, while the two single mutants were poorly or non-functional (Figure 3C).

Nanobody A07 inserts its CDR3 into the TMPRSS2 substrate-binding groove.

We next crystallized the complex between TMPRSS2^{S441A} and the inhibitory nanobody A07 (Figure 4A, FigureS3) and determined its structure to 1.8 Å resolution (Table 1). A07 covers the active site cleft and buries a large surface area (about 2600 Å^2, ~1400 Å^2 on the VHH and ~1200 Å^2 on TMPRSS2), interacting with residues in exposed loops of the SP domain (loops 1, 2, 3, A, B, C, D), some also involved in the

interaction with HKU1B RBD (Figure 4A). Superimposing the TMPRSS2^{S441A}/RBD structure on the TMPRSS2^{S441A}/A07 complex showed clashes between the nanobody and the RBD (Figure 4B), explaining the blocking activity and further validating the interaction site that we report here for the RBD.

Nanobody A07 contacts TMPRSS2 almost exclusively through its CDRs. The most notorious feature is the insertion of its long (21 residues) CDR3 in the substrate binding cleft in between the two lobes of the SP domain. Superposing the complex with the peptide-bound structures of the homologous hepsin and TMPRSS13 proteases shows that the side chain of R103$_{A07}$ occupies theTMPRSS2 P1 position, making contacts with residues D435$_T$, S436$_T$ and G464$_T$ (Figure 4C), which form the S1 site [21].

The TMPRSS2^{S441A}/A07 crystals displayed electron density for the TMPRSS2 LDLR-A domain, which was not resolved in the previous structure of TMPRSS2[21]. The LDLR-A domain includes a calcium-binding site formed by the side chains of D134$_T$, H138$_T$, D144$_T$, and E145$_T$, and the main chain carbonyl group of N131$_T$ and V136$_T$ (FigureS4A).

The SP domain of cleaved TMPRSS2 (TMPRSS2-CI) crystallized previously in complex with the Nafamostat inhibitor [21], superposes very well with its counterpart in the TMPRSS2^{S441A}/A07 crystals, with a root mean square deviation (RMSD) of 0.315 Å for 1502 atoms in 236 residues. In this structure, the residues immediately downstream the autocleavage site (I256$_T$ and V257$_T$) are found in an internal pocket where the free amino group of I256$_T$ forms a salt-bridge with the side chain of D440$_T$ (Figure 4D). This interaction can only be established after cleavage of the protease and is a characteristic feature of serine proteases in the active conformation [33]. This observation strongly suggests that TMPRSS2^{S441A} in the crystals of the complex with A07 underwent cleavage by a contaminating protease. The crystallographic data further indicated that the maturation cleavage occur in 53% of the molecules forming the crystal. Another conformation was detected, for which the high-resolution diffraction allowed the refinement of an atomic model (Table 1). In this second model, which refined to an occupancy of 47%, the loop bearing the cleavage site was disordered, with the first residue with clear electron density after the cleavage site being G259$_T$, as expected for an uncleaved form of TMPRSS2^{S441A}. We confirmed that this is indeed the case by crystallizing the TMPRSS2^{S441A}/A07 complex under different conditions, which yielded crystals that diffracted to 2.4 Å resolution. The residues I256$_T$ and V257$_T$ were not visible in the structure determined from these new crystals, in which TMPRSS2^{S441A} showed no evidence of proteolysis and was present with an occupancy of 100%. The resulting model aligned very well with the second conformation described above, supporting the hypothesis that it corresponds to the TMPRSS2 zymogen. Furthermore, this structure has the characteristic “zymogen triad” initially observed for chymotrypsinogen[34], corresponding to D440-H279-S272 in TMPRSS2 (FigureS4B). Upon activation, D440 is released from this triad to make the salt-bridge/hydrogen bond with the newly formed N-terminus at I256. This rearrangement leads to formation of the oxyanion hole required for cleavage of the scissile peptide bond of the substrate. Proteases that do not have a zymogen triad, such as the tissue-type plasminogen activator, have a high level of catalytic activity in the zymogen form[35].

Comparison of the structures of cleaved and zymogen TMPRSS2^{S441A} revealed that the main changes are localized to loops 1 (L430$_T$-D440$_T$) and 2 (G462$_T$-K467$_T$). In the zymogen, residues 430-440 occupy an external position and S1 site residues (D435$_T$, S436$_T$) are away from the active site (Figure 4D).

As mentioned above, in TMPRSS2^{S441A}-CI the free N-terminus of I256$_T$ flips to the interior of the molecule, displacing the 430-440 segment towards the core of the domain (Figure 4D). The disulfide bond between C437$_T$ and C465$_T$ propagates this movement to loop 2 (G462$_T$-K467$_T$), which also adopts a different conformation (Figure 4D). Consequently, the S1 site is formed without alteration of the catalytic triad. Of note, the conformation of loops 1 and 2 in TMPRSS2^{S441A}-CI is the as that observed in the active forms of hepsin[36] and TMPRSS13[37] (FigureS4C).

TMPRSS2 maturation affects binding affinity towards the HKU1-CoV RBD

As described above, maturation of TMPRSS2 into an active form implies changes in the conformation of loops 1 and 2. Since they are part of the HKU1B RBD binding site, we hypothesized that TMPRSS2 cleavage may impact the interaction with the RBD. We therefore conducted BLI experiments with immobilized RBD, soluble TMPRSS2^{S441A} and cleaved soluble TMPRSS2^{S441A} (TMPRSS2^{S441A}-CI). The latter was obtained by digestion of the zymogen with wild-type TMPRSS2. Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) confirmed the complete digestion of TMPRSS2^{S441A}, generating fragments of ~25-30 kDa associated by a disulfide bond (FigureS5A). The BLI curves showed that TMPRSS2^{S441A}-CI displays a dramatic increase in the association signal and reduced dissociation rate (Figure 5A). The curves corresponding to TMPRSS2^{S441A}-CI fitted well to a 1:1 binding model. We determined the kinetic association and dissociation rates and measured a dissociation constant (Kd) of 30 nM (Figure 5C). This value is ~5-fold higher than the one obtained for TMPRSS2^{S441A}-zymogen[1]. We also used TMPRSS2^{S441A}-CI to determine the affinity of some of the RBD mutants (Figure 5B). This confirmed that K487A$_R$, L510R$_R$, and Y528A$_R$ were the most impaired (Figure 5C), while S529A$_R$ showed increased affinity, mostly due to a lower dissociation rate. Additionally, we tested the effect of the RBD mutants on the *in vitro* TMPRSS2 activity and determined that their ability to inhibit the protease correlated with the affinity towards TMPRSS2 (FigureS5B).

We then compared the structures of TMPRSS2^{S441A}-CI and TMPRSS2^{S441A}-zymogen (obtained with A07) with that of TMPRSS2^{S441A} in complex with the RBD (Figure 5D). In the latter structure, residues I256$_T$ and V257$_T$ were not visible, indicating that TMPRSS2^{S441A} is in the uncleaved form. Overall, the protease conformation matched TMPRSS2^{S441A}-zymogen, except for local differences in the loops at residues around G432$_T$ and S463$_T$. In the complex with the RBD, these loops adopt a conformation closer to the one found in TMPRSS2 ^{S441A}-CI, where G432$_T$ moves towards the active site, further away from the RBD interface, and S463$_T$ flips outwards to form a hydrogen bond with the RBD (Figure 5D) explaining why the affinity of TMPRSS2^{S441A}-CI for the RBD is higher than that of the zymogen.

DISCUSSION

We report the structure of HKU1B RBD bound to its receptor TMPRSS2, revealing a unique binding mode among coronaviruses. The HKU1-CoV RBD uses a pincer plier motif at its distal end to recognize the TMPRSS2 substrate specificity loops. The interaction blocks access of substrates to the catalytic groove, inhibiting the proteolytic activity. Our results are in line with a recent report on the structure of the HKU1A RBD complexed with cleaved TMPRSS2 and the inhibition of TMPRSS2 proteolytic activity by HKU1A RBD[38].

The HKU1-CoV spike protein binds cell-surface disialoside glycans (9-O-Ac-Sia(α2,8)Sia) through a binding pocket in the NTD, causing the allosteric opening of the spike[13]. This observation led to the conclusion that the sialo glycan is a primary receptor, while a proteinaceous receptor would be secondary. This mechanism ensures that the spike will not undergo premature activation before reaching a cell, limiting exposure of the receptor binding motif to neutralizing antibodies. Moreover, when interacting with an up-RBD, the active site of TMPRSS2 is far from the spike core, indicating that a single TMPRSS2 molecule cannot simultaneously cleave the spike and act as receptor. By binding to up-RBD conformations, TMPRSS2 traps partially open intermediates of the HKU1 spike, displacing the conformational equilibrium towards the open form, where S2 is unshielded and primed for fusion, acting by the same mechanism as the angiotensin-converting enzyme 2 (ACE2) with the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) spike[39]. Considering that ~45% of HKU1-CoV spike with 9-O-Ac-Sia(α2,8)Sia were found closed or with only one RBD 'up' [13], a secondary protein receptor with high affinity towards the RBD would guarantee the unidirectionality of the conformational change towards the open spike.

We also determined the crystal structure of TMPRSS2 bound to nanobody A07, which competes with the RBD for binding[1]. A07 inserts its long CDR3 in the active site cleft, with $R103_{A07}$ occupying the P1 position (Figure 4E). It has been suggested that the predicted S1' site ($V280_T$, $H296_T$, $C297_T$) accepts small hydrophobic P1' residues[21]. Our structure shows $R104_{A07}$ in the S1' site and its side chain is well accommodated, interacting with main chain atoms from $H296_T$ and $C297_T$. Given that the peptide bond $R103_{A07}$-$R104_{A07}$ lies in the right position to undergo the attack of the active site S441, it is possible that the nanobody is cleaved by TMPRSS2. Nevertheless, we expect the extensive network of polar interactions and the large BSA to sustain binding even if the antibody is cleaved.

The structures of cleaved and uncleaved forms of TMPRSS2 highlight the internal reorganization of the protease upon autocleavage. In the zymogen form, loop 1 adopts a conformation where S1 residues are away from the active site and are not in an optimal position to allow substrate binding, as illustrated in Figure 4E (left panel). The proteolytic cleavage releases the N-terminus of $I256_T$, which flips towards the interior of the protein, with formation of a salt bridge between the charged amino group of $I256_T$ and $D440_T$. In turn, loop 1 is displaced and adopts a new conformation, transmitting this movement to loop 2 via a disulfide bond connecting them. Consequently, the S1 site residues ($D435_T$ and $S436_T$ from loop 1, and

G464$_T$ from loop 2) are brought into place to allow substrate binding. These changes agree with the general activation mechanism proposed for other serine proteases[40]. Furthermore, the TMPRSS2 loops of the cleaved form adopt the same position as in other enzymes of the same family (hepsin and TMPRSS13), suggesting a stable conformation in the active form. The structure of the A07/TMPRSS2 zymogen further shows how the substrate mimicking CDR3 is accommodated in the groove in the absence of the S1 site (Figure 4E, right panel), which forms only after activation. Together with the inherent plasticity of the zymogen, this data can potentially inform the development of TMPRSS2-specific small molecules targeting this site.

We also established that TMPRSS2 activation affects the interaction with the HKU1-CoV RBD. In zymogen TMPRSS2^{S441A}, residues 431-433 (loop 1) adopt a conformation that may not favor binding to the RBD, which has the j2 jaw of the pincer plier facing it. Moreover, S463 in loop 2 is too far to interact with D507$_R$. Upon cleavage, the 431-433 segment moves towards the core of TMPRSS2, facilitating the approach of the RBD, and S463 moves outwards, forming a hydrogen bond with D507$_R$. These changes in loops 1 and 2 appear to be important for binding, since the structure of zymogen TMPRSS2^{S441A} with the RBD shows that they adopt a position similar to that observed in activated TMPRSS2^{S441A}-CI. This 'mixed' conformation, with local changes restricted to segments 431-433 and 463-464 might represent a high-energy state, which would decrease the affinity. In addition, the zymogen TMPRSS2^{S441A} loops could adopt the ensemble of observed conformations, having a higher entropic content than TMPRSS2^{S441A}-CI and resulting in lower affinity towards the HKU1-CoV RBD. The BLI curves show that zymogen TMPRSS2^{S441A} does not fit to a 1:1 binding model, while TMPRSS2 ^{S441A}-CI does, in line with a heterogeneous population of conformations that becomes homogeneous upon cleavage.

In sum, the structure of the HKU1B RBD in complex with TMPRSS2 provides insights into the mode of action of the receptor and on the mechanism of spike activation. We also identify the most important residues of the interface, highlighting how they contribute to host species tropism, paving the way to characterization of potential animal reservoirs. Finally, we describe the maturation of TMPRSS2 from a zymogen to an active protease and how it impacts binding to HKU1-CoV. The observed plasticity of the TMPRSS2 zymogen stands out as a vulnerable feature of the protease that can be specifically targeted to prevent activation. Our results therefore open the way to the development of specific drugs against dysregulated TMPRSS2 in tumors without affecting other serine proteases to avoid toxicity.

Limitations of the study

We did not study the interaction of HKU1A with TMPRSS2. We previously reported that both HKU1A and HKU1B bind to and use TMPRSS2 as a functional receptor[1]. It is thus likely that the two HKU1 spikes similarly interact with TMPRSS2. The recent description by cryo-EM of a TMPRSS2 / HKU1A RBD complex[38], revealed a similar mode of interaction for both HKU1A and B. Also, we did not determine the structure of the trimeric HKU1-CoV spike in association with the receptor. We circumvented this problem

by superposing the structure of the RBD in complex with TMPRSS2 to the known structure of the HKU1 spike.

ACKNOWLEDGEMENTS

We thank the staff from the Crystallography platform at Institut Pasteur and the synchrotron source SOLEIL (Saint-Aubin, France) for granting access to the facility. We thank the staff of the beamlines Proxima 1, especially Pierre Legrand, and Proxima 2A for their advice and assistance during X-ray data collections. We thank Pablo Guardado-Calvo for access to the Octet-R8 instrument and Patrick England from the Molecular Biophysics platform at Institut Pasteur for useful discussions on biolayer interferometry data analysis.

This work was supported by the 'URGENCE COVID-19' fundraising campaign of Institut Pasteur "Allospike" project. F.A.R. is also funded by ANR grants ANR-18-CE92-0006 PHLEBO, ANR-10-LABX-62-10 IBEID, Wellcome Trust collaborative grant no. UNS22082 and the Institut Pasteur and the Centre national de la recherche scientifique. O.S. is funded by Institut Pasteur, Fondation pour la Recherche Médicale, Humanities in the European Research Area programme (DURABLE consortium and LEAPS consortium), the National Agency for AIDS Research–Emerging Infectious Diseases, the Vaccine Research Institute (ANR-10-LABX-77), Agence Nationale de la Recherche (ANR)/Fondation pour la Recherche Médicale Flash Covid PROTEO-SARS-CoV-2 and IDISCOVR, and IBEID Labex. Work with UtechS Photonic BioImaging is funded by grant no. ANR-10-INSB-04-01 and Région Ile-de-France programme DIM1-Health.

AUTHOR CONTRIBUTIONS

The experimental strategy was designed by I.F, N.S., S.D., J.B., O.S. and F.A.R. Construct design, protein expression and protein purification were performed by I.F., A.A and E.B.S. Crystallization and X-ray diffraction data collection were carried out by I.F. and A.H., while data processing was performed by I.F. and S.D. Cell-cell fusion, spike binding and pseudovirus entry experiments, as well as the mutagenesis of the plasmids used for these assays were performed by N.S., W.H.B. and J.B. The biolayer interferometry experiments were performed by I.F. The original manuscript draft was written by I.F. and F.A.R., with input from N.S., S.D., J.B. and O.S. The funding was acquired by O.S. and F.A.R.

DECLARATION OF INTERESTS

I.F., N.S., E.B., P.L. J.B., O.S. and F.A.R. have a provisional patent on anti-TMPRSS2 nanobodies.

F.A.R. is a founder of *Meletios Therapeutics* and a member of its scientific advisory board. The other authors declare no competing interests.

FIGURE LEGENDS

Figure 1. The HKU1B-RBD recognizes TMPRSS2 with a unique binding mode

(A) Schematic representation of the HKU1B SD1-RBD and TMPRSS2 constructs used for crystallization experiments. The different (sub)domains are indicated with colors, and the residues at the boundaries are below the scheme. TMPRSS2 serine-peptidase (SP) domain is formed by two β-barrels, shown in different colors. Important TMPRSS2 residues, such as H296-D345-S441 (catalytic triad, indicated with a star) and D435 (part of the S1 site) are on top of the scheme. (S441) indicates that this active site residue was mutated to alanine in the crystallized structure. The scissor indicates the TMPRSS2 autocleavage site.

(B) Crystal structure of the HKU1B-RBD in complex with TMPRSS2^{S441A}. Crystals were obtained for the ternary complex with VHH-A01 but for better clarity the nanobody is not displayed (the structure of the ternary complex is displayed in Figure S1B). Both proteins are colored according to the (sub)domain code presented in panel (A). Important elements on the RBD, such as the loops that form jaws 1 (j1) and 2 (j2) are indicated. TMPRSS2 loops at the interface are labeled L1 (loop 1, residues 427-441), L2 (loop 2, residues 462-471), L3 (loop 3, residues 412-423) and LC (loop C, residues 334-346).

(C) Description of relevant contacts at the interface between the RBD and TMPRSS2^{S441A}. The left panel shows polar residues that form salt bridges or hydrogen bonds (dashed lines), while the right panel shows hydrophobic and aromatic residues. The subscript of each residue indicates if it is part of the RBD (R) or TMPRSS2 (T). The residues that form the catalytic triad (H296, D345, S441A) are colored in purple.

(D) Open-book representation of the RBD-TMPRSS2^{S441A} complex. The contact surface was colored indicating residues that form polar interactions (dark blue for RBD residues, light magenta for TMPRSS2), hydrophobic and aromatic residues (white), and residues that are at the interface and establish van der Waals contacts (light blue for the RBD and pink for TMPRSS2^{S441A}).

(E) Scheme highlighting residues buried at the RBD-TMPRSS2^{S441A} interface (blue/pink shade). Amino acids that form salt bridges are indicated with filled triangles, while those that form hydrogen bonds are indicated with the empty symbol. For comparison, the sequence from the HKU1A RBD (isolate N1) is aligned below the one from HKU1B. Residues that are important for TMPRSS2 activity are indicated with stars (purple: catalytic triad; red: S1 site; orange: S2 site). Loops and β-strands are indicated, as in the previous panels.

(F) Comparison between the structures of different coronavirus RBDs (light blue) in complex with their receptors (dark pink). Two examples of betacoronaviruses (SARS-CoV-2 in complex with ACE2, PDB: 6M0J; MERS-CoV in complex with DPP4, PDB: 4L72), two of alphacoronaviruses (NL63-CoV with ACE2, PDB: 3KBH; HCoV-229E with APN, PBD: 6U7G) and one of a deltacoronavirus (PDCoV and human APN, PDB: 7VPQ) were selected. The buried surface area (BSA) of each complex is indicated below.

(G) The structure of the RBD-TMPRSS2^{S441A} complex was aligned to the RBD from the closed HKU1A Spike (apo-closed, PDB: 8OHN), the closed Spike with a disialoglycan (Sia) in the NTD (holo-closed, PDB: 8OPM) and the Spike in an open form with a disialoglycan (Sia) (holo-1-RBD-up, PDB: 8OPN). An inset on the first panel zooms into the region indicated with an oval to better visualize the clashes between TMPRSS2 bound to a protomer in the 'down' conformation and amino acids and glycans from the other chains. Each protomer of the Spike is indicated with different colors (dark blue, light blue and white) and glycans are colored in green.

(H) 293T cells expressing either GFP1-10 or GFP11 (GFP split system) were transfected with TMPRSS2 and HKU1B (wild-type or harboring the W89A mutation). Cell-cell fusion was quantified by measuring the GFP area after 20 h. Data are mean ± s.d. of three independent experiments. The dotted line indicates a normalized fusion of 1.0 relative to wild-type Spike. Statistical analysis: (H) One-way ANOVA with Dunnett's multiple comparison test compared to the WT TMPRSS2 on non-normalized log-transformed data. *** $p<0.001$.

See also Figure S1 and Table S1.

Figure 2. Functional experiments validate the RBD-TMPRSS2 crystal structure

(A) 293T cells expressing either GFP1-10 or GFP11 (GFP split system) were transfected with HKU1B Spike and TMPRSS2, and fusion was quantified by measuring the GFP area after 20 h. Data are mean ± s.d. of three independent experiments. The dotted line indicates a normalized fusion of 1.0 relative to wild-type TMPRSS2.

(B) Biolayer interferometry (BLI) experiment performed with soluble HKU1B RBDs containing mutations in the binding interface. The RBDs were immobilized to Ni-NTA sensors and their interaction with soluble TMPRSS2^{S441A} (800 nM) was followed in real-time. One representative experiment of three is presented.

(C) Cell-cell fusion experiment carried out transfecting the wild-type HKU1B Spike with TMPRSS2 WT or different mutants. Fusion was quantified by measuring the GFP area after 20 h. Data are mean ± s.d. of four independent experiments. The dotted line indicates a normalized fusion of 1.0 relative to wild-type TMPRSS2.

(D) 293T cells were transfected with plasmids coding for TMPRSS2 or mutated variants, they were incubated with biotinylated soluble HKU1B Spike and binding to cells was determined by flow cytometry using labeled streptactin. Data are mean ± s.d. of three independent experiments. The upper dotted line indicates the percentage of transfected cells with wild-type TMPRSS2 that were considered positive, while the bottom dotted line shows the background levels in non-transfected cells.

(E) 293T cells transfected with TMPRSS2 WT or mutant variants were infected by luciferase-encoding HKU1B pseudoviruses. Luminiscence was read 48 h postinfection. RLU, relative light unit. The

upper dotted line indicates the mean RLU obtained when cells were transfected with wild-type TMPRSS2, while the lower dotted line shows the background levels in non-transfected cells. Data are mean ± s.d. of three independent experiments.

(F) Heatmaps summarizing the functional data from panels A to E, along with the results from controls of expression and enzymatic activity (supplementary information).

Statistical analysis: (A, C, E) One-way ANOVA with Dunnett's multiple comparison test compared to the WT on non-normalized log-transformed data. (D) One-way ANOVA with Dunnett's multiple comparison test compared to the WT. * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

See also Figure S2.

Figure 3. TMPRSS2 residues at positions 417 and 469 influence binding to different homologues

(A) Frequency of different amino acids occupying positions 409 to 473 (numbers at the bottom) of TMPRSS2 orthologs. The logo was obtained using 201 sequences from mammals. Purple inverted triangles indicate the positions identified as relevant for human TMPRSS2 in functional experiments (Figure 2).

(B) Cell-cell fusion (left panel) and pseudovirus entry (right panel) experiments performed with cells transfected with TMPRSS2 from different species. Data are mean ± s.d. of three independent assays. The dotted line indicates a normalized fusion of 1.0 relative to human TMPRSS2 or the mean RLU obtained with human TMPRSS2.

(C) Cell-cell fusion (left panel) and pseudovirus entry (right panel) experiments performed with cells transfected with wild-type or mutant TMPRSS2 from selected mammals. Data are mean ± s.d. of three independent experiments. The dotted line indicates a normalized fusion of 1.0 relative to human TMPRSS2 or the mean RLU obtained with human TMPRSS2.

(D) Sequence alignment of the TMPRSS2 homologues used for the experiments in panel (B), highlighting the residues that are part of the interface formed with the HKU1-RBD (pink shade) and indicating TMPRSS2 loops L1 (loop 1), L2 (loops 2), L3 (loop 3) and LC (loop C).

Statistical analysis: (B, C) One-way ANOVA with Šídák's multiple comparison test * $p<0.05$, *** $p<0.001$.

See also Figure S2.

Figure 4. VHH-A07 blocks TMPRSS2 activity and binding to the HKU1-RBD by inserting its CDR3 in the active site

(A) Crystal structure of the nanobody VHH-A07 (yellow) complexed to TMPRSS2^{S441A}. The two β-barrels that form the SP domain are colored in gray (β-barrel 1) and light red (β-barrel 2). Important structural elements are indicated in the nanobody ($CDR1_{A07}$, $CDR2_{A07}$, $CDR3_{A07}$), as well as in

TMPRSS2^{S441A} (loops B, C, D, 1, 2, 3). Residues from the catalytic triad are shown in purple. Subscripts in the labels identify the protein. The active site S441 was mutated to alanine in the crystallized construct, so it is annotated between parentheses.

(B) Superposition of the TMPRSS2^{S441A}+RBD and TMPRSS2^{S441A}+VHH-A07 complexes. For simplicity, TMPRSS2 from the complex with the RBD is not shown.

(C) Superposition of TMPRSS2^{S441A} in complex with VHH-A07 (protease in light red, nanobody in yellow) with the structures of TMPRSS13 (green, PDB: 6KD5) and Hepsin (gray, PDB: 1Z8G) bound to substrate-analog inhibitors (decanoyl-RVKR-chloromethylketone shown in green sticks and Ac-KQLR-chloromethylketone in gray ticks, respectively). For clarity, the proteases are shown in ribbon representation and the nanobody in cartoon. Residues from the TMPRSS2 catalytic triad are indicated in purple, with S441 labeled with parenthesis to indicate that it is mutated to alanine in the crystalized construct. Residues that form the S1 site (D435, S436, G464) are indicated, as well as R103 from nanobody A07.

(D) Superposition of different TMPRSS2^{S441A} structures obtained in complex with VHH-A07. A first crystal form (crystal 1) allowed building two atomic models, where one has the characteristic features of a cleaved (Cl, light red) serine protease, while the other represents the zymogen (UNCl, wheat) form. The second crystal form (crystal 2) allowed building the model shown in orange that corresponds to uncleaved TMPRSS2^{S441A}. A previously reported structure of active TMPRSS2 is shown in white for comparison (PDB: 7MEQ). The I256 residue, which follows the autocleavage site, is indicated. Important TMPRSS2 loops are designated as well as the positions of the catalytic triad (H296, D345, S441A) and D435 from the S1 site. The disulfide C437-C465 is shown in yellow sticks. Arrows with tapered lines indicate the repositioning on relevant segments upon cleavage.

(E) Comparison between the TMPRSS2^{S441A} active site in the uncleaved (left) and cleaved (right) forms found in crystal 1. The surface of the protease was colored according to the electrostatic potential and the residues R99$_{A07}$-W106$_{A07}$ from A07 CDR3 are shown in sticks.

See also Figures S3 and S4.

Figure 5. TMPRSS2 cleavage affects binding to the HKU1 RBD.

(A) Curves of a BLI experiments performed immobilizing the HKU1B RBD and measuring the response upon incubation with 150 nM TMPRRS2^{S441A} or with 150 nM TMPRSS2^{S441A} previously cleaved by incubation with the wild-type protease. Two separate cleavage reactions were carried out. The colored curves correspond to the experimental data and the thin black lines are the curves fitted to a 1:1 binding model. One representative experiment of three is presented.

(B) BLI curves of experiments performed immobilizing different RBDs and measuring the response after incubation with 120 nM cleaved TMPRSS2^{S441A}. The colored curves correspond to the

experimental data and the thin black lines are the curves fitted to a 1:1 binding model. One representative experiment of three is presented.

(C) Kd values of different RBDs towards cleaved TMPRSS2^{S441A}. Three experiments were performed for each RBD using a TMPRSS2 concentration range from 7.5 to 240 nM.

(D) Superposition of the TMPRSS2^{S441A} structure determined in complex with the HKU1-RBD (TMPRSS2 in pink, RBD in blue surface) or in complex with VHH-07 (crystal 1: atomic model of the cleaved form shown in light red, and the zymogen form in wheat). Loops that form the binding site and change upon cleavage (loops 1 and 2) are labeled. The positions where conformational changes were observed are indicated with the residues that are implicated (backbone atoms of G432 and side chain of S463). Arrows with tapered lines indicate the movement of relevant segments upon TMPRSS2 autocleavage.

See also Figures S5 and S6.

TABLES

Table 1. Data collection and refinement statistics

	RBD+TMPRSS2+VHH-A01	TMPRSS2+VHH-A07	TMPRSS2+VHH-A07
PDB code	8S0M	8S0L	8S0N
Data Collection			(staraniso)
Space Group	P 6_1	C 1 2 1	C 1 2 1
a, b, c (Å)	201.9, 201.9, 210.3	182.7, 53.8, 65.5	161.9, 54.5, 165.8
a, b ,g (deg)	90.0, 90.0, 120.0	90.0, 100.7, 90.0	90.0, 108.4, 90.0
Resolution (Å)	25-3.55 (3.64-3.55)	25-1.80 (1.85-1.80)	25-2.30 (2.39-2.30)
Rmerge	0.265 (12.876)	0.106 (2.643)	0.240 (2.236)
Mean((I)/sd(I))	14.6 (0.4)	9.2 (0.6)	6.6 (0.9)
Number of reflections	58457 (4358)	58045 (4194)	50880 (2545)
Completeness (%)	99.6 (100.0)	99.8 (98.3)	82.7 (39.4)
Comp. (ellipsoidal)			95.0 (79.2)
Multiplicity	43.3 (41.4)	6.4 (6.3)	7.0 (6.9)
CC(1/2)	1.00 (0.223)	0.998 (0.293)	0.990 (0.316)
Refinement			
Resolution (Å)	25 - 3.55	25 - 1.80	25 - 2.30
Number of reflections	58058	57982	50824
Rwork / Rfree	19.29 / 22.13	17.69 / 21.54	20.87 / 24.71
NCS	2		2
No. of atoms			
Protein	12567	4003	7032
Ligand / Carb / Water	95	389	207
B-factor			
Clashscore	6.67	3.70	3.92
R.m.s deviations			
Bond lenghts (Å)	0.004	0.010	0.002
Bond angles (°)	0.788	1.013	0.493
Ramachandran			
Favored (%)	96.05	96.15	96.52
Outliers (%)	0.13	0.20	0.11
Rotamer outliers	1.00	0.46	0.26

Diffraction limit #1: 2.796:(0.8883, 0.0000, 0.4594): 0.984 **a*** + 0.177 **c***
Diffraction limit #2: 2.101:(0.0000, 1.0000, 0.0000): **b***
Diffraction limit #3: 2.241:(-0.4594, 0.0000, 0.8883):-0.413 **a*** + 0.911 **c***

Table 2. Buried surface area and number of polar interactions in the crystal structures.

TMPRSS2 Partner	Partner ∆ASA (Å^2)	TMPRSS2 ∆ASA (Å^2)	Number of H-Bonds	Number of salt bridges
HKU1B RBD*	843	820	10	2
	850	840	8	1
VHH-A01*	1242	1143	11	1
	1287	1195	10	1
VHH-A07 crystal 1 Cleaved	1439	1219	13	5
VHH-A07 crystal 1 Zymogen	1303	1145	9	3
VHH-A07 crystal 2*	1044	1193	11	3
	1047	1182	11	4

* 2 complexes in crystal asymmetric unit

SUPPLEMENTAL FIGURE LEGENDS

Figure S1 (Related to Figure 1). Purification of the ternary complex TMPRSS2^{S441A}/RBD-SD1/VHH-A01.

(A) Size-exclusion chromatography of the binding reaction prepared with HKU1B RBD, TMPRSS2^{S441A} and VHH-A01. The eluate was collected in different fractions (numbers indicated above the chromatogram) and 10 µl aliquots of some of them were analyzed under reducing SDS-PAGE (bottom panel). MWM: molecular weight marker.

(B) Crystal structure of the ternary complex.

(C) *In vitro* TMPRSS2 enzymatic activity. Cleavage of a fluorescent substrate (Boc-QAR-AMC) by soluble TMPRSS2 was evaluated upon incubation with different concentrations of the HKU1B RBD or SARS-CoV-2 RBD (control). The initial kinetic slope of the reaction was measured and normalized to the non-treated condition. Data are mean ± s.d. of three independent assays. The plot represents a subset of a larger panel of RBDs presented in Figure S5.

Figure S2 (Related to Figures 2 and 3). Expression and activity of Spike mutants, TMPRSS2 mutants and TMPRSS2 orthologs.

(A) Expression of HKU1B Spike constructs determined by FACS using an anti-S2 antibody. Data are mean ± s.d. of two to four independent assays.

(B) Surface expression of TMPRSS2 mutants assessed on live cells by FACS with anti-TMPRSS2 VHH-A01-Fc (top). Cell-cell fusion assay performed by co-transfecting cells with HCoV-229E Spike, APN (aminopeptidase N, the receptor) and a TMPRSS2 variant from the panel (bottom panel). Data are mean ± s.d. of three to four independent experiments.

(C) Expression of TMPRSS2 orthologs and their point mutants evaluated by intracellular staining of their c-myc epitope (left panel). Cell-cell fusion assay performed by co-transfecting cells with HCoV-229E Spike, APN and one TMPRSS2 ortholog or point mutant (right panel). Data are mean ± s.d. of three (top panels) or four (bottom panels) independent experiments.

(D) Statistical analysis: (A, B, C) One-way ANOVA with Dunnett's multiple comparison test compared to the WT. For fusion assay One-Way ANOVA was performed on the non-normalized log transformed data. * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

Figure S3. (Related to Figure 4). Purification of the TMPRSS2^{S441A}+VHH-A07 complex.

TMPRSS2^{S441A} was incubated with an excess of VHH-A07 and the binding reaction was injected onto a size-exclusion chromatography column. The elution profile is shown, with two arrows indicating fractions that were analyzed by reducing SDS-PAGE (bottom). MWM: molecular weight marker.

Figure S4 (Related to Figure 4). TMPRSS2^{S441A} structure.

(A) Structure of the LDLR-A domain obtained from crystal form 1 of the TMPRSS2^{S441A}+VHH-A07 complex. A central calcium ion is shown, as well as the two backbone carbonyls (subscript CO) and side chains that coordinate it. Disulfides C133-C148 and C120-C139 are shown with yellow sticks.

(B) Superposition of the structures obtained for cleaved TMPRSS2^{S441A} in the crystal 1 (TMPRSS2^{S441A}-Cl, red), TMPRSS2^{S441A} zymogen in crystal 1 (TMPRSS2^{S441A}-UNCl, wheat) and TMPRSS2^{S441A} zymogen from crystal 2 (TMPRSS2^{S441A}-UNCl, orange) showing the zymogen triad (S272, H279, D440). Upon autocleavage, the conformational change in loop 1 shifts the position of D440 (indicated by the subscript 'Cl').

(C) Superposition of cleaved TMPRSS2^{S441A} (from crystal 1) with active proteases from its subfamily (TMPRSS13, green; hepsin, gray). The loops 1, 2 and 3 (L1, L2 and L3, respectively) are labeled.

Figure S5 (Related to Figure 5). TMPRSS2^{S441A} proteolysis by TMPRSS2 wild-type and effect of mutations on the RBD ability to inhibit TMPRSS2 activity.

(A) SDS-PAGE under reducing and non-reducing conditions to show the cleavage of TMPRSS2^{S441A} upon incubation with the wild-type protease. Two reactions were performed under different conditions (producing 'batch 1'and 'batch 2') and 2 µl of each reaction (corresponding to ~2.5 µg of substrate) were analyzed by SDS-PAGE. For comparison, wild-type (WT) TMPRSS2 and untreated TMPRSS2^{S441A} were also loaded in the gel. M: molecular weight marker.

(B) *In vitro* TMPRSS2 enzymatic activity was evaluated upon incubation with different concentrations of the HKU1B RBDs harboring mutations at the binding interface. SARS-CoV-2 RBD was used as a control. The initial kinetic slope of cleavage of a fluorescent substrate (Boc-QAR-AMC) was measured and normalized to the non-treated condition. Data are mean ± s.d. of three independent assays.

Figure S6 (Related to Figure 5). Biolayer interferometry (BLI) experiments to determine the Kd of RBD mutants against cleaved TMPRSS2^{S441A}**.** Three independent assays (each represented in a column) were performed. The initial and final concentrations tested are indicated to the right of each plot, as well as the serial dilutions between them.

Figure S7 (Related to Methods and to Figure 2). Gating strategy for the detection of surface expression of the transfected proteins by flow cytometry.

STAR METHODS

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Félix Rey (felix.rey@pasteur.fr)

Materials availability

Reagents generated in this study are available on request from the Lead Contact with a completed Materials Transfer Agreement.

Data and code availability

Coordinates and structure factors of the crystal structures have been deposited in the Protein Data Bank (PDB) and are publicly available as of the data of publication. Accession numbers are listed in the key resources table. This paper does not report original code. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL

Cells

Functional experiments were performed using human female embryonic kidney 293 T cells (HEK293T or 293T) from the American Type Culture Collection (ATCC) that were cultured in Dulbecco's modified Eagle medium (DMEM) with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin (100X stock: 10,000 units penicillin and 10 mg/ml streptomycin) at 37 °C. GFP-split cells were previously described [41] and cultured with 1 µg/mL of puromycin (InvivoGen) at 37 °C. Cells were routinely screened for mycoplasma. Cells were authenticated by genotyping (Eurofins).

Recombinant proteins were expressed by stably transfected Drosophila Schneider line 2 (S2) cell lines or by transiently transfected Expi293F™ cells (Thermo-Fischer), as indicated in Method details. S2 cell lines were cultured in serum-free insect cell medium (HyClone, GE Healthcare) containing 7 µg/ml puromycin and 1% penicillin/streptomycin at 28 °C. Expi293F™ cells were maintained in Expi293™ expression medium (Gibco) without antibiotics at 37 °C. These cell lines were not authenticated. The supplier does not inform their sex and this information is not relevant for our study.

METHOD DETAILS

Construct design

For production of recombinant proteins

Codon-optimized synthetic genes coding for the HKU1B RBD (residues 330-614 of the spike protein, isolate N5P8, NCBI accession Q0ZME7), the RBD and subdomain-1 (RBD-SD1, residues 307-675), the spike ectodomain (residues 14-1276) and the human TMPRSS2 ectodomain (residues 107-492, NCBI accession O15393) were obtained from Genscript. Cloning and mutagenesis of these genes were also performed by Genscript. The RBD was cloned into pCAGGS, following a murine immunoglobulin kappa signal peptide, and upstream of a thrombin cleavage site and *in-tandem* Hisx8, Strep and Avi-tags. The Spike ectodomain was stabilized in the prefusion form by introducing mutations in the furin site (^{752}RRKRR^{756} to ^{752}GGSGS^{756}) and in the S2 subunit (^{1067}NL^{1068} to ^{1067}PP^{1068}) and adding a Foldon trimerization motif at the C-terminus. This construct was cloned into pcDNA3.1(+), with an Ig kappa signal peptide, a thrombin cleavage site at the C-terminus followed by a His-tag. RBD-SD1, TMPRSS2 wild-type (WT) and TMPRSS2^{S441A} were cloned into a modified pMT/BiP plasmid (Invitrogen; hereafter termed pT350), which translates the protein in frame with an enterokinase cleavage site and a double strep-tag at the C-terminal end. Selected VHHs have been previously obtained [1] and were cloned into the bacterial expression vector pET23 with a C-terminal His-tag and N-terminal Met-Ala residues introduced during subcloning.

For cell transfection and functional assays

Codon-optimized synthetic genes coding for the full-length spike of HKU1 B/C isolate N5P8 (referred to as HKU1B, UniProtKB/Swiss-Prot: Q0ZME7.1) and those coding for TMPRSS2 from mouse (Mouse C57BL/6 - UniPropt: Q3UKE3), ferret (Ferret – UniProt: A0A8U0SMZ2), hamster (Syrian Hamster Isoform X1 – UniProt: A0A1U8BWQ2) and macaque (Macaque – UniProt: F6SVR2) with an N-terminal cMYC-tag were ordered to GeneArt (Thermo Fisher Scientific) and cloned into a phCMV backbone (GeneBank: AJ318514) by replacing the VSV-G gene. pQCXIP-Empty control plasmid was previously described [41]. pQCXIP-BSR-GFP11 and pQCXIP-GFP1-10 were a kind gift from Yutaka Hata [42] (Addgene plasmid #68716 and #68715). pCSDest-TMPRSS2 was a kind gift from Roger Reeves [43] (Addgene plasmid # 53887). Human pCCAGS N-terminal cMYC-epitope tagged TMPRSS2 was a kind gift from Stefan Pöhlmann [44]. Mutations in the HKU1 spike and TMPRSS2 were introduced using the NEB Q5 Side-Directed mutagenesis kit. Plasmids were sequenced before usage.

Protein expression and purification

Protein expression and purification for X-ray crystallography

Plasmids encoding RBD-SD1 or TMPRSS2^{S441A} were co-transfected with the pCoPuro plasmid[45] for puromycin selection in Drosophila Schneider line 2 cells (S2) using the Effectene transfection reagent (Qiagen). The cell lines underwent selection in serum-free insect cell medium (HyClone, GE Healthcare) containing 7 µg/ml puromycin and 1% penicillin/streptomycin. For protein production, the cells were grown in spinner flasks until the density reached 10^7 cells/mL, at which point the protein expression was induced

with 4 μM $CdCl_2$. After 6 days, the cultures were centrifuged, and the supernatants were concentrated and used for affinity purification in a Strep-Tactin column (IBA). The strep tags were removed by incubating the proteins with 64 units of Enterokinase light chain (BioLabs) in 10 mM Tris, 100 mM NaCl, 2 mM $CaCl_2$, pH 8.0, at room temperature, overnight. The proteolysis reactions were buffer-exchanged into 10 mM Tris, 100 mM NaCl, pH 8.0, and subjected to a second affinity purification, recovering the flow-through fraction containing the untagged proteins. The proteins were concentrated and the enzymatic deglycosylation with endoglycosidase D (EndoD) and endoglycosidase H (EndoH) was set up at room-temperature following overnight incubation with 1000 units of each enzyme in 50 mM Na-acetate, 200 mM NaCl, pH 5.5. The proteins were further purified on a size exclusion chromatography (SEC) Superdex 200 16/600 (Cytiva) column in 10 mM Tris, 100 mM NaCl, pH 8.0, and concentrated in VivaSpin concentrators.

Escherichia coli BL21pLysS cells were transformed with the plasmids encoding the VHHs, which were expressed in the cytoplasm after overnight induction with 0.5 mM isopropyl β-d-1-thiogalactopyranoside (IPTG) at 16° C. The cultures were centrifuged, the bacterial pellets were resuspended in 40 mL of lysis buffer (20 mM Tris-HCl, 200 mM NaCl, 20 mM imidazole, pH 8.0) containing complete protease inhibitor cocktail (Roche) and they were frozen at -80 °C until used. On the purification day, the resuspended pellets were thawed, sonicated (15 minutes, 9s on-pulse, 5s off-pulse), centrifuged and loaded onto a HisTrap column. Bound proteins were eluted with a linear gradient of buffer B (20 mM Tris-HCl, 200 mM NaCl, 500 mM imidazole, pH 8.0) and analysed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Fractions with higher purity were pooled, concentrated and further purified by SEC on a Superdex 75 16/600 column (Cytiva) pre-equilibrated in 10 mM Tris-HCl, 100 mM NaCl, pH 8.0.

The purity of the final protein samples was analyzed by SDS-PAGE followed by Coomasie Blue staining.

Purification of complexes used for crystallization screenings.

The RBD-SD1 construct was incubated with $TMPRSS2^{S441A}$ and A01 at final concentrations of 47.4 μM, 71.1 μM and 107 μM, respectively. After over-night incubation at 4 °C, the reaction was loaded onto a Superdex 200 10/300 column (Cytiva) equilibrated in 10 mM Tris-HCl, 100 mM NaCl (pH 8.0) to isolate the complex by SEC. Eluted fractions were analyzed by SDS-PAGE and those corresponding to the ternary complex were pooled, concentrated to 8.5 mg/mL and used in crystallization trials.

$TMPRSS2^{S441A}$ was incubated with A07 at final concentrations of 68 μM and 102 μM, respectively, over-night at 4 °C. Then, the mix was loaded onto a Superdex 200 10/300 column (Cytiva) equilibrated in 10 mM Tris-HCl, 100 mM NaCl (pH 8.0) and eluted fractions were analyzed by SDS-PAGE. Fractions of the binary complex were pooled, concentrated to 6.1 mg/mL and used in crystallization trials.

Protein expression and purification for biophysical assays

RBD and Spike ectodomain-encoding plasmids were transiently transfected into Expi293F™ cells (Thermo-Fischer) using FectoPro® DNA transfection reagent (PolyPlus). After 5 days at 37 °C, cells were

harvested by centrifugation and proteins from the supernatants were purified using a HisTrap-Excel column (Cytiva). Eluted fractions were pooled, concentrated and injected onto a Superdex 200 10/300 column (Cytiva) equilibrated in 10 mM Tris-HCl, 100 mM NaCl (pH 8.0) to perform size-exclusion chromatography. Fractions from the main peak were concentrated and frozen. The HKU1B spike protein used for flow cytometry was biotinylated using an EZ-Link Sulfo-NHS-Biotinylation Kit (Thermo Fisher Scientific).

TMPRSS2 (WT) was expressed from a stable S2 cell line as indicated before, and it was purified by affinity (the tag and glycans were not removed).

The purity of the final protein samples was analyzed by SDS-PAGE followed by Coomasie Blue staining.

Cleavage of TMPRSS2^{S441A} for biophysical assays

To prepare the first batch of cleaved TMPRSS2^{S441A}, 600 µg were incubated at room temperature with 1 µg of TMPRSS2 WT in 600 µL of buffer (10 mM Tris-HCl, 100 mM NaCl, pH 8.0) for 21 hours. The final concentration of TMPRSS2^{S441A} in the reaction was 21 µM. Then, the mix was stored at 4 °C for 8 hours, aliquoted, flash-frozen in liquid nitrogen, and stored at -80 °C until used.

The second batch of cleaved TMPRSS2^{S441A} was prepared by incubating at room temperature 230 µg with 0.3 µg of TMPRSS2 WT in 200 µL of buffer (10 mM Tris-HCl, 100 mM NaCl, pH 8.0) for 7 hours. The final concentration of TMPRSS2^{S441A} in the reaction was 25 µM. Then, the mix was aliquoted, flash-frozen in liquid nitrogen, and stored at -80 °C until used.

Crystallization and structural determination

The RBD-SD1/TMPRSS2^{S441A}/A01 complex crystallized in 0.35 M NaH_2PO_4, 0.65 M K_2HPO_4 at 4 °C using the sitting-drop vapor diffusion method. The TMPRSS2^{S441A}/A07 complex crystallized in 20 %w/v polyethylene glycol (PEG) 3350, 0.05 M 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) (pH 7.0), 1 %w/v Tryptone, 0.001 %w/v NaN_3 (crystal form 1) and in 10 %w/v PEG 3000, 0.1 M imidazole (pH 8.0), 0.2 M lithium sulfate (crystal form 2) at 18 °C using the sitting-drop vapor diffusion method. Crystals were flash-frozen by immersion into a cryo-protectant containing the crystallization solution supplemented with 33% (v/v) glycerol, followed by rapid transfer into liquid nitrogen.

The X-ray diffraction data of both complexes were collected at the SOLEIL synchrotron source (Saint Aubin, France). Collections were carried out at 100 K at the Proxima-1 beamline [46].

Data were processed, scaled and reduced with XDS [47,48] and AIMLESS [49]. The structures were determined by molecular replacement using Phaser from the PHENIX suite [50] with search ensembles obtained from AlphaFold2 (HKU1B-RBD-SD1 and A07) or from previously deposited structures (7MEQ for TMPRSS2^{S441A}, 7KN5 for A01). The final models were built by combining real space model building in Coot [51] with reciprocal space refinement with phenix.refine.

The RBD/ TMPRSS2^{S441A}/A01 complex crystallized in hexagonal $P6_1$ space group and diffracted to 3.5 Å resolution with two ternary complexes in the asymmetric unit. Refinement was carried out using

constraints provided by non-crystallographic symmetry (NCS) and model targets (5KWB for the RBD, our high-resolution structure of TMPRSS2^{S441A} in complex with A07 for the TMPRSS2^{S441A} protein, and 7KN5 and 7VOA for the nanobody A01) allowing us to build a model with good geometry (Table 1).

The TMPRSS2^{S441A}/A07 complex provided two crystal forms belonging to the C2 orthorhombic space group. One form diffracted to 1.8 Å with one complex in the asymmetric unit, and the other diffracted to 2.4 Å with 2 complexes in the asymmetric unit (Table 1). From the high-resolution crystal we built a model of cleaved TMPRSS2^{S441A}, while the second crystal form allowed us to build the uncleaved form. Nevertheless, many strong positive peaks remained in the difference electron density map of the high-resolution data, particularly near the active site. These densities were easily explained by superposing the uncleaved TMPRSS2^{S441A} structure, clearly showing the presence of an alternative conformation on two loops in this crystal (residues 431-440 and 462-467, linked by the disulfide bond C437-C465), which refined with final occupancies of 0.55 and 0.45 (cleaved and uncleaved forms, respectively). In addition, the cleavage released a new N-terminus (residues 256-258) that became ordered and inserted deeply into the core of the protease to interact with the side chain of the buried D440, as it is observed in this family of proteases upon activation.

The final models were validated with Molprobity [52]. The analyses of the macromolecular surfaces were carried out in PDBePISA [53]. Figures were created using Pymol [54] and BioRender.com.

Biolayer interferometry (BLI)

Affinity of recombinant RBDs towards the purified ectodomain of TMPRSS2^{S441A} was assessed in real-time using a bio-layer interferometry Octet-R8 device (Sartorius). Nickel-nitriloacetic acid (Ni-NTA) capture sensors (Sartorius) were loaded for 10 min at 1,000 rpm shaking speed with the different RBDs at 400 nM in phosphate-buffered saline (PBS). The sensors were then blocked with PBS containing bovine serum albumin (BSA) at 1.0 mg/mL (assay buffer) and were incubated at 1,000 rpm with two-fold serially diluted concentrations (800 nM to 25 nM) of TMPRSS2^{S441A} in assay buffer. Association and dissociation were monitored for 300 s and 240 s, respectively. Measurements for a reference were recorded using a sensor loaded with an unrelated protein (CD147) that was dipped at each analyte concentration. A sample reference measurement was recorded from a sensor loaded with each RBD and dipped in the assay buffer. Specific signals were calculated by double referencing, substracting nonspecific signals obtained for the sensor and sample references from the signals recorded for the RBD-loaded sensors dipped in TMPRSS2^{S441A} solutions. Three independent experiments were performed but only the curves obtained at 800 nM in the first experiment were chosen for preparing Figure 2.

Affinity of the recombinant RBDs towards cleaved TMPRSS2^{S441A} (first batch) was determined following a similar protocol, although the range of ligand concentrations assayed went from 240 nM to 7.5 nM. Association and dissociation were monitored for 240 s and 180 s, respectively. Specific signals were calculated by subtracting the nonspecific signal of the sample reference from the signals recorded for the RBD-loaded sensors. Association and dissociation profiles were fitted assuming a 1:1 binding model. Three

independent experiments were performed (Figure S6) and the dissociation constant (Kd) values from each of them were averaged and used to calculate the standard deviation.

Sequence alignment

Multiple sequence alignments were performed using Clustal Omega [55]. The sequence logo was created with WebLogo [56] (https://weblogo.berkeley.edu/).

GFP-split fusion assay

Cell–cell fusion assays were performed as previously described [41]. Briefly, 293T cells stably expressing GFP1-10 and GFP11 were co-cultured at a 1:1 ratio (6×10^4 cells/well) and transfected in suspension with Lipofectamine 2000 (Thermo) in a 96-well plate (uClear, #655090) (20 ng of spike plasmid, 20ng of TMPRSS2 plasmids adjusted to 100 ng DNA with pQCXIP-Empty). At 20 h post-transfection, images covering 90% of the well surface, were acquired per well on an Opera Phenix High-Content Screening System (PerkinElmer). The GFP area was quantified on Harmony High-Content Imaging and Analysis Software.

Pseudovirus generation and infection

Pseudoviruses were produced by transfection of 293T cells as previously described [1]. Briefly, cells were cotransfected with plasmids encoding for lentiviral proteins, a luciferase reporter, and the HKU1 spike plasmid. Pseudotyped virions were harvested 2 and 3 days after transfection. Production efficacy was assessed by measuring infectivity or HIV Gag p24 concentration using the commercial ELISA Lenti-X p24 Rapid titer kit (Takara). For infection assays, 293T cells (6×10^4) were transfected in suspension with Lipofectamine 2000 (Thermo) in a 96-well white plate (20ng of TMPRSS2 plasmids adjusted to 100 ng DNA with pQCXIP-Empty). 24 h post-transfection, cells were passed in 2 wells, and infected with indicated amount of virus (5-10 ng of p24) in 100 µL. The next day, 100 µL of media was added. 48 h post-infection, 125 µL of media was carefully removed, and 75 µL of Bright-Glo™ lysis buffer (ProMega) was added. After 10 min, luminescence was acquired using the EnSpire (PerkinElmer).

Flow cytometry

For spike binding, 293T cells were transiently transfected with TMPRSS2 and incubated with Camostat (10 µM) for 2 h. The cells were incubated with soluble biotinylated spike diluted in magnetic-activated cell sorting (MACS) buffer (PBS, 5 g/L BSA, 2 mM ethylenediaminetetraacetic acid (EDTA) at 2 µg/mL) for 30 min at 4°C. The cells were then washed twice with PBS and then incubated with Alexa Fluor 647-conjugated streptavidin (Thermo Fisher Scientific, S21374, 1:400 in MACS buffer) for 30 min at 4°C.

For the spike, transfection efficiency was measured at the surface of live cells using mAb10 diluted in MACS buffer for 30 min at 4°C, and Alexa Fluor 647 anti-human IgG (Thermo Scientific, 1:500 in MACS

buffer). mAb10 is an antibody generated from a SARS-CoV-2 infected patient which cross-reacts with HKU1 [57].

Surface expression of TMPRSS2 was assessed on live cells by staining with anti-TMPRSS2 A01-Fc [1] at 1 µg/mL, for 30 min at 4°C in MACS buffer, followed by staining with Alexa Fluor 647-conjugated Goat anti-Human antibody (Thermo Fisher Scientific, A-21445, 1:500 in MACS buffer).

All cells were washed twice with PBS and fixed with 4% paraformaldehyde. The results were acquired using an Attune Nxt Flow Cytometer (Life Technologies, software v3.2.1). Gating strategies are described in Figure S7.

TMPRSS2-myc expression was assessed on fixed cells by staining intracellularly with anti-cMyc 9E10 (Thermo - M4439, 1:400), for 30 min at RT in PBS/BSA 1%/NaN_3 0.05% with saponin 0.05% followed by Alexa Fluor 647 Goat anti-mouse antibody (Thermo Fisher Scientific, A-21242, 1:500 in PBS/BSA 1%/NaN_3 with saponin 0.05%).

In vitro TMPRSS2 enzymatic activity

For enzymatic assays black 96 well plates were used. Soluble TMPRSS2 WT (5 nM final concentration) was mixed with the indicated RBDs at different concentrations in a buffer containing 50 mM Tris-HCl pH 8.0, 150 mM NaCl and 0.01% Tween-20. The plate was incubated 15 minutes at 37 °C. The substrate Boc-QAR-AMC (t-Butyloxycarbonyl-Glutamine-Alanine-Arginine-7-Amino-4-methylcoumarin) was added (100 µM final concentration, 100 µL final volume). Fluorescence was read every three minutes for 3 hours (excitation wavelength: 380 nm; emission wavelength: 460 nm) at 37°C, using the EnSpire (PerkinElmer). The initial kinetic slope was taken over the 12 first minutes.

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analysis were performed using GraphPad Prism. Details of tests and number of replicates can be found in the figure legends. We defined statistical significance as $p<0.05$.

KEY RESOURCES TABLE

SUPPLEMENTARY INFORMATION

SUPPLEMENTAL TABLE

Table S1 (Related to Figure 1). List of residues involved in hydrogen bonds and salt bridges.

REFERENCES

1. Saunders, N., Fernandez, I., Planchais, C., Michel, V., Rajah, M.M., Baquero Salazar, E., Postal, J., Porrot, F., Guivel-Benhassine, F., Blanc, C., et al. (2023). TMPRSS2 is a functional receptor for human coronavirus HKU1. Nature *624*, 207-214. 10.1038/s41586-023-06761-7.
2. Ko, C.J., Huang, C.C., Lin, H.Y., Juan, C.P., Lan, S.W., Shyu, H.Y., Wu, S.R., Hsiao, P.W., Huang, H.P., Shun, C.T., and Lee, M.S. (2015). Androgen-Induced TMPRSS2 Activates Matriptase and Promotes Extracellular Matrix Degradation, Prostate Cancer Cell Invasion, Tumor Growth, and Metastasis. Cancer Res *75*, 2949-2960. 10.1158/0008-5472.CAN-14-3297.
3. Mukai, S., Yorita, K., Kawagoe, Y., Katayama, Y., Nakahara, K., Kamibeppu, T., Sugie, S., Tukino, H., Kamoto, T., and Kataoka, H. (2015). Matriptase and MET are prominently expressed at the site of bone metastasis in renal cell carcinoma: immunohistochemical analysis. Hum Cell *28*, 44-50. 10.1007/s13577-014-0101-3.
4. Tomlins, S.A., Rhodes, D.R., Perner, S., Dhanasekaran, S.M., Mehra, R., Sun, X.W., Varambally, S., Cao, X., Tchinda, J., Kuefer, R., et al. (2005). Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science *310*, 644-648. 10.1126/science.1117679.
5. Wang, Z., Wang, Y., Zhang, J., Hu, Q., Zhi, F., Zhang, S., Mao, D., Zhang, Y., and Liang, H. (2017). Significance of the TMPRSS2:ERG gene fusion in prostate cancer. Mol Med Rep *16*, 5450-5458. 10.3892/mmr.2017.7281.
6. Lucas, J.M., Heinlein, C., Kim, T., Hernandez, S.A., Malik, M.S., True, L.D., Morrissey, C., Corey, E., Montgomery, B., Mostaghel, E., et al. (2014). 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. 10.1158/2159-8290.CD-13-1010.
7. Woo, P.C., Lau, S.K., Chu, C.M., Chan, K.H., Tsoi, H.W., Huang, Y., Wong, B.H., Poon, R.W., Cai, J.J., Luk, W.K., et al. (2005). Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. J Virol *79*, 884-895. 10.1128/JVI.79.2.884-895.2005.
8. Kahn, J.S., and McIntosh, K. (2005). History and recent advances in coronavirus discovery. Pediatr Infect Dis J *24*, S223-227, discussion S226. 10.1097/01.inf.0000188166.17324.60.
9. Tortorici, M.A., and Veesler, D. (2019). Structural insights into coronavirus entry. Adv Virus Res *105*, 93-116. 10.1016/bs.aivir.2019.08.002.
10. Huang, X., Dong, W., Milewska, A., Golda, A., Qi, Y., Zhu, Q.K., Marasco, W.A., Baric, R.S., Sims, A.C., Pyrc, K., et al. (2015). Human Coronavirus HKU1 Spike Protein Uses O-Acetylated Sialic Acid as an Attachment Receptor Determinant and Employs Hemagglutinin-Esterase Protein as a Receptor-Destroying Enzyme. J Virol *89*, 7202-7213. 10.1128/JVI.00854-15.
11. Hulswit, R.J.G., Lang, Y., Bakkers, M.J.G., Li, W., Li, Z., Schouten, A., Ophorst, B., van Kuppeveld, F.J.M., Boons, G.J., Bosch, B.J., et al. (2019). Human coronaviruses OC43 and HKU1 bind to 9-O-acetylated sialic acids via a conserved receptor-binding site in spike protein domain A. Proc Natl Acad Sci U S A *116*, 2681-2690. 10.1073/pnas.1809667116.

12. Li, Z., Lang, Y., Liu, L., Bunyatov, M.I., Sarmiento, A.I., de Groot, R.J., and Boons, G.J. (2021). Synthetic O-acetylated sialosides facilitate functional receptor identification for human respiratory viruses. Nat Chem *13*, 496-503. 10.1038/s41557-021-00655-9.
13. Pronker, M.F., Creutznacher, R., Drulyte, I., Hulswit, R.J.G., Li, Z., van Kuppeveld, F.J.M., Snijder, J., Lang, Y., Bosch, B.J., Boons, G.J., et al. (2023). Sialoglycan binding triggers spike opening in a human coronavirus. Nature *624*, 201-206. 10.1038/s41586-023-06599-z.
14. Benton, D.J., Wrobel, A.G., Xu, P., Roustan, C., Martin, S.R., Rosenthal, P.B., Skehel, J.J., and Gamblin, S.J. (2020). Receptor binding and priming of the spike protein of SARS-CoV-2 for membrane fusion. Nature *588*, 327-330. 10.1038/s41586-020-2772-0.
15. Walls, A.C., Xiong, X., Park, Y.J., Tortorici, M.A., Snijder, J., Quispe, J., Cameroni, E., Gopal, R., Dai, M., Lanzavecchia, A., et al. (2020). Unexpected Receptor Functional Mimicry Elucidates Activation of Coronavirus Fusion. Cell *183*, 1732. 10.1016/j.cell.2020.11.031.
16. Hurdiss, D.L., Drulyte, I., Lang, Y., Shamorkina, T.M., Pronker, M.F., van Kuppeveld, F.J.M., Snijder, J., and de Groot, R.J. (2020). Cryo-EM structure of coronavirus-HKU1 haemagglutinin esterase reveals architectural changes arising from prolonged circulation in humans. Nat Commun *11*, 4646. 10.1038/s41467-020-18440-6.
17. de Groot, R.J. (2006). Structure, function and evolution of the hemagglutinin-esterase proteins of corona- and toroviruses. Glycoconj J *23*, 59-72. 10.1007/s10719-006-5438-8.
18. Chen, Y.W., Lee, M.S., Lucht, A., Chou, F.P., Huang, W., Havighurst, T.C., Kim, K., Wang, J.K., Antalis, T.M., Johnson, M.D., and Lin, C.Y. (2010). 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. 10.2353/ajpath.2010.090665.
19. Afar, D.E., Vivanco, I., Hubert, R.S., Kuo, J., Chen, E., Saffran, D.C., Raitano, A.B., and Jakobovits, A. (2001). Catalytic cleavage of the androgen-regulated TMPRSS2 protease results in its secretion by prostate and prostate cancer epithelia. Cancer Res *61*, 1686-1692.
20. Tseng, C.C., Jia, B., Barndt, R., Gu, Y., Chen, C.Y., Tseng, I.C., Su, S.F., Wang, J.K., Johnson, M.D., and Lin, C.Y. (2017). Matriptase shedding is closely coupled with matriptase zymogen activation and requires de novo proteolytic cleavage likely involving its own activity. PLoS One *12*, e0183507. 10.1371/journal.pone.0183507.
21. Fraser, B.J., Beldar, S., Seitova, A., Hutchinson, A., Mannar, D., Li, Y., Kwon, D., Tan, R., Wilson, R.P., Leopold, K., et al. (2022). Structure and activity of human TMPRSS2 protease implicated in SARS-CoV-2 activation. Nat Chem Biol *18*, 963-971. 10.1038/s41589-022-01059-7.
22. Bertram, S., Dijkman, R., Habjan, M., Heurich, A., Gierer, S., Glowacka, I., Welsch, K., Winkler, M., Schneider, H., Hofmann-Winkler, H., et al. (2013). TMPRSS2 activates the human coronavirus 229E for cathepsin-independent host cell entry and is expressed in viral target cells in the respiratory epithelium. J Virol *87*, 6150-6160. 10.1128/JVI.03372-12.
23. Glowacka, I., Bertram, S., Muller, M.A., Allen, P., Soilleux, E., Pfefferle, S., Steffen, I., Tsegaye, T.S., He, Y., Gnirss, K., et al. (2011). 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. 10.1128/JVI.02232-10.

24. Hoffmann, M., Kleine-Weber, H., Schroeder, S., Kruger, N., Herrler, T., Erichsen, S., Schiergens, T.S., Herrler, G., Wu, N.H., Nitsche, A., et al. (2020). SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell *181*, 271-280 e278. 10.1016/j.cell.2020.02.052.
25. Limburg, H., Harbig, A., Bestle, D., Stein, D.A., Moulton, H.M., Jaeger, J., Janga, H., Hardes, K., Koepke, J., Schulte, L., et al. (2019). 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*. 10.1128/JVI.00649-19.
26. Shirato, K., Kawase, M., and Matsuyama, S. (2013). Middle East respiratory syndrome coronavirus infection mediated by the transmembrane serine protease TMPRSS2. J Virol *87*, 12552-12561. 10.1128/JVI.01890-13.
27. Perona, J.J., and Craik, C.S. (1995). Structural basis of substrate specificity in the serine proteases. Protein Sci *4*, 337-360. 10.1002/pro.5560040301.
28. Goettig, P., Brandstetter, H., and Magdolen, V. (2019). Surface loops of trypsin-like serine proteases as determinants of function. Biochimie *166*, 52-76. 10.1016/j.biochi.2019.09.004.
29. Perona, J.J., and Craik, C.S. (1997). Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold. J Biol Chem *272*, 29987-29990. 10.1074/jbc.272.48.29987.
30. Lawrence, M.C., and Colman, P.M. (1993). Shape complementarity at protein/protein interfaces. J Mol Biol *234*, 946-950. 10.1006/jmbi.1993.1648.
31. Kirchdoerfer, R.N., Cottrell, C.A., Wang, N., Pallesen, J., Yassine, H.M., Turner, H.L., Corbett, K.S., Graham, B.S., McLellan, J.S., and Ward, A.B. (2016). Pre-fusion structure of a human coronavirus spike protein. Nature *531*, 118-121. 10.1038/nature17200.
32. Yeager, C.L., Ashmun, R.A., Williams, R.K., Cardellichio, C.B., Shapiro, L.H., Look, A.T., and Holmes, K.V. (1992). Human aminopeptidase N is a receptor for human coronavirus 229E. Nature *357*, 420-422. 10.1038/357420a0.
33. Pasternak, A., Ringe, D., and Hedstrom, L. (1999). Comparison of anionic and cationic trypsinogens: the anionic activation domain is more flexible in solution and differs in its mode of BPTI binding in the crystal structure. Protein Sci *8*, 253-258. 10.1110/ps.8.1.253.
34. Freer, S.T., Kraut, J., Robertus, J.D., Wright, H.T., and Xuong, N.H. (1970). Chymotrypsinogen: 2.5-angstrom crystal structure, comparison with alpha-chymotrypsin, and implications for zymogen activation. Biochemistry *9*, 1997-2009. 10.1021/bi00811a022.
35. Madison, E.L., Kobe, A., Gething, M.J., Sambrook, J.F., and Goldsmith, E.J. (1993). Converting tissue plasminogen activator to a zymogen: a regulatory triad of Asp-His-Ser. Science *262*, 419-421. 10.1126/science.8211162.
36. Herter, S., Piper, D.E., Aaron, W., Gabriele, T., Cutler, G., Cao, P., Bhatt, A.S., Choe, Y., Craik, C.S., Walker, N., et al. (2005). 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 *390*, 125-136. 10.1042/BJ20041955.

37. Ohno, A., Maita, N., Tabata, T., Nagano, H., Arita, K., Ariyoshi, M., Uchida, T., Nakao, R., Ulla, A., Sugiura, K., et al. (2021). Crystal structure of inhibitor-bound human MSPL that can activate high pathogenic avian influenza. Life Sci Alliance *4*. 10.26508/lsa.202000849.
38. McCallum, M., Park, Y.J., Stewart, C., Sprouse, K.R., Brown, J., Tortorici, M.A., Gibson, C., Wong, E., Ieven, M., Telenti, A., and Veesler, D. (2024). Human coronavirus HKU1 recognition of the TMPRSS2 host receptor. bioRxiv. 10.1101/2024.01.09.574565.
39. Ke, Z., Oton, J., Qu, K., Cortese, M., Zila, V., McKeane, L., Nakane, T., Zivanov, J., Neufeldt, C.J., Cerikan, B., et al. (2020). Structures and distributions of SARS-CoV-2 spike proteins on intact virions. Nature *588*, 498-502. 10.1038/s41586-020-2665-2.
40. Kerr, M.A., Walsh, K.A., and Neurath, H. (1976). A proposal for the mechanism of chymotrypsinogen activation. Biochemistry *15*, 5566-5570. 10.1021/bi00670a022.
41. Buchrieser, J., Dufloo, J., Hubert, M., Monel, B., Planas, D., Rajah, M.M., Planchais, C., Porrot, F., Guivel-Benhassine, F., Van der Werf, S., et al. (2020). Syncytia formation by SARS-CoV-2-infected cells. EMBO J *39*, e106267. 10.15252/embj.2020106267.
42. Kodaka, M., Yang, Z., Nakagawa, K., Maruyama, J., Xu, X., Sarkar, A., Ichimura, A., Nasu, Y., Ozawa, T., Iwasa, H., et al. (2015). A new cell-based assay to evaluate myogenesis in mouse myoblast C2C12 cells. Exp Cell Res *336*, 171-181. 10.1016/j.yexcr.2015.06.015.
43. Edie, S., Zaghloul, N.A., Leitch, C.C., Klinedinst, D.K., Lebron, J., Thole, J.F., McCallion, A.S., Katsanis, N., and Reeves, R.H. (2018). Survey of Human Chromosome 21 Gene Expression Effects on Early Development in Danio rerio. G3 (Bethesda) *8*, 2215-2223. 10.1534/g3.118.200144.
44. Hoffmann, M., Hofmann-Winkler, H., Smith, J.C., Kruger, N., Arora, P., Sorensen, L.K., Sogaard, O.S., Hasselstrom, J.B., Winkler, M., Hempel, T., et al. (2021). Camostat mesylate inhibits SARS-CoV-2 activation by TMPRSS2-related proteases and its metabolite GBPA exerts antiviral activity. EBioMedicine *65*, 103255. 10.1016/j.ebiom.2021.103255.
45. Iwaki, T., Figuera, M., Ploplis, V.A., and Castellino, F.J. (2003). Rapid selection of Drosophila S2 cells with the puromycin resistance gene. Biotechniques *35*, 482-484, 486. 10.2144/03353bm08.
46. Chavas, L.M.G., Gourhant, P., Guimaraes, B.G., Isabet, T., Legrand, P., Lener, R., Montaville, P., Sirigu, S., and Thompson, A. (2021). PROXIMA-1 beamline for macromolecular crystallography measurements at Synchrotron SOLEIL. J Synchrotron Radiat *28*, 970-976. 10.1107/S1600577521002605.
47. Kabsch, W. (2010). Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr D Biol Crystallogr *66*, 133-144. 10.1107/S0907444909047374.
48. Kabsch, W. (2010). Xds. Acta Crystallogr D Biol Crystallogr *66*, 125-132. 10.1107/S0907444909047337.
49. Evans, P.R., and Murshudov, G.N. (2013). How good are my data and what is the resolution? Acta Crystallogr D Biol Crystallogr *69*, 1204-1214. 10.1107/S0907444913000061.
50. Liebschner, D., Afonine, P.V., Baker, M.L., Bunkoczi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.W., Jain, S., McCoy, A.J., et al. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr D Struct Biol *75*, 861-877. 10.1107/S2059798319011471.

51. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr *66*, 486-501. 10.1107/S0907444910007493.
52. Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant, M.G., Videau, L.L., Deis, L.N., Verma, V., Keedy, D.A., Hintze, B.J., Chen, V.B., et al. (2018). MolProbity: More and better reference data for improved all-atom structure validation. Protein Sci *27*, 293-315. 10.1002/pro.3330.
53. Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. J Mol Biol *372*, 774-797. 10.1016/j.jmb.2007.05.022.
54. DeLano, W.L. (2002). The PyMOL Molecular Graphics System. . DeLano Scientific, San Carlos, CA, USA, 2002.
55. Madeira, F., Pearce, M., Tivey, A.R.N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A., and Lopez, R. (2022). Search and sequence analysis tools services from EMBL-EBI in 2022. Nucleic Acids Res *50*, W276-W279. 10.1093/nar/gkac240.
56. Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: a sequence logo generator. Genome Res *14*, 1188-1190. 10.1101/gr.849004.
57. Planchais, C., Fernandez, I., Bruel, T., de Melo, G.D., Prot, M., Beretta, M., Guardado-Calvo, P., Dufloo, J., Molinos-Albert, L.M., Backovic, M., et al. (2022). Potent human broadly SARS-CoV-2-neutralizing IgA and IgG antibodies effective against Omicron BA.1 and BA.2. J Exp Med *219*. 10.1084/jem.20220638.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
mAb10	Planchais C, et al. Ref 57	N/A
anti-TMPRSS2 A01-Fc	Saunders N, et al. Ref 1	N/A
Alexa Fluor 647-conjugated Goat anti-Human antibody	Thermo Fisher Scientific	Cat#A-21445
anti-cMyc 9E10	Thermo Fisher Scientific	Cat#M4439
Alexa Fluor 647-conjugated Goat anti-Human antibody	Thermo Fisher Scientific	Cat#A-21242
Bacterial and virus strains		
Escherichia coli BL21(DE3)pLysS competent cells	Promega	Cat# L1195
Biological samples		
Chemicals, peptides, and recombinant proteins		
Puromycin	InvivoGen	Cat#ant-pr-1
Penicillin-streptomycin	Gibco	Cat#15140122
Serum-free insect cell medium	HyClone	Cat#SH30913.02
Enterokinase light chain	New England Biolabs	Cat#P8070L
Endoglycosidase D	New England Biolabs	Cat#P0742L
Endoglycosidase H	New England Biolabs	Cat#P0702L
Isopropyl β-d-1-thiogalactopyranoside (IPTG)	EUROMEDEX	Cat# EU0008-B
Complete protease inhibitor cocktail	Roche	Cat#11873580001
Effectene transfection reagent	Qiagen	Cat#301427
FectoPro® DNA transfection reagent	Polyplus	Cat#101000007
Lipofectamine 2000	Thermo Fisher Scientific	Cat#11668019
Camostat	Sigma-Aldrich	Cat#SML0057
Alexa Fluor 647-conjugated streptavidin	Thermo Fisher Scientific	Cat#S21374
t-Butyloxycarbonyl-Glutamine-Alanine-Arginine-7-Amino-4-methylcoumarin (Boc-QAR-AMC)	Biotechne	Ca#ES014
Recombinant protein: HKU1B RBD	This study	N/A
Recombinant protein: HKU1B RBD K487A	This study	N/A
Recombinant protein: HKU1B RBD D507A	This study	N/A
Recombinant protein: HKU1B RBD L510R	This study	N/A
Recombinant protein: HKU1B RBD W515A	Saunders N, Ref 1	N/A
Recombinant protein: HKU1B RBD T527A	This study	N/A
Recombinant protein: HKU1B RBD Y528A	This study	N/A
Recombinant protein: HKU1B RBD S529A	This study	N/A
Recombinant protein: SARS-CoV-2 RBD	Saunders N, Ref 1	N/A
Recombinant protein: HKU1B RBD-SD1	This study	N/A
Recombinant protein: HKU1B Spike ectodomain	This study	N/A

Recombinant protein: Human TMPRSS2 ectodomain	This study	N/A
Recombinant protein: Human TMPRSS2^{S441A} ectodomain	This study	N/A
Recombinant protein: Nanobody A01	This study	N/A
Recombinant protein: Nanobody A07	This study	N/A
HIV Gag p24 concentration kit	Saunders N, Ref 1	N/A
Critical commercial assays		
Bright-Glo™ lysis buffer	Promega	Cat#E2620
Q5 Site-Directed mutagenesis kit	New England Biolabs	Cat# E0554S
Lenti-X p24 Rapid titer kit	Takara	Cat#632200
EZ-Link Sulfo-NHS-Biotinylation kit	Thermo Fisher Scientific	Cat#21217
Deposited data		
Crystal structure of the HKU1B RBD in complex with TMPRSS2^{S441A} and nanobody A01	This study	8S0M
Crystal structure of TMPRSS2^{S441A} in the zymogen and mature forms complexed to nanobody A07 (Crystal form 1)	This study	8S0L
Crystal structure of TMPRSS2^{S441A} in the zymogen form complexed to nanobody A07 (Crystal form 2)	This study	8S0N
Experimental models: Cell lines		
Drosophila melanogaster S2 cell line	Thermo Fisher Scientific	Cat#R690-07
Expi293™ cells	Thermo Fisher Scientific	Cat#A14527
293T cells	ATCC	CRL-3216
293T GFP-split cells (GFP1-10 and GFP11)	Buchrieser J, et al. Ref 41	N/A
Experimental models: Organisms/strains		
Oligonucleotides		
Recombinant DNA		
HKU1B RBD (residues 330-614) with Hisx8, Strep and Avi-tags cloned in pCAGGs	Saunders N, et al. (ref 1) and this study	N/A
HKU1B RBD (residues 330-614) K487A with Hisx8, Strep and Avi-tags cloned in pCAGGs	This study	N/A
HKU1B RBD (residues 330-614) D507A with Hisx8, Strep and Avi-tags cloned in pCAGGs	This study	N/A
HKU1B RBD (residues 330-614) L510R with Hisx8, Strep and Avi-tags cloned in pCAGGs	This study	N/A
HKU1B RBD (residues 330-614) W515A with Hisx8, Strep and Avi-tags cloned in pCAGGs	Saunders N, et al. (ref 1) and this study	N/A
HKU1B RBD (residues 330-614) T527A with Hisx8, Strep and Avi-tags cloned in pCAGGs	This study	N/A
HKU1B RBD (residues 330-614) Y528A with Hisx8, Strep and Avi-tags cloned in pCAGGs	This study	N/A

HKU1B RBD (residues 330-614) S529A with Hisx8, Strep and Avi-tags cloned in pCAGGs	This study	N/A
SARS-CoV-2 RBD (residues 331-528) with Hisx8, Strep and Avi-tags cloned in pcDNA3.1(+)	Saunders N, et al. Ref 1	N/A
HKU1B RBD-SD1 (residues 307-675) with an enterokinase site and double-Strep tag cloned in a modified pMT/BiP plasmid	This study	N/A
HKU1B stabilized Spike ectodomain (residues residues 14–1276) with a His-tag, cloned in pcDNA3.1(+)	Saunders N, et al. Ref 1	N/A
Human TMPRSS2^{S441A} ectodomain (residues 107-492) with an enterokinase site and double-Strep tag cloned in a modified pMT/BiP plasmid	This study	N/A
Human TMPRSS2 ectodomain (residues 107-492) with an enterokinase site and double-Strep tag cloned in a modified pMT/BiP plasmid	This study	N/A
pCoPURO plasmid	Iwaki et al. (Ref. 45)	Addgene #17533
Nanobody A01 with a His-tag cloned in pET23	Saunders N, et al. (Ref 1)	N/A
Nanobody A07 with a His-tag cloned in pET23	Saunders N, et al. (Ref 1)	N/A
Human TMPRSS2 with an N-terminal cMYC-epitope, cloned in pCCAGS	Hoffmann, M, et al (Ref 44)	N/A
Human TMPRRS2 with an N-terminal cMYC-tag, cloned into a phCMV backbone	This study	N/A
Mouse TMPRRS2 with an N-terminal cMYC-tag, cloned into a phCMV backbone	This study	N/A
Ferret TMPRRS2 with an N-terminal cMYC-tag, cloned into a phCMV backbone	This study	N/A
Hamster TMPRRS2 with an N-terminal cMYC-tag, cloned into a phCMV backbone	This study	N/A
Macaque TMPRRS2 with an N-terminal cMYC-tag, cloned into a phCMV backbone	This study	N/A
pQCXIP-Empty	Buchrieser J, et al. (Ref 41)	N/A
pCSDest-TMPRSS2	Edie S, et al. (Ref 43)	Addgene #53887
Software and algorithms		
XDS	Kabsch, W. (refs 47 and 48)	https://xds.mr.mpg.de
Aimless	Evans, P.R. and Murshudov, G.N. (ref 49)	https://www.ccp4.ac.uk/html/aimless.html
Phenix suite	Liebschner, D, et al. (Ref 50)	https://www.phenix-online.org/
AlphaFold2	Deepmind	https://www.deepmind.com
Coot	Emsley, P., et al (ref 51)	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
Molprobity	Williams, C.J. (Ref 52)	http://molprobity.biochem.duke.edu/
PDBePISA	Krissinel, E., and Henrick, K. (ref 53)	https://www.ebi.ac.uk/pdbe/pisa/
PyMOL	Schrodinger, LLC	https://pymol.org/2

Clustal Omega web server	Madeira, F, et al. (Ref 55)	https://www.ebi.ac.uk/jdispatcher/msa/clustalo
WebLogo web server	Crooks, G.E, et al (Ref 56)	https://weblogo.berkeley.edu
Prism Version 9	GraphPad	https://www.graphpad.com
Other		
Strep-Tactin Superflow high-capacity column	IBA Life Sciences	Cat#2-1238-001
HiLoad™ 16/600 Superdex™ 200 pg	Cytiva	Cat#28989335
HiLoad™ 16/600 Superdex™ 75 pg	Cytiva	Cat#28989333
HisTrap™ HP	Cytiva	Cat#17524801
HisTrap™ Excel	Cytiva	Cat#17371206
Superdex 200 increase 10/300 GL	Cytiva	Cat#28990944
Octet® NTA capture biosensors	Sartorius	Cat#18-5101

Figure 1

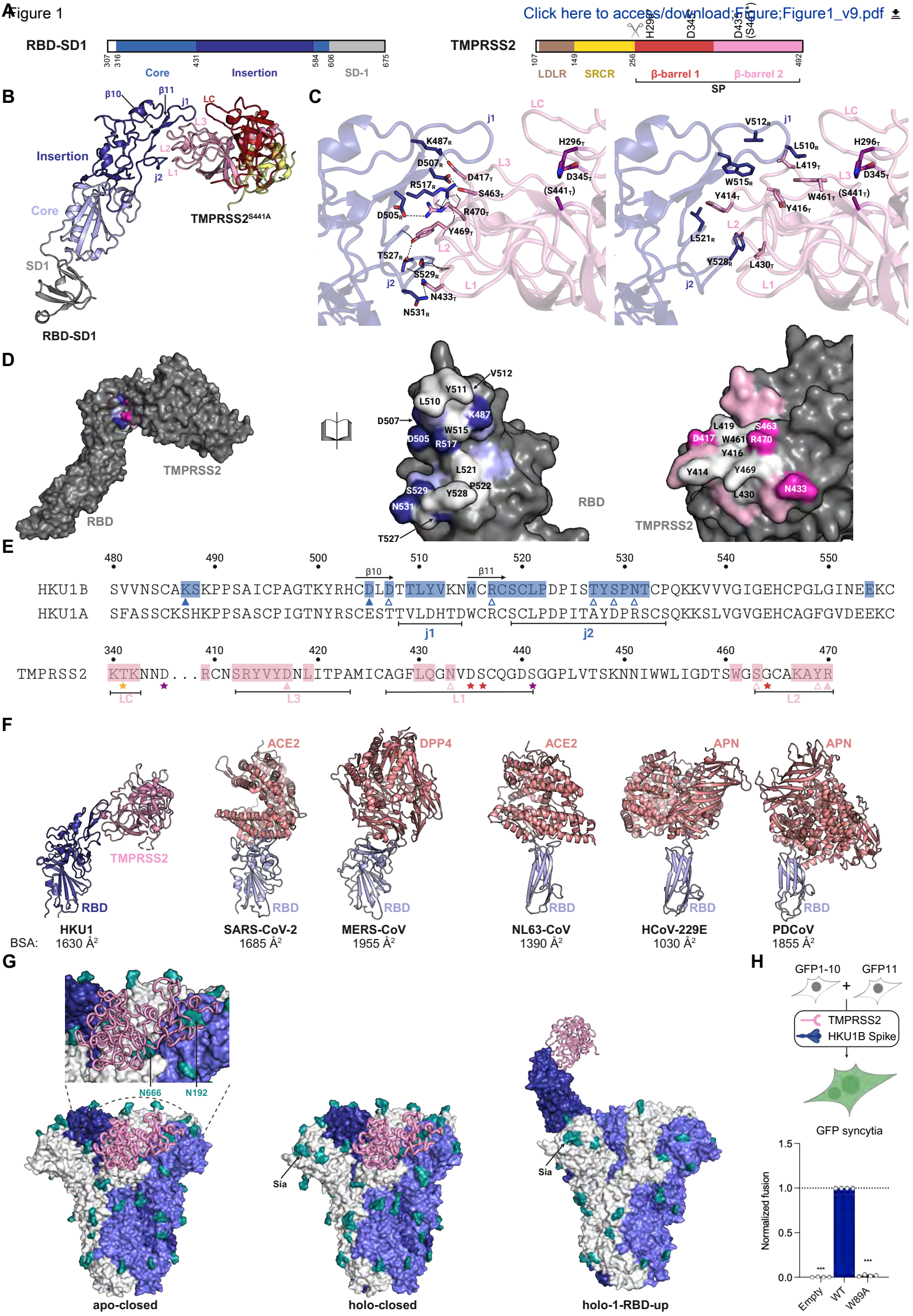


Figure 2

Click here to access/download;Figure;Figure_2_v2.pdf

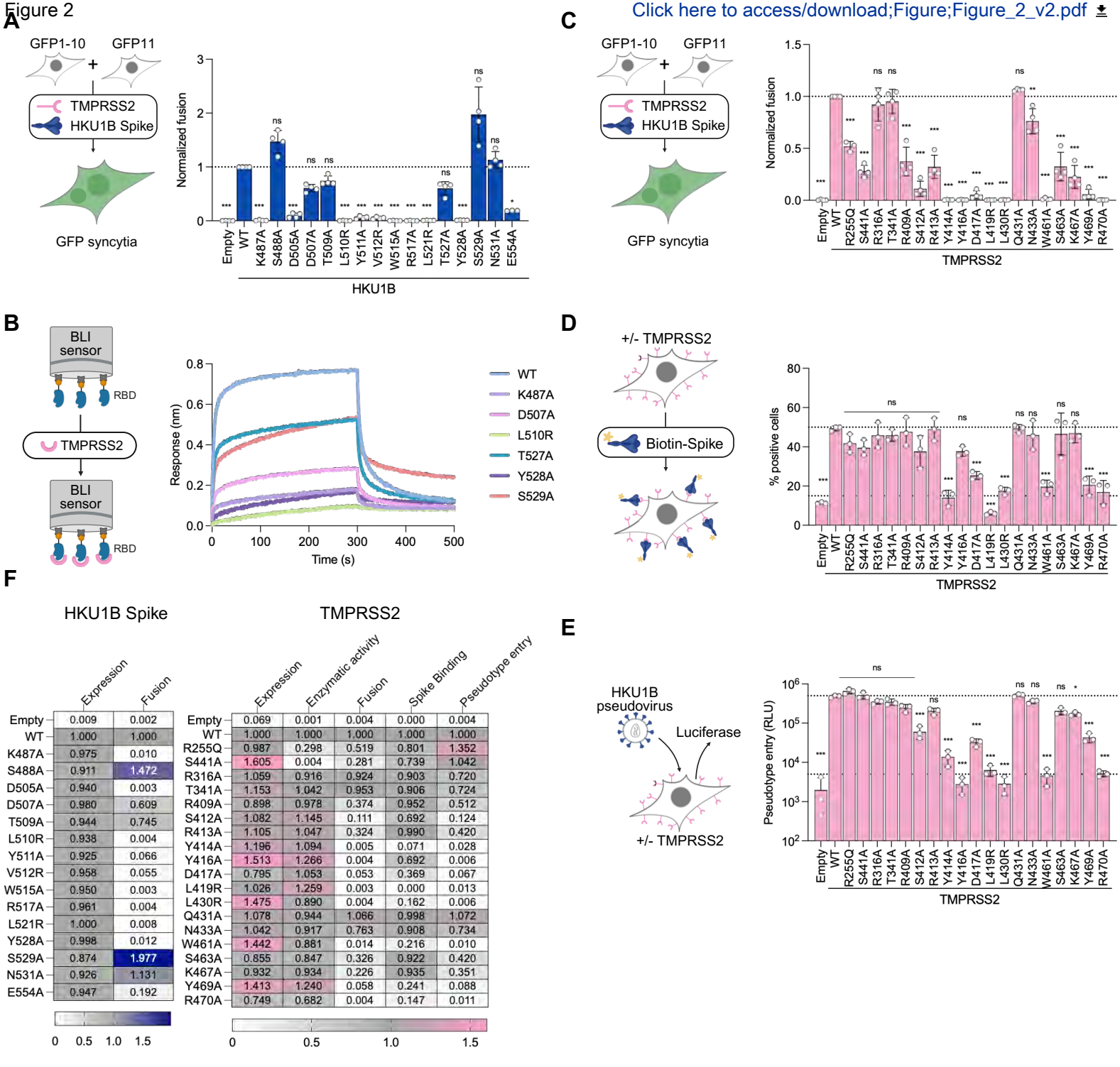


Figure 3

Click here to access/download;Figure;Figure_3.pdf

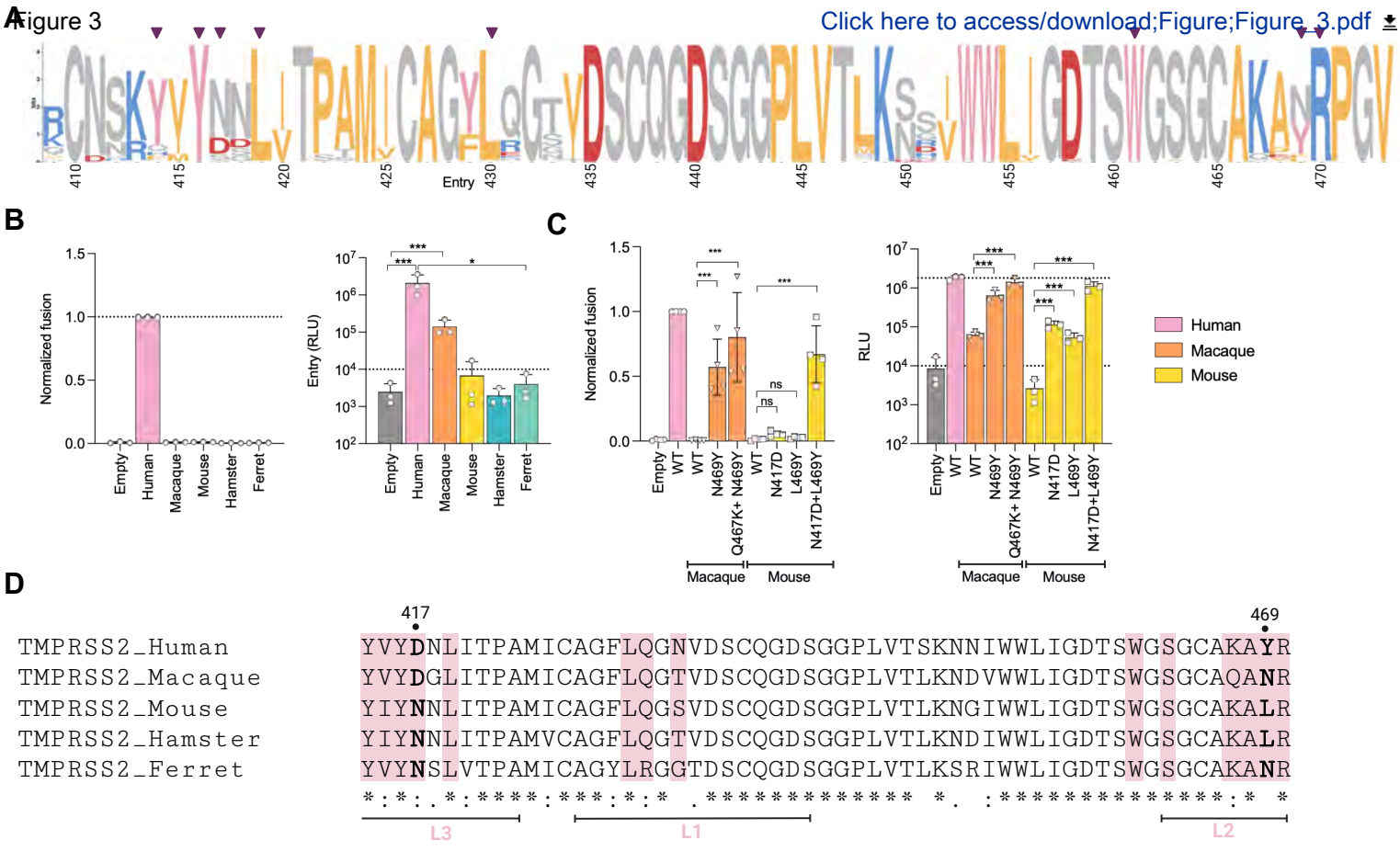


Figure 4

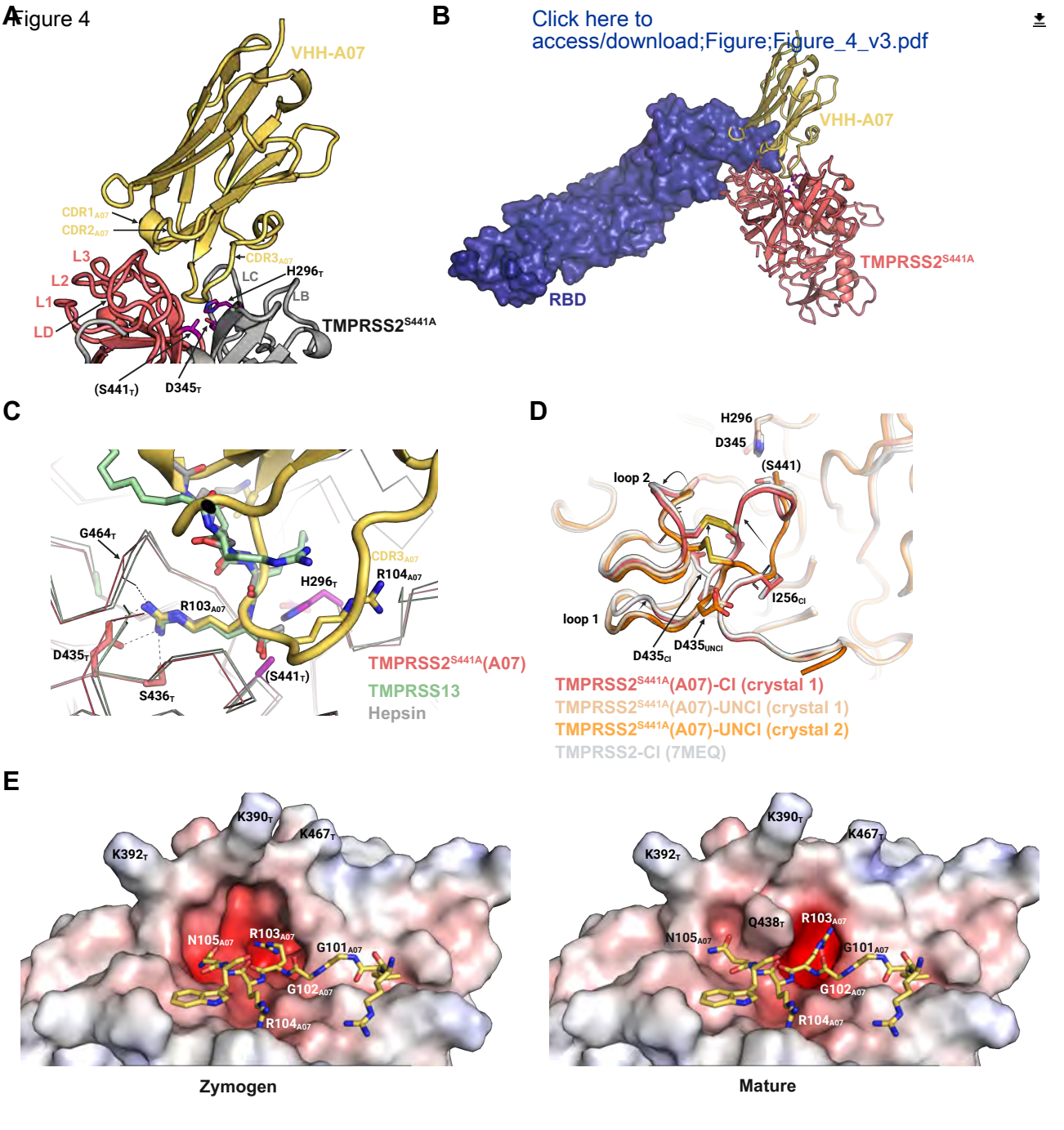
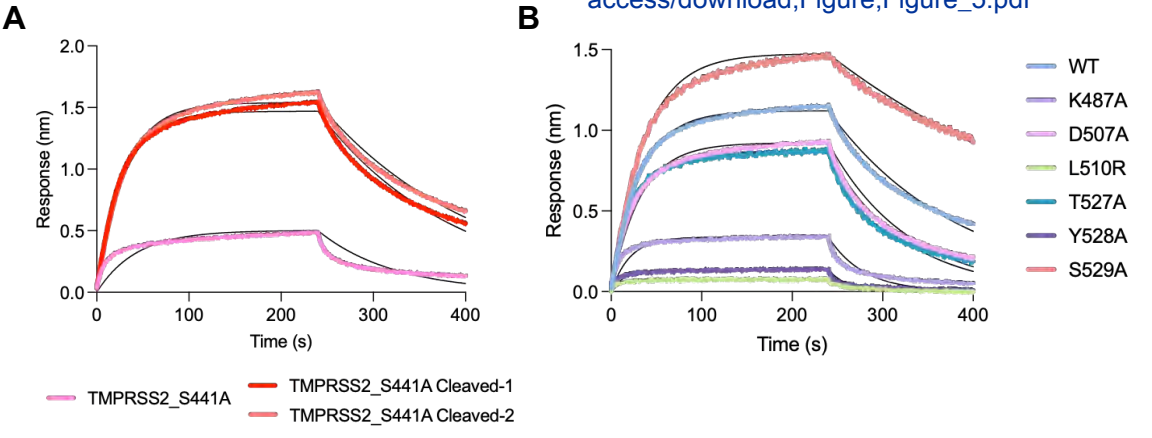

A
VHH-A07
CDR1A07
CDR2A07
CDR3A07
L3
L2
L1
LD
LC
LB
H296T
TMPRSS2S441A
(S441T)
D345T
B
Click here to access/download;Figure;Figure_4_v3.pdf
VHH-A07
TMPRSS2S441A
RBD
C
G464T
CDR3A07
H296T
R104A07
R103A07
D435T
(S441T)
S436T
TMPRSS2S441A(A07)
TMPRSS13
Hepsin
D
H296
D345
(S441)
loop 2
I256CI
loop 1
D435UNCI
D435CI
TMPRSS2S441A(A07)-CI (crystal 1)
TMPRSS2S441A(A07)-UNCI (crystal 1)
TMPRSS2S441A(A07)-UNCI (crystal 2)
TMPRSS2-CI (7MEQ)
E
K390T
K467T
K392T
N105A07
R103A07
G101A07
G102A07
R104A07
Q438T
Zymogen
Mature

Figure 5

Click here to access/download;Figure;Figure_5.pdf



C

Protein	Mean Kd (nM)	SD (nM)
RBD-WT	33.3	2.0
RBD-K487A	146.5	12.2
RBD-D507A	67.4	2.8
RBD-L510R	N.D.	N.D.
RBD-T527A	58.2	3.3
RBD-Y528A	222.1	67.5
RBD-S529A	13.0	1.0

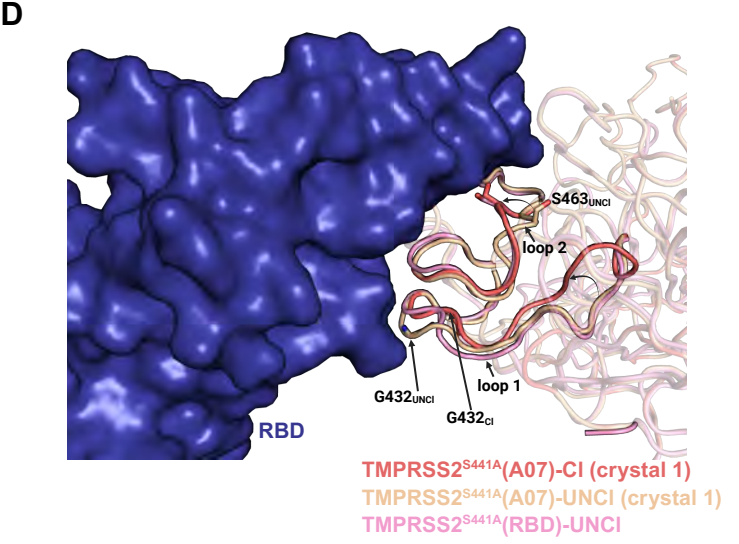


Table SI1 (Related to Figures 1 and 4). List of residues involved in hydrogen bonds and salt bridges.

Donor (D)	Acceptor (A)	D-A distance (Å)
TMPRSS2 S441A (chain A) VHH-A07 (chain B) Crystal 1 – Alternate location (alt loc) A		
A LYS 342 NZ	B VAL 100 O	2.637
A GLN 438 NE2 (alt loc A)	B ARG 104 O	3.407
A GLY 439 N (alt loc A)	B ARG 103 O	3.019
A SER 463 OG (alt loc A)	B HIS 31 ND1	2.358
A ARG 470 NE	**B GLU 30 OE1**	**3.019**
A ARG 470 NE	**B GLU 30 OE2**	**3.381**
A ARG 470 NH2	**B GLU 30 OE2**	**2.776**
B HIS 31 ND1	A SER 463 OG (alt loc A)	2.358
B VAL 100 N	A LYS 340 O	2.928
B ARG 103 N	A SER 460 O	3.270
B ARG 103 NH1 (alt loc A)	**A ASP 435 OD1 (alt loc A)**	**3.138**
B ARG 103 NH2 (alt loc A)	**A ASP 435 OD2 (alt loc A)**	**3.166**
B ARG 103 NH2 (alt loc A)	A GLY 464 O (alt loc A)	2.980
B ARG 104 NH1	A CYS 297 O	3.237
B ASN 105 N	A VAL 280 O	3.138
B ASN 105 ND2 (alt loc A)	A HIS 279 O	2.945
B ASN 115 ND2	A ASP 338 OD1	3.121
B ASN 115 ND2	A SER 339 OG	3.159
TMPRSS2 S441A (chain A) VHH-A07 (chain B) Crystal 1 - Alternate location (alt loc) B		
A LYS 342 NZ	B VAL 100 O	2.637
A SER 463 OG (alt loc B)	B THR 53 OG1	2.656
A ARG 470 NE	**B GLU 30 OE1**	**3.019**
A ARG 470 NE	**B GLU 30 OE2**	**3.381**
A ARG 470 NH2	**B GLU 30 OE2**	**2.776**
B THR 53 OG1	A SER 463 OG (alt loc B)	2.656
B VAL 100 N	A LYS 340 O	2.928
B ARG 103 N	A SER 460 O	3.270
B ARG 104 NH1	A CYS 297 O	3.237
B ASN 105 N	A VAL 280 O	3.138
B ASN 115 ND2	A ASP 338 OD1	3.121
B ASN 115 ND2	A SER 339 OG	3.159
TMPRSS2 S441A (chains A and C) VHH-A07 (chain B and D) Crystal 2		
A LYS 342 NZ	B VAL 100 O	2.372
A SER 463 OG	B THR 53 OG1	2.752
A ARG 470 NE	**B GLU 30 OE2**	**3.212**
A ARG 470 NH2	**B GLU 30 OE2**	**2.447**
B THR 53 OG1	A SER 463 OG	2.752
B VAL 100 N	A LYS 340 O	3.086
B ARG 103 N	A HIS 296 NE2	3.142
B ARG 104 N	A HIS 296 NE2	3.241
B ARG 104 NH1	A CYS 297 O	3.049
B ARG 104 NH1	**A GLU 299 O**	**3.232**
B ARG 104 NH2	A LYS 300 O	3.136
B ASN 105 N	A VAL 280 O	3.443
B ASN 105 ND2	A HIS 279 O	3.471
C LYS 342 NZ	D VAL 100 O	2.641
C SER 463 OG	D THR 53 OG1	2.559
C ARG 470 NE	**D GLU 30 OE1**	**3.332**
C ARG 470 NE	**D GLU 30 OE2**	**3.354**
C ARG 470 NH2	**D GLU 30 OE2**	**2.679**
D THR 53 OG1	C SER 463 OG	2.559

D VAL 100 N	C LYS 340 O	2.951
D ARG 103 N	C HIS 296 NE2	3.079
D ARG 104 NH1	C CYS 297 O	3.089
D ARG 104 NH1	**C GLU 299 O**	**3.551**
D ARG 104 NH2	C LYS 300 O	3.141
D ASN 105 N	C VAL 280 O	3.492
D ASN 105 ND2	C HIS 279 O	3.464
D ASN 115 ND2	C ASP 338 OD1	3.107
D ASN 115 ND2	C SER 339 OG	3.084

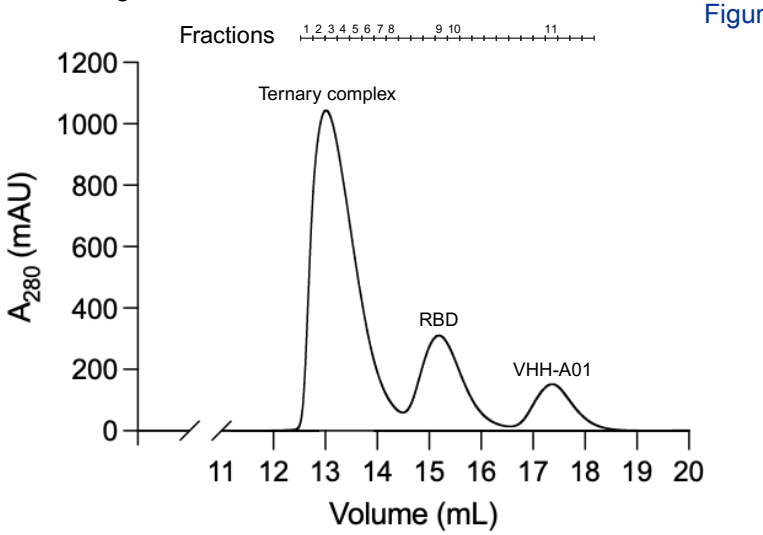
RBD (chains A and D)
TMPRSS2 S441A (chains B and E)
VHH-A07 (chain U and V)

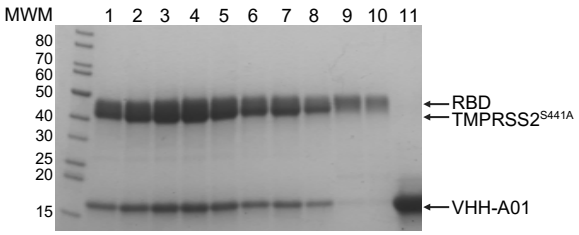
A LYS 487 NZ	**B ASP 417 OD2**	**3.036**
A ARG 517 NH1	B VAL 415 O	2.196
A ARG 517 NH2	B VAL 415 O	2.836
A LEU 521 N	B TYR 469 OH	2.908
A THR 527 OG1	B TYR 469 OH	3.032
A SER 529 OG	B GLN 431 O	2.696
A ASN 531 ND2	B ASN 433 OD1	2.991
B SER 463 OG	A ASP 507 OD1	2.402
B SER 463 OG	A ASP 507 OD2	2.677
B TYR 469 OH	A LEU 521 O	2.765
B TYR 469 OH	A THR 527 OG1	3.032
B ARG 470 NH2	**A ASP 505 OD2**	**3.281**
B ARG 150 NH1	U ASP 30 O	3.068
B TYR 152 OH	U ASP 105 OD2	2.535
B LYS 166 NZ	**U ASP 73 OD1**	**2.330**
B TRP 168 N	U SER 54 O	3.103
B THR 207 N	U ASP 56 OD2	3.146
B GLY 370 N	U ALA 101 O	3.119
B LEU 373 N	U LEU 102 O	2.602
B GLN 408 N	U ASP 116 OD1	3.155
B GLN 408 NE2	U PRO 112 O	3.500
U TYR 31 OH	B ASP 482 OD2	2.725
U TYR 32 OH	B PRO 422 O	2.688
U TYR 117 OH	B ILE 420 O	3.278
D LEU 521 N	E TYR 469 OH	2.840
D THR 527 OG1	E TYR 469 OH	3.096
D SER 529 OG	E GLN 431 O	2.704
D ASN 531 ND2	E ASN 433 OD1	2.893
E SER 463 OG	D ASP 507 OD1	2.493
E SER 463 OG	D ASP 507 OD2	2.727
E TYR 469 OH	D LEU 521 O	2.736
E TYR 469 OH	D THR 527 OG1	3.096
E ARG 470 NH2	**D ASP 505 OD2**	**3.371**
E ARG 150 NH1	V ASP 30 O	3.029
E TYR 152 OH	V ASP 105 OD2	2.895
E LYS 166 NZ	**V ASP 73 OD1**	**3.038**
E TRP 168 N	V SER 54 O	3.167
E SER 206 OG	V ASP 56 OD2	2.832
E LEU 373 N	V LEU 102 O	2.357
E GLN 408 N	V ASP 116 OD1	3.159
E GLN 408 NE2	V PRO 112 O	3.491
V TYR 31 OH	E ASP 482 OD2	2.693
V TYR 32 OH	E PRO 422 O	2.798
V TYR 117 OH	E ILE 420 O	3.342

Supplemental Figure 1

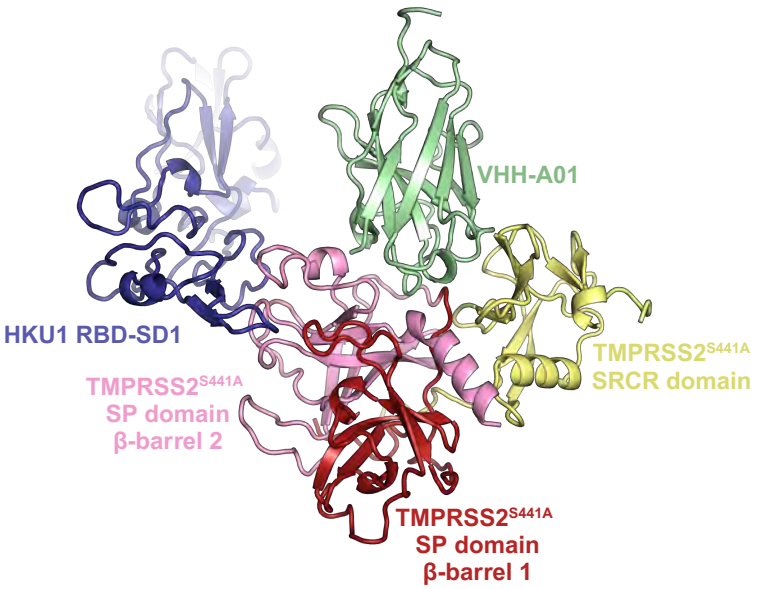
Click here to access/download;Supplemental Figure;Figure_SI_1.pdf

A

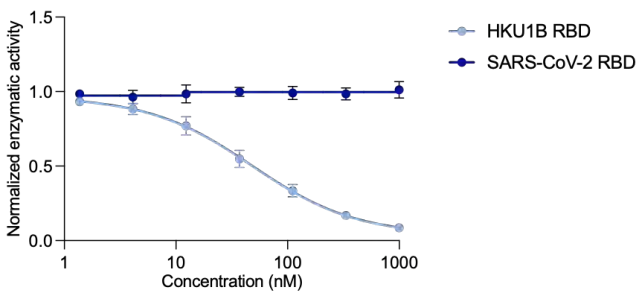




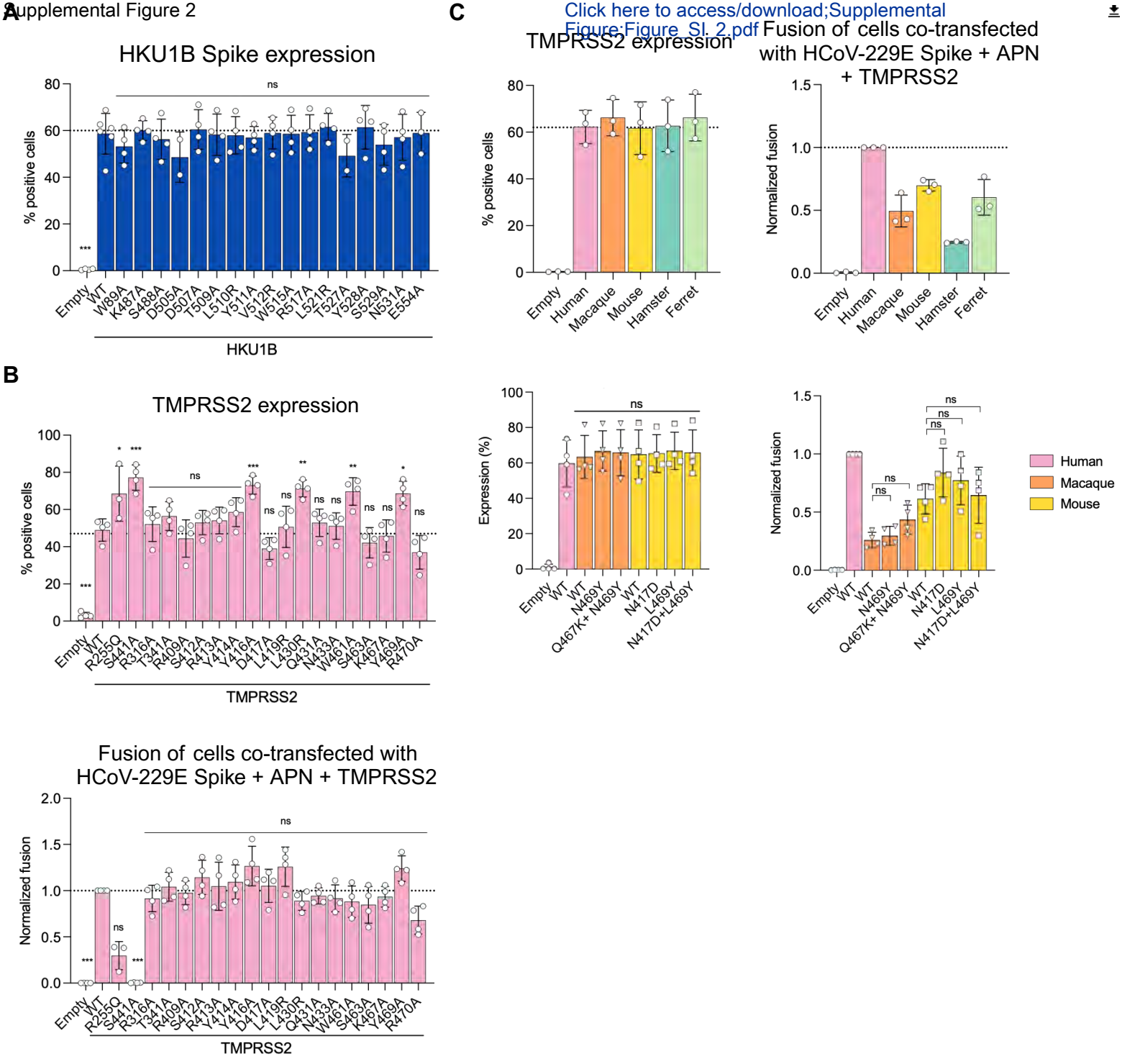
B



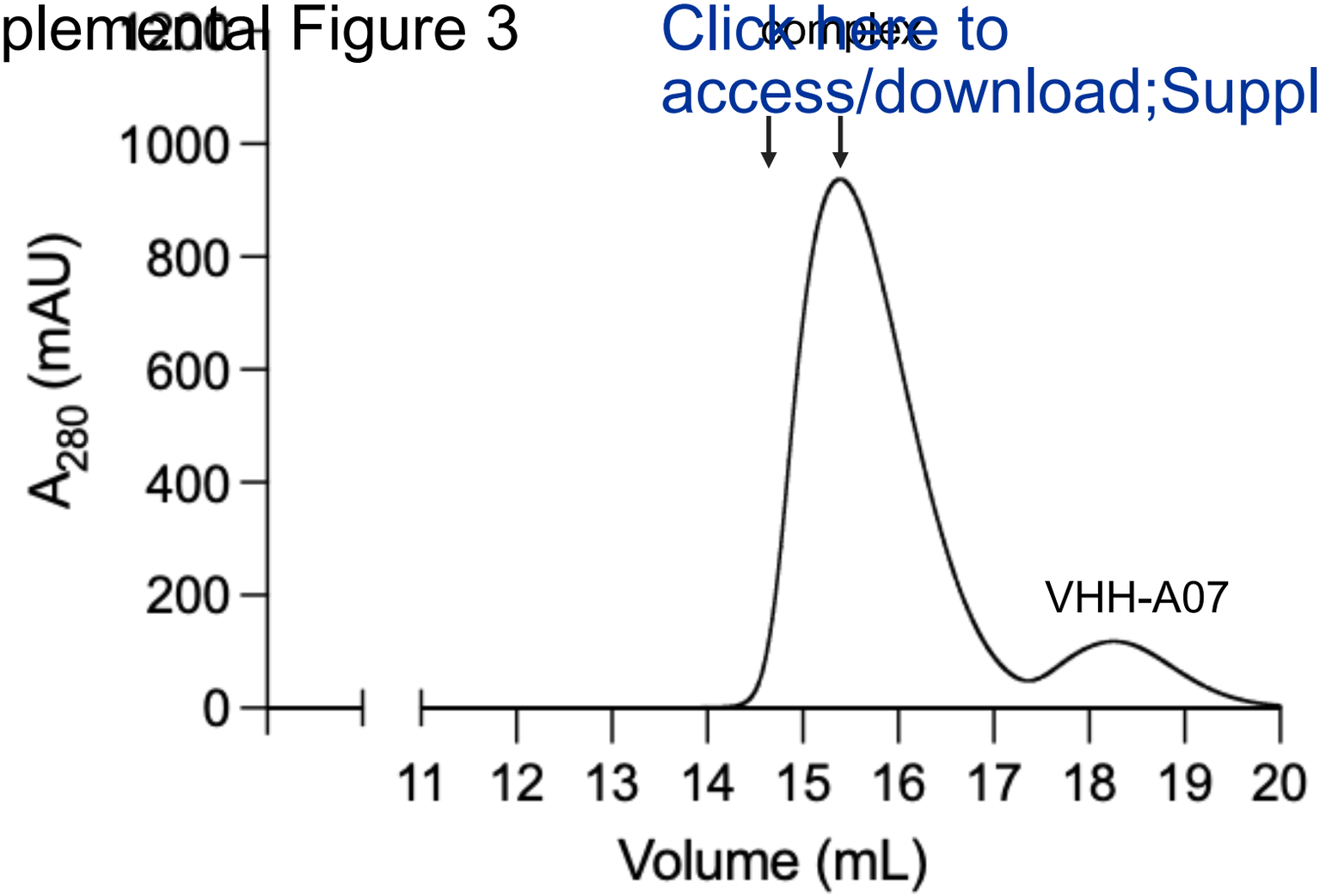
C

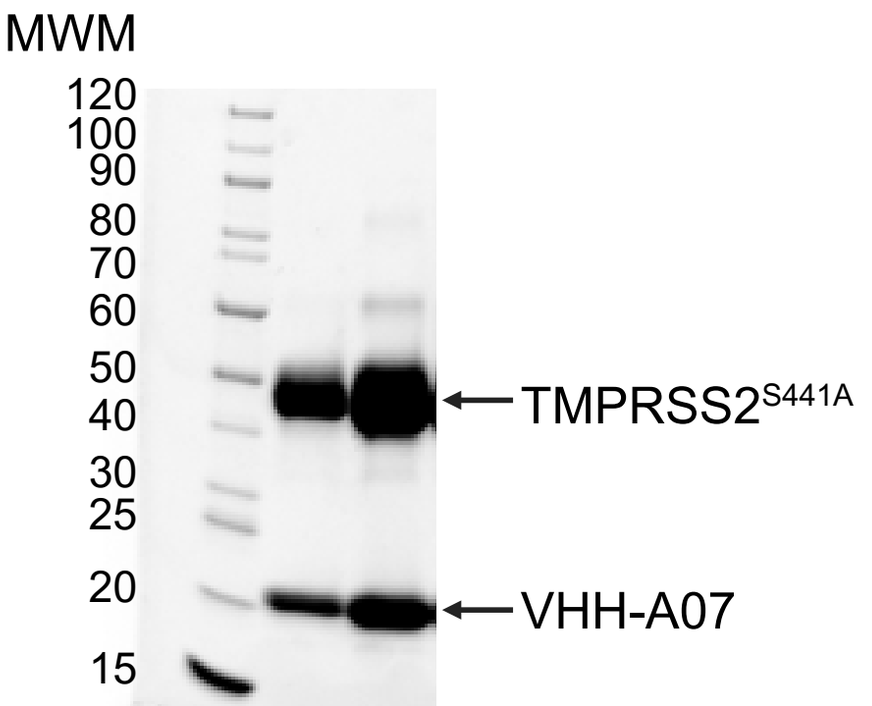


Supplemental Figure 2



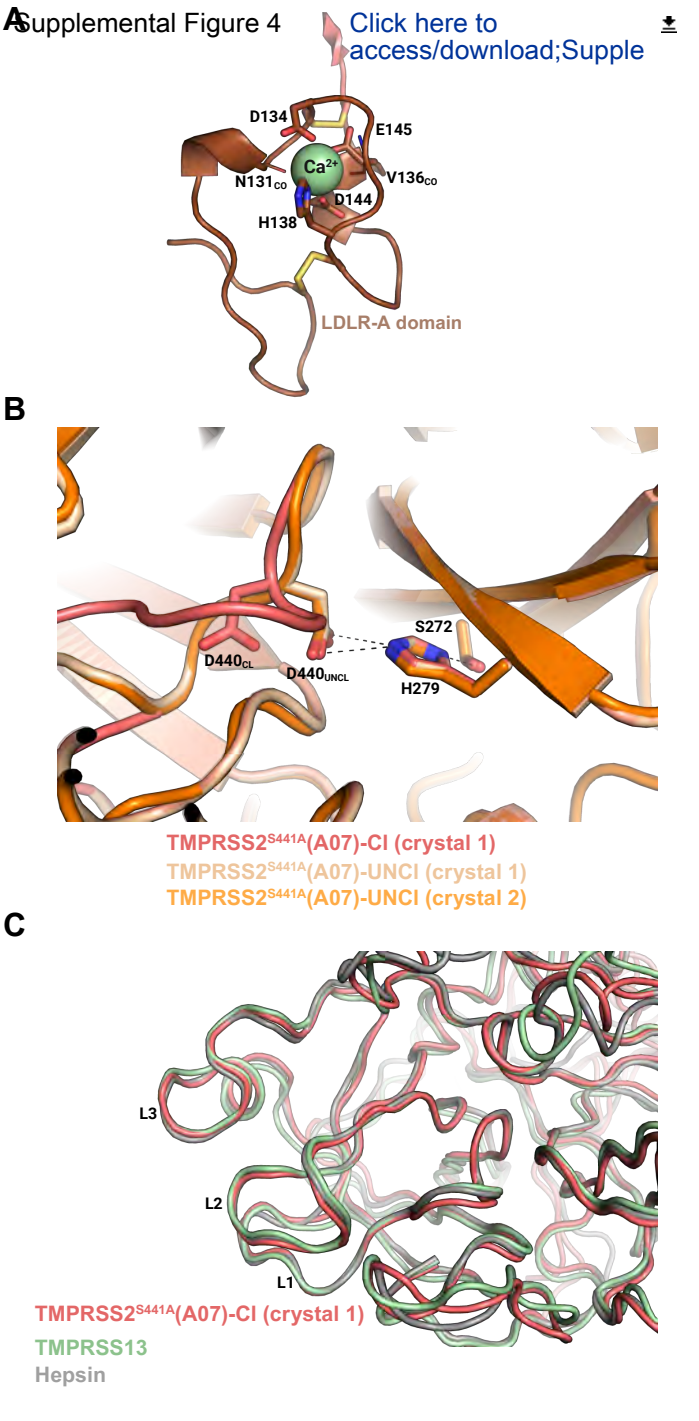
Supplemental Figure 3



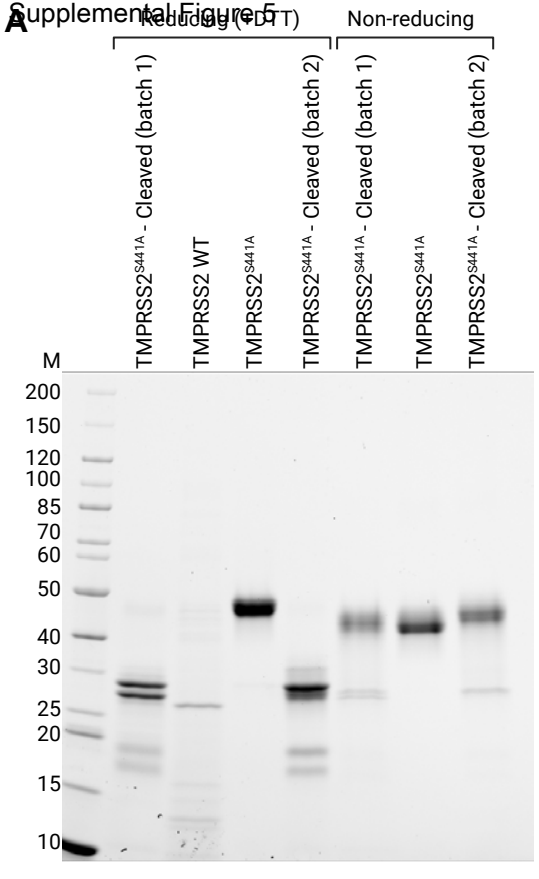


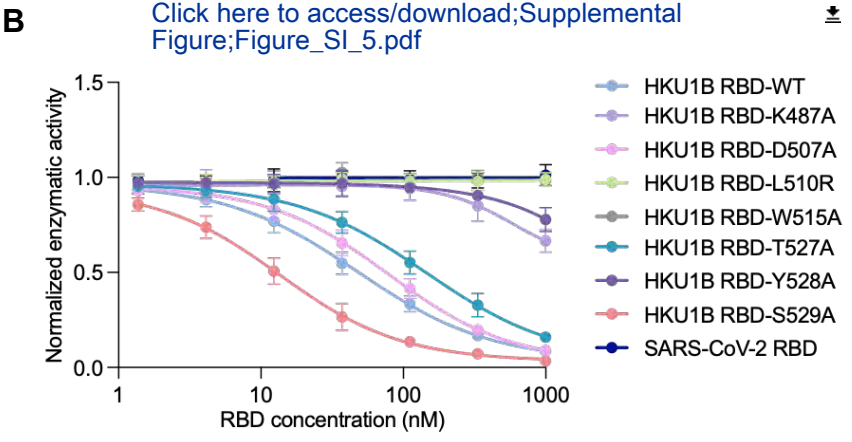
Supplemental Figure 4

Click here to access/download;Supple

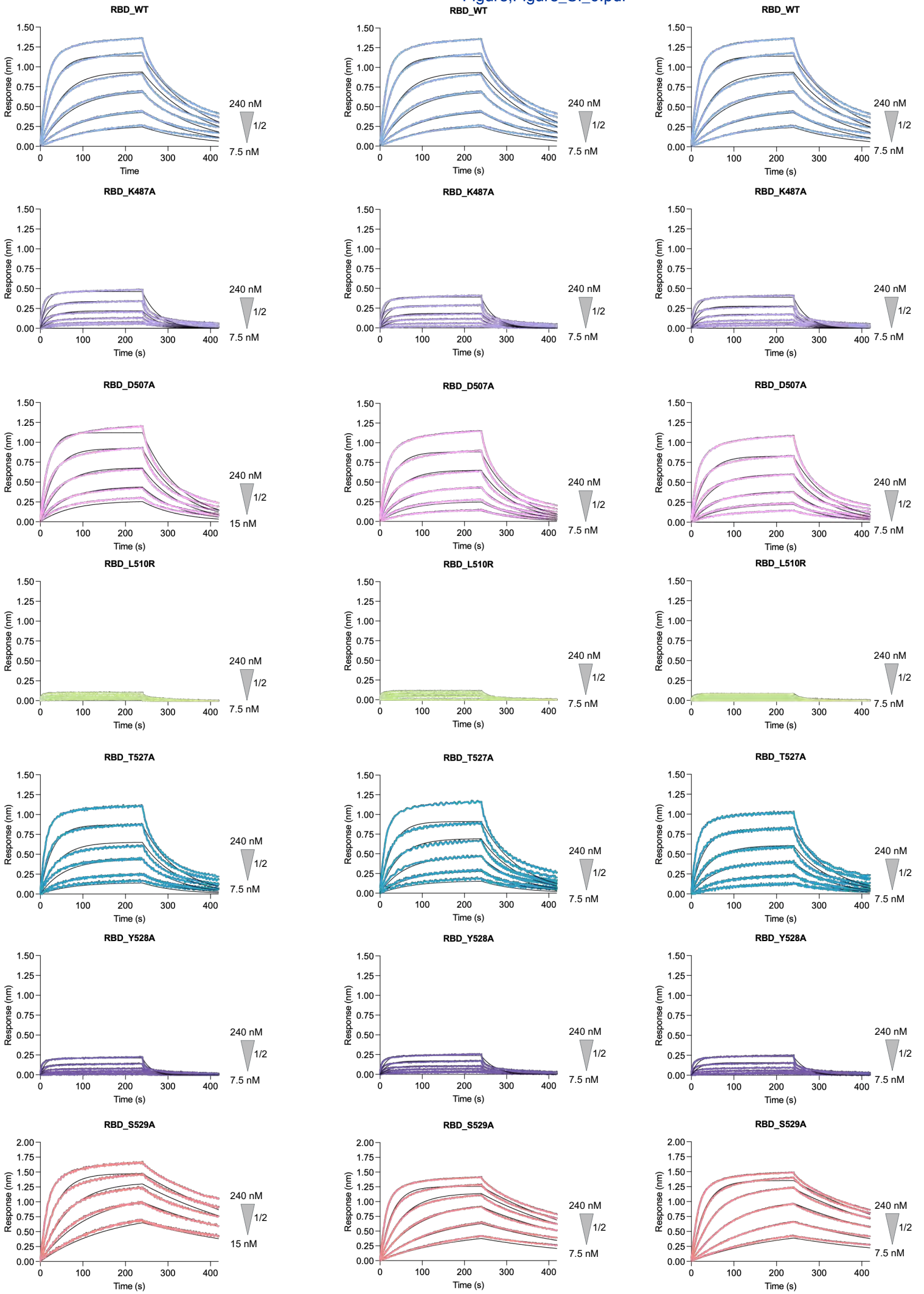


Supplemental Figure 5





Supplemental Figure 6



Supplemental Figure 7

Click here to access/download;Supplemental Figure;Figure_SI_7.tif