

# Human coronavirus HKU1 recognition of the TMPRSS2 host receptor

Matthew McCallum<sup>1,&</sup>, Young-Jun Park<sup>1,2,&</sup>, Cameron Stewart<sup>1</sup>, Kaitlin R. Sprouse, Jack Brown<sup>1</sup>, M. Alejandra Tortorici<sup>1</sup>, Cecily Gibson<sup>1,2</sup>, Emily Wong<sup>3</sup>, Margareta Ieven<sup>4</sup>, Amalio Telenti<sup>3</sup>, David Veesler<sup>1,2,\*</sup>

<sup>1</sup>Department of Biochemistry, University of Washington, Seattle, Washington, USA

<sup>2</sup>Howard Hughes Medical Institute, Seattle, WA 98195, USA

<sup>3</sup>Vir Biotechnology, San Francisco, CA 94158, USA

<sup>4</sup>Laboratory of Clinical Microbiology, Vaccine & Infectious Disease Institute, University of Antwerp, Antwerp, Belgium

<sup>&</sup>These authors contributed equally

\*Correspondence: [dveesler@uw.edu](mailto:dveesler@uw.edu)

**The human coronavirus HKU1 spike (S) glycoprotein engages host cell surface sialoglycans and transmembrane protease serine 2 (TMPRSS2) to initiate infection. The molecular basis of HKU1 binding to TMPRSS2 and determinants of host receptor tropism remain elusive. Here, we designed an active human TMPRSS2 construct enabling high-yield recombinant production in human cells of this key therapeutic target. We determined a cryo-electron microscopy structure of the HKU1 RBD bound to human TMPRSS2 providing a blueprint of the interactions supporting viral entry and explaining the specificity for TMPRSS2 among human type 2 transmembrane serine proteases. We found that human, rat, hamster and camel TMPRSS2 promote HKU1 S-mediated entry into cells and identified key residues governing host receptor usage. Our data show that serum antibodies targeting the HKU1 RBD TMPRSS2 binding-site are key for neutralization and that HKU1 uses conformational masking and glycan shielding to balance immune evasion and receptor engagement.**

HKU1 is one of the four  $\beta$ -coronaviruses known to circulate in humans seasonally. It was initially discovered in 2005 in nasopharyngeal aspirates obtained from a 71-year-old man with pneumonia in Hong Kong<sup>1</sup>. Retrospective investigations identified HKU1 in nasopharyngeal aspirates collected between 1995 and 2002 from symptomatic individuals on several continents<sup>1–4</sup> and more HKU1 strains were sequenced and isolated thereafter<sup>5,6</sup>. Although HKU1 is largely considered to be a ‘common cold’ coronavirus, mainly causing mild respiratory infections, it also causes severe illness, particularly (but not only) for children, the elderly and immunocompromised individuals<sup>7–9</sup>. Furthermore, its high seroprevalence in multiple independent studies indicates that HKU1 continuously circulates in the human population and is therefore endemic<sup>8</sup>.

Coronavirus infections are initiated by interaction of the trimeric spike (S) glycoprotein with a host cell receptor leading to fusion of the viral and host membrane and genome delivery<sup>10</sup>. To date, only a finite number of proteinaceous receptors have been identified as promoting coronavirus infection and all of them are host transmembrane proteases: angiotensin-converting

enzyme 2 (ACE2) for SARS-CoV-1<sup>11</sup>, SARS-CoV-2<sup>12–14</sup>, NL63<sup>15</sup> and a few divergent bat-borne merbecoviruses<sup>16</sup>; aminopeptidase N (APN) for 229E<sup>17</sup>, CCoV-HuPn-2018<sup>18</sup>, PDCoV<sup>19</sup> and TGEV/PEDV<sup>20</sup>; dipeptidyl-peptidase 4 (DPP4) for MERS-CoV<sup>21</sup> and HKU4-related merbecoviruses<sup>22–24</sup>; and carcinoembryonic antigen-related cell adhesion molecule 1a for MHV A59<sup>25</sup>. Furthermore, OC43 and HKU1 recognize 9-O-acetylated sialosides<sup>26–28</sup> present at the host cell surface and a recent study showed that transmembrane protease serine 2 (TMPRSS2) is an entry receptor for HKU1<sup>29</sup>, which is likely engaged upon S conformational changes triggered by  $\alpha$ 2,8-linked 9-O-acetylated disialosides binding<sup>30</sup>. The molecular basis of HKU1 recognition of the human TMPRSS2 receptor along with the determinants of host receptor tropism, however, remain elusive.

TMPRSS2 is part of the type 2 transmembrane serine protease family that encompasses several cell surface-expressed trypsin-like serine proteases participating in epithelial homeostasis and TMPRSS2 genomic rearrangements are common in prostate cancer<sup>31</sup>. Furthermore, TMPRSS2 has been shown to proteolytically activate the S glycoprotein of many coronaviruses, including SARS-CoV-2<sup>32,33</sup>, SARS-CoV-1<sup>34,35</sup>, MERS-CoV<sup>36,37</sup>, 229E<sup>38</sup>, as well as influenza virus hemagglutinins<sup>39,40</sup>. As a result, TMPRSS2 inhibitors have been proposed as candidates for preventing or treating these viral infections<sup>41–43</sup>. Recombinant production of TMPRSS2 and related type 2 transmembrane serine proteases has proven exceptionally challenging, hindering studies of these important therapeutic targets.

Here, we designed a human TMPRSS2 glycoprotein construct enabling high-yield recombinant production of the active mature enzyme ectodomain in human cells, which will expedite biochemical studies and drug screening globally. We determined a cryo-electron microscopy (cryoEM) structure of the human TMPRSS2 ectodomain bound to the HKU1 S receptor-binding domain (RBD), providing a blueprint of the interactions mediating receptor engagement, specificity and host receptor tropism. Analysis of HKU1 infection-elicited serum antibodies in humans reveal that the TMPRSS2-binding site is a key site of vulnerability for HKU1 that becomes exposed upon sialoglycan-mediated RBD opening. Our findings reveal how this human coronavirus gains access to target host cells and balances receptor recognition and immune evasion, informing vaccine design.

### **Design of an active human TMPRSS2 construct**

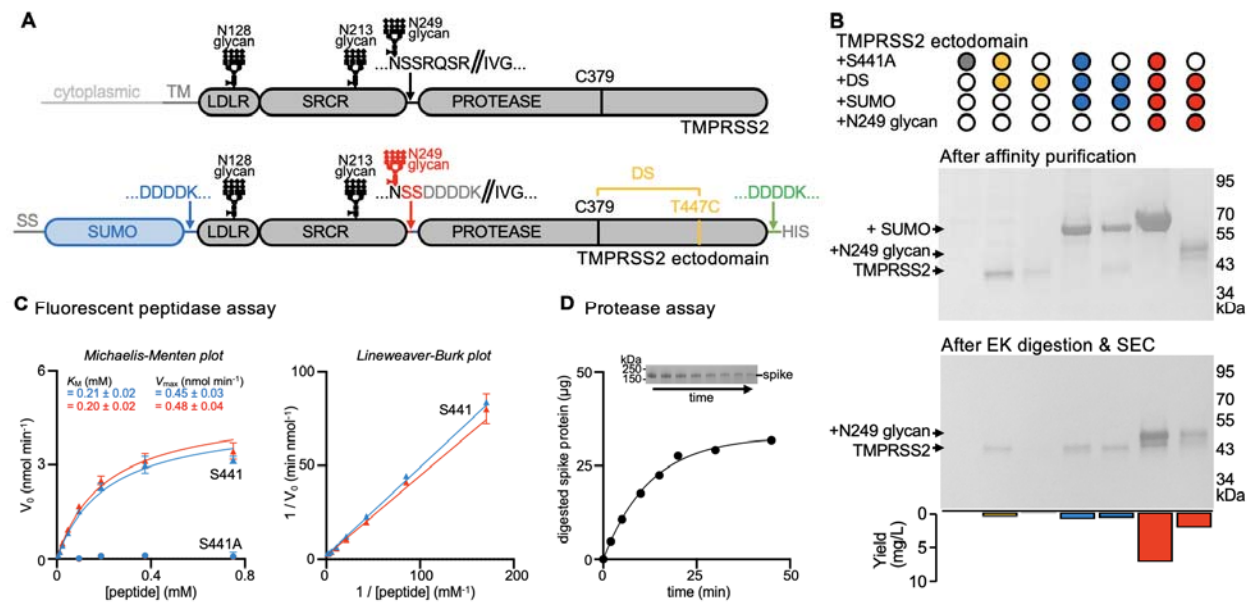
Given the therapeutic relevance of human TMPRSS2, we sought to design a strategy enabling recombinant production of native enzyme from human cells. Although a recent study reported producing the active TMPRSS2 ectodomain (dasTMPRSS2) using Sf9 cells<sup>44</sup>, insect cell expression is inherently time-consuming and does not faithfully recapitulate the chemical composition of protein glycosylations. Furthermore, the replacement of the native TMPRSS2 autoactivation cleavage site with an enterokinase cleavage sequence in dasTMPRSS2 removes the endogenous N249 glycosylation motif and also requires an additional proteolytic activation step, adding to the laborious nature of this purification.

To overcome these limitations, we designed an ectodomain construct enabling high-yield and rapid recombinant production of active human TMPRSS2. We first observed that we could not

detect dasTMPRSS2 expression in Expi293 human cells even after introducing the S441A residue mutation to disable proteolytic activity and potential cytotoxicity during expression (**Figure 1A-B, Figure S1 and Table S1**). Human TMPRSS2 harbors an unpaired and solvent-exposed cysteine at position 379 in the protease domain, which otherwise participates in a disulfide bond with the residue equivalent to T447 in most related type 2 transmembrane serine proteases<sup>44</sup>. We therefore introduced the T447C residue substitution to allow formation of a C379-C447 disulfide bond, which was sufficient to detect TMPRSS2 expression, albeit with low yield (**Figure 1A-B, Figure S1 and Table S1**). Fusion of an enterokinase-cleavable N-terminal SUMO tag increased production yield approximately 2-fold, relative to the construct lacking the SUMO fusion (**Figure 1A-B, Figure S1 and Table S1**). Finally, reintroduction of two native serine residues, corresponding to positions 250 and 251 of the human TMPRSS2 sequence, restored the N-linked glycan at position N249 and improved production yield 8-fold compared to the glycan knockout construct (**Figure 1A-B, Figure S1 and Table S1**).

Reintroduction of the S441 catalytic residue reduced yields ~3 fold relative to S441A-harboring TMPRSS2 (**Figure 1A-B, Figure S1 and Table S1**). We noted that the TMPRSS2 construct harboring the wildtype S441 residue, T447C, the N-terminal SUMO fusion, and the native N249 glycan autocatalytically processed the enterokinase cleavage sequences during expression, leading to zymogen activation and SUMO cleavage (**Figure 1A-B, Figure S1 and Table S1**). Consistent with this interpretation, the catalytically inactive S441A-harboring TMPRSS2 mutant remains uncleaved. Overall, our designed constructs streamline TMPRSS2 production yielding 7 mg of purified S441A TMPRSS2 or 2 mg of self-activating S441 TMPRSS2 per liter of human Expi293 cells, 4 days after transfection, which will accelerate studies of this key therapeutic target.

Purified TMPRSS2 is highly active in a fluorescent peptidase assay, using the fluorogenic Boc-Gln-Ala-Arg-7-aminomethylcoumarine (Boc-QAR-AMC) peptide substrate, and its activity can be modeled with Michaelis-Menten kinetics (**Figure 1C and Figure S1**). We determined a  $K_M$  of 200  $\mu$ M and  $V_{max}$  of 0.5 nmol/min for Boc-QAR-AMC, similar to dasTMPRSS2 purified from insect cells<sup>44</sup>. Given that reintroduction of the N249 glycan did not significantly alter TMPRSS2 enzyme kinetics, the observed autocatalytic cleavage likely resulted from a high TMPRSS2 concentration during production rather than enhanced activity (**Figure 1C and Figure S1**). TMPRSS2 is also active using the SARS-CoV-2 S ectodomain trimer as a substrate (**Figure 1D**).



**Figure 1. Design and functional characterization of an enzymatically active human TMPRSS2 construct**

(A), Schematic representation of the domain architecture of human TMPRSS2 and of the ectodomain construct designed for mammalian expression. The sequences of the constructs tested can be found in **Table S1**. Designed features are a C379-T447C disulfide bond (DS, yellow), an enterokinase-cleavable N-terminal SUMO fusion (blue), and reintroduction of the N249 glycosylation motif (red) relative to a previously described insect cell expression construct (dasTMPRSS2)<sup>44</sup>. LDLR: low-density lipoprotein receptor domain; SRCR: Scavenger Receptor Cysteine-Rich domain; PROTEASE: trypsin-like serine protease domain. SS: signal sequence. DDDDK: enterokinase cleavage sequence. (B) Non-reducing SDS-PAGE analysis of purified TMPRSS2 constructs with or without S441A, the C379-T447C disulfide bond (DS), an N-terminal SUMO fusion, and the native glycan at position N249. Proteins were expressed and affinity purified in parallel, and then normalized to a standard volume to enable comparisons of relative yields and purity after affinity purification (top) and after enterokinase (EK) digestion and SEC purification (middle). The final yield of purified TMPRSS2 per liter of Expi293 expression medium is indicated as a bar graph (bottom). (C) Michaelis–Menten plot (left) and Lineweaver–Burke plot (right) of initial reaction velocities with various concentrations of fluorescent Boc-QAR-AMC peptide substrate, at 22 °C, in the presence of 6.8 nM of the TMPRSS2 ectodomain harboring either the DS and SUMO (blue) or the DS, SUMO, and N249 glycan (red) modifications.  $K_M$  and  $V_{max}$  were determined through fitting using GraphPad Prism. Data are shown as the geometric mean and standard deviation of three technical replicates. (D) Proteolytic processing of 40  $\mu$ g of SARS-CoV-2 S 2P (S2P) by 0.5  $\mu$ M TMPRSS2 over 45 min at 37°C, visualized by reducing SDS-PAGE analysis (inset) and measured by densitometry. The inset on top shows the decreasing intensity of the S2P band over time used for these calculations.

## Structure of the TMPRSS2-bound HKU1 RBD

To understand the molecular basis of HKU1 engagement of its host receptor, we characterized an HKU1 isolate N1 (genotype A) RBD in complex with the human TMPRSS2 ectodomain harboring the S441A substitution of the catalytic nucleophile. We used cryoEM to determine a structure of this 85kDa complex at 2.9 Å resolution revealing that the TMPRSS2 catalytic domain is engaged by the HKU1 receptor-binding motif (RBM) which is inserted in between two β-strands of the RBD core β-sheet (**Figure 2A-D, Figure S2 and Table S2**). No major conformational changes occur upon binding besides small scale structural rearrangements of the HKU1 RBM mostly centered on the β-hairpin spanning residues 505-517, as compared to the isolated RBD structure<sup>45</sup> (**Figure S3**).

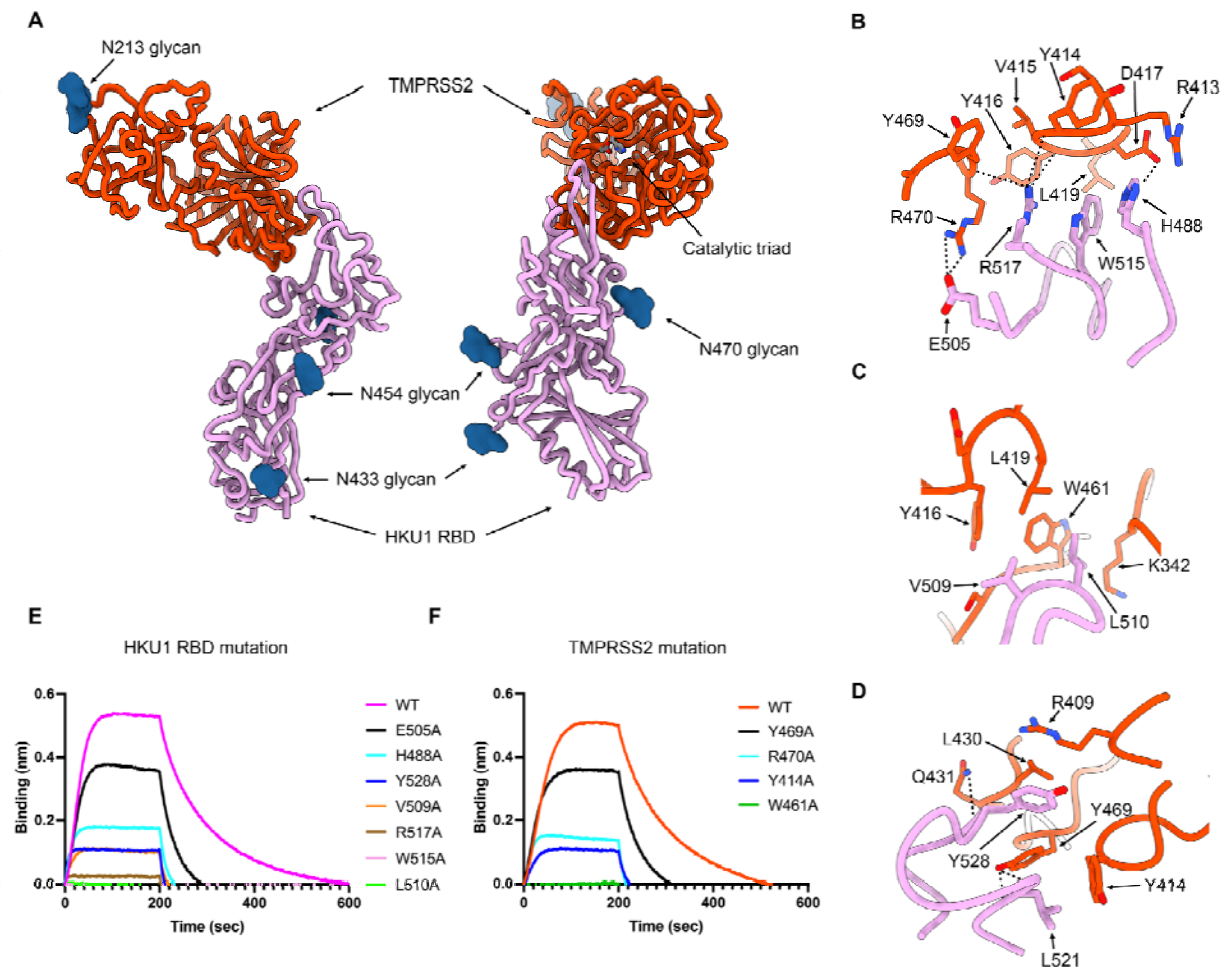
Receptor engagement leads to burial of surfaces of approximately 800 Å<sup>2</sup> from each of the two binding partners using electrostatic interactions and surface complementarity. The HKU1 RBM comprises residues 488, 505, 507-512, 515, 517-522, 527-532, and 554 to recognize TMPRSS2 residues 340-342, 409, 412-417, 419, 430-431, 433, 461-463 and 468-470 forming loops protruding from one of the two β-barrels of the protease domain (**Figure 2A**). Key interactions include (i) the H488<sub>HKU1</sub> imidazole side chain salt-bridged to the D417<sub>TMPRSS2</sub> side chain carboxylate (**Figure 2B**); (ii) the E505<sub>HKU1</sub> side chain carboxylate forming a salt bridge with the R470<sub>TMPRSS2</sub> side chain guanidinium (**Figure 2B**); (iii) the V509<sub>HKU1</sub> side chain establishing van der Waals contacts with the Y416<sub>TMPRSS2</sub>, L419<sub>TMPRSS2</sub> and W461<sub>TMPRSS2</sub> side chains (**Figure 2C**); (iv) the L510<sub>HKU1</sub> side chain inserting in a surface-exposed cleft and contacting the aliphatic part of the K342<sub>TMPRSS2</sub> side chain and the W461<sub>TMPRSS2</sub> indol ring (**Figure 2C**); (v) the W515<sub>HKU1</sub> side chain forming van der Waals contacts with Y416<sub>TMPRSS2</sub>, D417<sub>TMPRSS2</sub> and L419<sub>TMPRSS2</sub> (**Figure 2A**); (vi) the R517<sub>HKU1</sub> side chain electrostatically interacting with the Y414<sub>TMPRSS2</sub>, V415<sub>TMPRSS2</sub> and Y469<sub>TMPRSS2</sub> backbone carbonyls (**Figure 2A**); (vii) the L521<sub>HKU1</sub> side chain which is sandwiched between the Y414<sub>TMPRSS2</sub> and Y469<sub>TMPRSS2</sub> side chains and the L521 backbone amide and carbonyl that are hydrogen-bonded to the Y469<sub>TMPRSS2</sub> side chain phenol (**Figure 2D**); and (viii) the Y528<sub>HKU1</sub> side chain interacting with the R409<sub>TMPRSS2</sub>, L430<sub>TMPRSS2</sub> and Y469<sub>TMPRSS2</sub> side chains and the Y528<sub>HKU1</sub> backbone carbonyl which is hydrogen-bonded to the Q431<sub>TMPRSS2</sub> amide side chain (**Figure 2D**). As none of the HKU1 S or TMPRSS2 N-linked glycans contribute to the binding interface, the previously reported HKU1 S pseudovirus sensitivity to neuraminidase treatment of target cells<sup>29</sup> likely resulted from removal of sialoglycan receptors, thereby dampening entry.

## Validation of the TMPRSS2-bound HKU1 RBD cryoEM structure

To functionally validate the role of key interacting residues identified in our cryoEM structure, we evaluated binding of catalytically active or inactive TMPRSS2 to seven isolate N1 HKU1 RBD point mutants harboring an alanine substitution at the receptor-binding interface. The HKU1 RBD L510A and W515A mutations had the strongest effect among the mutants tested, completely preventing binding (**Figure 2E**). R517A had the next strongest effect, severely abrogating recognition, followed by V509A, Y528A, H488A and E505A (**Figure 2E**). The deleterious impact of the HKU1 RBD W515A and R517A mutations on TMPRSS2 binding concur with prior biochemical data<sup>29,45</sup>, underscoring the key role of these two viral residues. Furthermore, we found that the TMPRSS2 W461A, Y414A, R470A and Y469A interface



mutants had no or reduced binding to the HKU1 RBD (**Figure 2F**). These data demonstrate that the HKU1 S and TMPRSS2 residues identified as main contributors to binding in our cryoEM structure are mediating attachment of the RBD to its host receptor and that mutations of these residues to alanine is deleterious for TMPRSS2 binding.



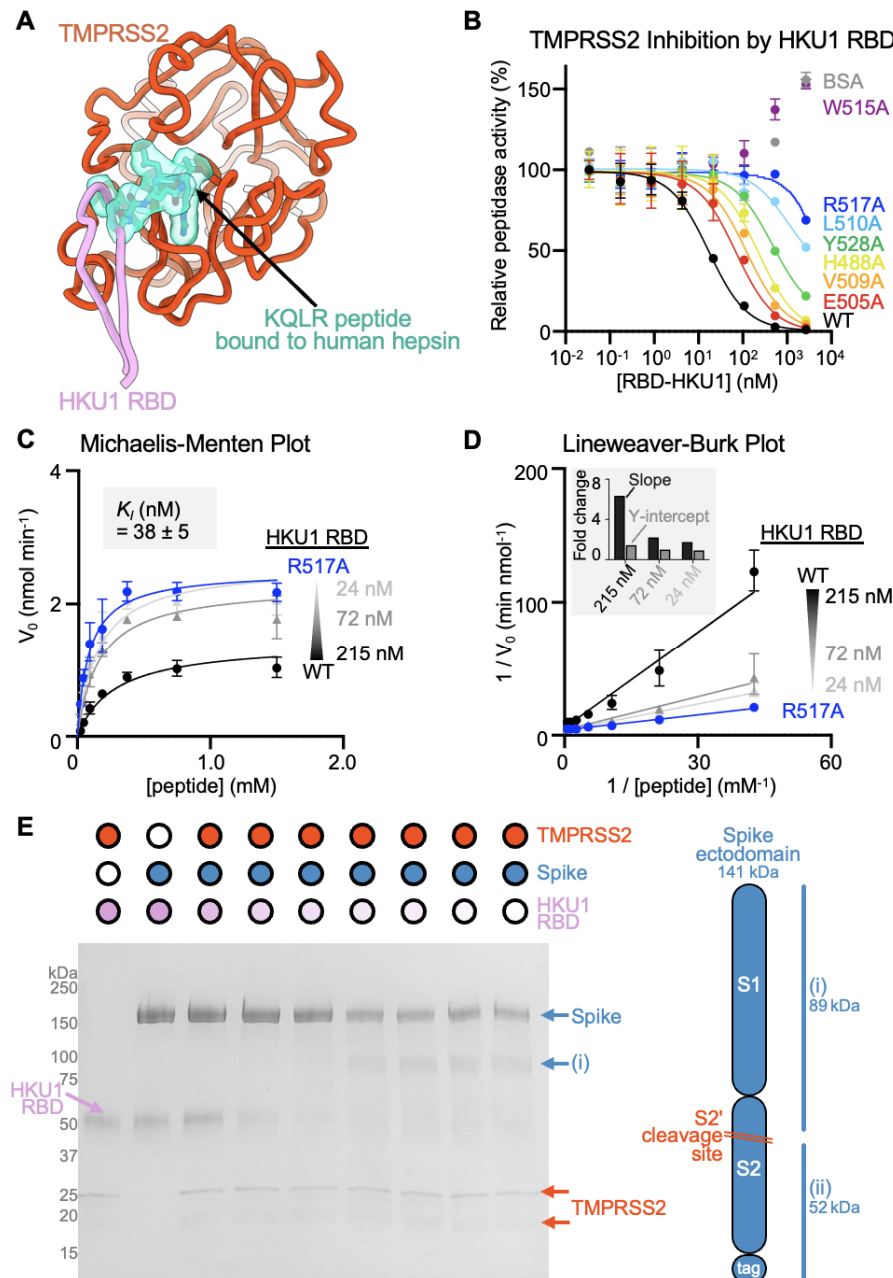
**Figure 2. Molecular basis of human TMPRSS2 recognition by the HKU1 RBD**

(A) Ribbon diagrams in two orthogonal orientations of the cryoEM structure of the HKU1 RBD (purple) bound to the human TMPRSS2 ectodomain (orange) at 2.9Å resolution. The TMPRSS2 catalytic triad residues (H296, D345 and S441) are shown as sticks colored by heteroatom and N-linked glycans are rendered as blue spheres. (B-D) Zoomed-in views of the interface highlighting key interactions between the HKU1 RBD and the human TMPRSS2 peptidase domain. Select electrostatic interactions are shown as black dotted lines. (E) Binding of S441A TMPRSS2 to the wildtype (WT) isolate N1 and to the H488A, E505A, V509A, L510A, W515A, R517A and Y528A HKU1 RBD interface mutants immobilized on biolayer interferometry SA biosensors. (F) Binding of the wildtype (WT) and the Y414A, W461A, Y469A and R470A human

TMPRSS2 (active S441) interface mutants to the wildtype isolate N1 HKU1 RBD immobilized on biolayer interferometry SA biosensors.

### HKU1 RBD binding inhibits TMPRSS2 activity

Upon binding, the HKU1 RBM  $\beta$ -hairpin (comprising residues 505-517) is positioned in close proximity to the active site albeit without contacting the catalytic triad residues (H296, D345 and S441) (**Figure 2A**). Superimposing the HKU1 RBD-bound TMPRSS2 structure with that of the KQLR chloromethylketone-bound human hepsin<sup>46</sup> indicates that the HKU1 RBM would compete with binding of this substrate to the TMPRSS2 active site through steric hindrance, assuming an identical binding mode of the ligand (**Figure 3A**). To assess a possible impact on TMPRSS2 enzymatic activity, we quantified proteolytic processing of the Boc-QAR-AMC peptide substrate in the presence of various concentrations of the HKU1 RBD. TMPRSS2 proteolytic activity was inhibited by the HKU1 RBD in a dose-dependent manner with an inhibition constant  $K_i$  of 38  $\pm$  5 nM (**Figure 3B-D and Figure S4**). Furthermore, the seven HKU1 RBD alanine point mutants at the receptor-binding interface had dampened inhibitory effect on TMPRSS2 activity (**Figure 3B and Figure S4**), concurring with the binding reductions observed by biolayer interferometry (**Figure 2E**). The W515A HKU1 RBD mildly increased TMPRSS2 peptidase activity, as was the case for bovine serum albumin (BSA), suggesting non-specific TMPRSS2 stabilization. Michaelis-Menten kinetics analysis of TMPRSS2 activity in the presence of various concentrations of the HKU1 RBD suggested a mechanism of action through competitive inhibition (**Figure 3C-D**), in line with partial occluding of the active site observed structurally. Furthermore, the HKU1 RBD inhibited cleavage of the purified SARS-CoV-2 S 2P trimer in a dose-dependent manner, as evaluated by SDS-PAGE (**Figure 2F**). These results underscore the lack of requirement for TMPRSS2 enzymatic activity to promote HKU1 entry, as the R255Q and the S441A inactive mutants were shown to support HKU1 S-mediated pseudovirus entry with comparable levels to wildtype TMPRSS2<sup>29</sup>.



**Figure 3. HKU1 RBD binding inhibits human TMPRSS2 activity**

(A) Superimposition of the KQLR chloromethylketone-bound hepsin structure (PDB ID 1Z8G<sup>46</sup>) to the HKU1-bound TMPRSS2 structure showing that the ligand would be precluded sterically from binding in the active site upon attachment of the HKU1 RBD (assuming an identical binding mode of the ligand). Only the TMPRSS2 peptidase domain is shown and hepsin is omitted for clarity. (B) Assessment of inhibition of human TMPRSS2 activity by the wildtype (WT) isolate N1 and interface mutant HKU1 RBDs using the fluorescent Boc-QAR-AMC peptide substrate in the presence of 3.4 nM of the TMPRSS2 ectodomain harboring the DS, SUMO, and N249 glycan. Data are shown as the geometric mean and standard deviation of 2-3 technical replicates. (C-D) Michaelis-Menten plot (C) and Lineweaver-Burke plot (D) of initial reaction velocities in function of the concentration of fluorescent Boc-QAR-AMC peptide substrate, at 22 °C, in the presence



of 5 nM TMPRSS2 and the wildtype (WT) isolate N1 or R517A interface mutant HKU1 RBDs. The inhibitor constant  $K_i$  was determined through fitting competitive inhibition using GraphPad Prism. Data are shown as the geometric mean and standard deviation of three technical replicates. The inset in panel D shows the relative change in slope ( $K_M/V_{max}$ ) and y axis intercept ( $1/V_{max}$ ) for the different concentrations of wildtype HKU1 RBD relative to the R517A RBD. (E) Reducing SDS-PAGE analysis of proteolytic processing at 37°C of 5 µg of SARS-CoV-2 S 2P (S2P) by 0.5 µM TMPRSS2 preincubated with various concentrations of the HKU1 isolate N1 RBD.

### Molecular basis of HKU1 specificity for TMPRSS2

Our structural data provide a blueprint to understand the HKU1 specificity for the TMPRSS2 receptor. Amino acid sequence alignment of TMPRSS2 orthologs reveals that the HKU1 binding site is mostly conserved among mammals, partially conserved in reptiles and birds, and poorly conserved in amphibians and other vertebrates (**Figure 4A-C**). We scored the compatibility with the HKU1 RBD of residues mutated relative to human TMPRSS2 using proteinMPNN<sup>47</sup> and observed high probabilities for mammalian TMPRSS2 residues whereas other TMPRSS2s had lower scores. Among weakly conserved residues, mutations at positions Y414, D417, and Y469 were predicted to have the greatest reductions of HKU1 RBD binding to mammalian TMPRSS2 orthologs (**Figure 4A-C**). Moreover, mutations at positions T341, Y416, L419, and S463 were anticipated to particularly impact HKU1 RBD binding to non-mammalian TMPRSS2s.

To test this hypothesis, we expressed human TMPRSS2 harboring the K340D, T341S, R409T, S412N, R413K, Y414S, V415I, D417N, L419A/M, S463T/Y/F, A468P, Y469N, or R470K mutations and analyzed binding to the HKU1 isolate N1 wildtype RBD (**Figure 4D and Figure S5**). All mutants bound more weakly to the HKU1 RBD than wildtype TMPRSS2 with the Y414S/L, D417N, L419A, S463T/Y/F, and Y469N causing marked binding reductions, consistent with the compatibility probability score (**Figure 4C**). Contrary to its predicted score, R470K in TMPRSS2 was also severely impaired in its ability to bind the HKU1 RBD, possibly due to a requirement of R470 pi-stacking interactions with HKU1 R517 to overcome charge repulsion effects<sup>48</sup>. These findings likely narrow the TMPRSS2 species tropism of HKU1 as mammalian TMPRSS2s frequently harbor D417N and Y469N substitutions sometimes accompanied by mutations of residue Y414. Furthermore, S463 and L419 are not conserved for non-mammalian TMPRSS2s and R470 is only partially conserved among reptile/bird TMPRSS2s and not conserved for TMPRSS2s found in amphibians and other vertebrates. The combination of these residue substitutions along with the Y469 and Y414 mutations, are expected to limit the ability of the HKU1 RBD to engage non-mammalian TMPRSS2 orthologs.

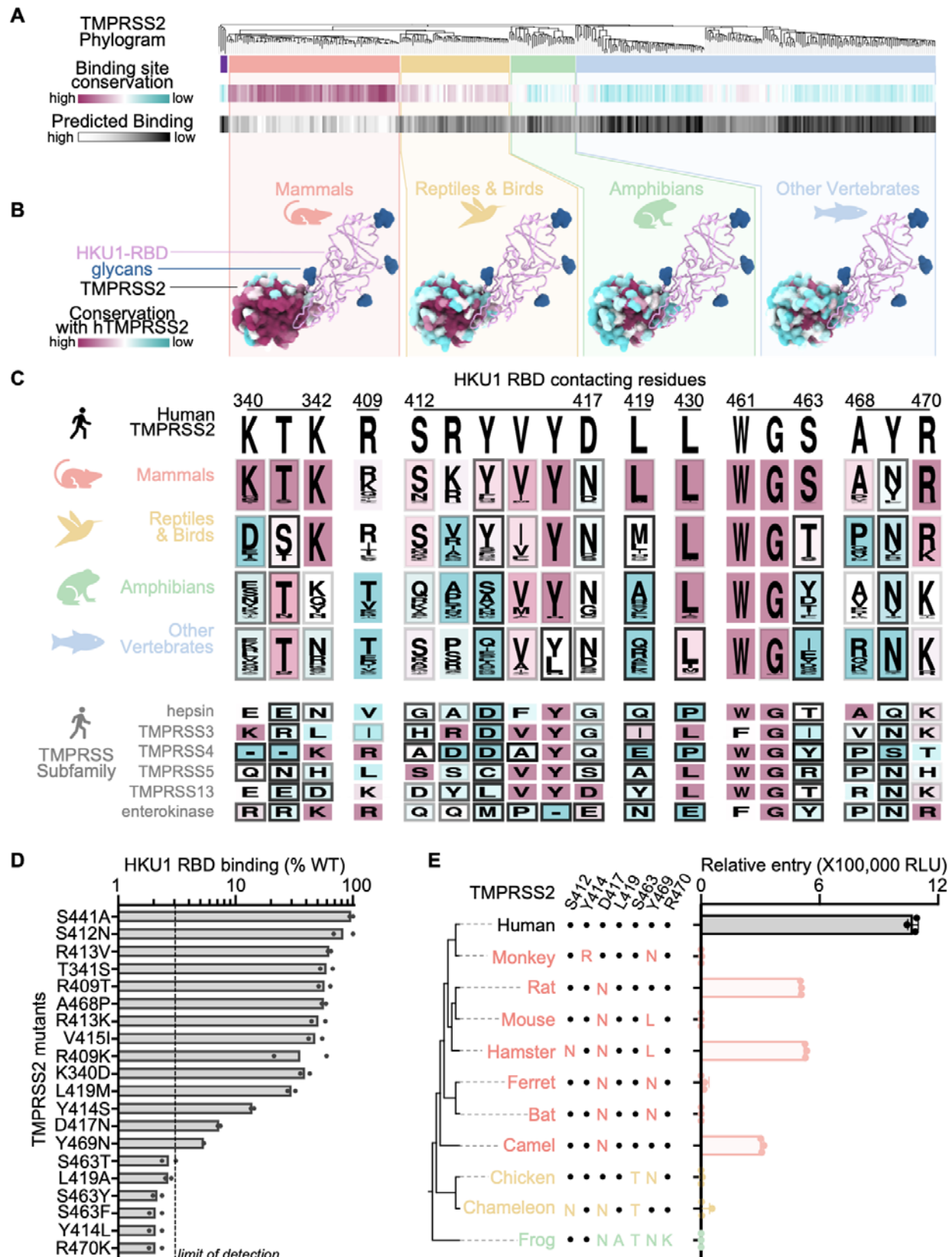
Using a vesicular stomatitis virus (VSV) pseudotyped with HKU1 isolate N1 S, we found that transient transfection of rat, hamster and camel TMPRSS2s promote HKU1 S-mediated entry into HEK293T cells, in addition to human TMPRSS2 (**Figure 4F**). Each of these three non-human TMPRSS2 orthologs harbors the D417N residue substitution, suggesting that this mutation alone does not prevent entry. Although rat TMPRSS2 (possessing D417N) and hamster TMPRSS2 (possessing D417N, Y469L, and S412N mutations) rendered cells

susceptible to HKU1 S VSV, mouse TMPRSS2 (possessing D417N and Y469L mutations) did not (**Figure 4F**). We predict that a mutation (potentially S412N) in hamster TMPRSS2 epistatically offsets a possible binding defect from Y469L. The lack of HKU1 S VSV entry into cells transiently transfected with ferret, greater horseshoe bat, and African green monkey TMPRSS2s is rationalized by the reduced HKU1 RBD binding of the TMPRSS2 Y469N and D417N mutants and likely the Y414R mutant. As anticipated for TMPRSS2 orthologs harboring S463T, R470K, or L419A mutations, non-mammalian (chicken, African clawed frog, and chameleon) TMPRSS2s did not support HKU1 S VSV entry. Collectively, our data explain the efficient HKU1 use of human TMPRSS2 as entry receptor and reveal that rat, hamster and camel TMPRSS2 orthologs support HKU1 S-mediated entry into cells.

### **Functional consequences of TMPRSS2 site mutants in humans**

We found several low frequency TMPRSS2 single nucleotide polymorphisms (SNPs) at residue positions involved in HKU1 binding and several of them correspond to missense mutations observed in TMPRSS2 paralogs and orthologs (**Figure 4C and Data S1-S2**). Among the SNPs we characterized experimentally, some of the mutations reduced binding to the HKU1 RBD (T341S, R409T/K, S412N, V415I and A468P) whereas others almost entirely abrogated binding (S463F and R470K). The most frequently sampled SNP is V415I (rs148125094, allele frequency:  $1.1 \times 10^{-3}$ ) which led to reduced but detectable binding to the HKU1 isolate N1 RBD relative to wildtype human TMPRSS2 (**Figure 4D, Figure S5 and Data S1-S2**). The S463F SNP (rs2146419758, allele frequency:  $2.1 \times 10^{-6}$ ), however, abrogated binding to the wildtype HKU1 RBD, concurring with the fact that none of the energetically favored phenylalanine side chain rotamers at this position would be sterically compatible with the binding interface in our cryoEM structure (**Figure 4D, Figure S5 and Data S1-S2**). In summary, although several TMPRSS2 SNPs map to the site of HKU1 attachment, some of them (e.g. S463F and R470K) possibly reducing susceptibility to HKU1 infection, the low frequency of TMPRSS2 SNPs is unlikely to majorly impact HKU1 transmission in the human population.

Residues involved in binding to the HKU1 RBD, markedly diverge across human type 2 transmembrane serine proteases with G462 being the sole residue strictly conserved among them (**Figure 4C**). Members of the TMPRSS protease subfamily harbor non-conservative mutations at several positions involved in HKU1 binding relative to TMPRSS2. This includes L419A (TMPRSS5), S463T/Y (hepsin, TMPRSS4, TMPRSS13, and enterokinase), and Y469N (TMPRSS3, TMPRSS5, TMPRSS13, and enterokinase), which we found to abrogate TMPRSS2 binding to the HKU1 RBD (**Figure 4C-D and Figure S5**). These data therefore rationalize the lack of HKU1 S-mediated fusion observed in the presence of TMPRSS3, TMPRSS4, and TMPRSS13<sup>29</sup> and explain the HKU1 specificity for TMPRSS2.



**Figure 4. Molecular determinants of HKU1 utilization of human TMPRSS2 and host receptor tropism.** (A, top) Neighbor-Joining phylogram of TMPRSS2 orthologs, clustered in red (mammals), yellow (reptiles/birds), green (amphibians), and blue (other vertebrates). The tree is rooted (far left, purple) on TMPRSS3 4, 5, and 13. (A, middle) The average amino acid sequence conservation (BLOSUM62) of the HKU1 RBD-contacting residues in each TMPRSS2 ortholog, relative to human TMPRSS2, is rendered from maroon (conserved) to cyan (not conserved). (bottom) As a proxy for predicted binding compatibility, the ProteinMPNN<sup>47</sup> probability score of the HKU1 RBD-contacting residues in each TMPRSS2 ortholog was calculated and rendered from white (highest score, eg. human TMPRSS2) to black (lowest score). (B) Average amino acid sequence conservation (BLOSUM62) with human TMPRSS2 (hTMPRSS2) of residues from the TMPRSS2 clusters in (A) plotted on the surface of the human TMPRSS2 structure and rendered from maroon (conserved) to cyan (not conserved). The HKU1 RBD is shown as pink ribbons with N-linked glycans rendered as blue surfaces). (C) Logo plots of amino acids equivalent to the TMPRSS2 residues interacting with the HKU1 RBD from the ortholog clusters in (A) and from members of the TMPRSS protease subfamily colored by conservation with human TMPRSS2 from maroon (conserved) to cyan (not conserved). The shade of the frame for each residue represents the predicted binding score from proteinMPNN. (D) Binding of TMPRSS2 point mutants (corresponding to polymorphisms identified in TMPRSS2 orthologs or in humans) to the HKU1 RBD immobilized on biolayer interferometry SA biosensors. Data are shown as the area under the binding response curve normalized to the signal of WT TMPRSS2. Data are shown as the mean of two technical replicates presented in **Figure S5**. Dashed line shows the practical limit of detection. (E) HKU1 S VSV entry into HEK293T cells transiently transfected with one of the indicated TMPRSS2 orthologs. Key TMPRSS2 amino acid residues interacting with the HKU1 RBD are indicated to highlight conserved and divergent positions relative to human TMPRSS2. Data are shown as the geometric mean and standard deviation of three technical replicates.

### Glycan shielding and conformational masking mediates immune evasion

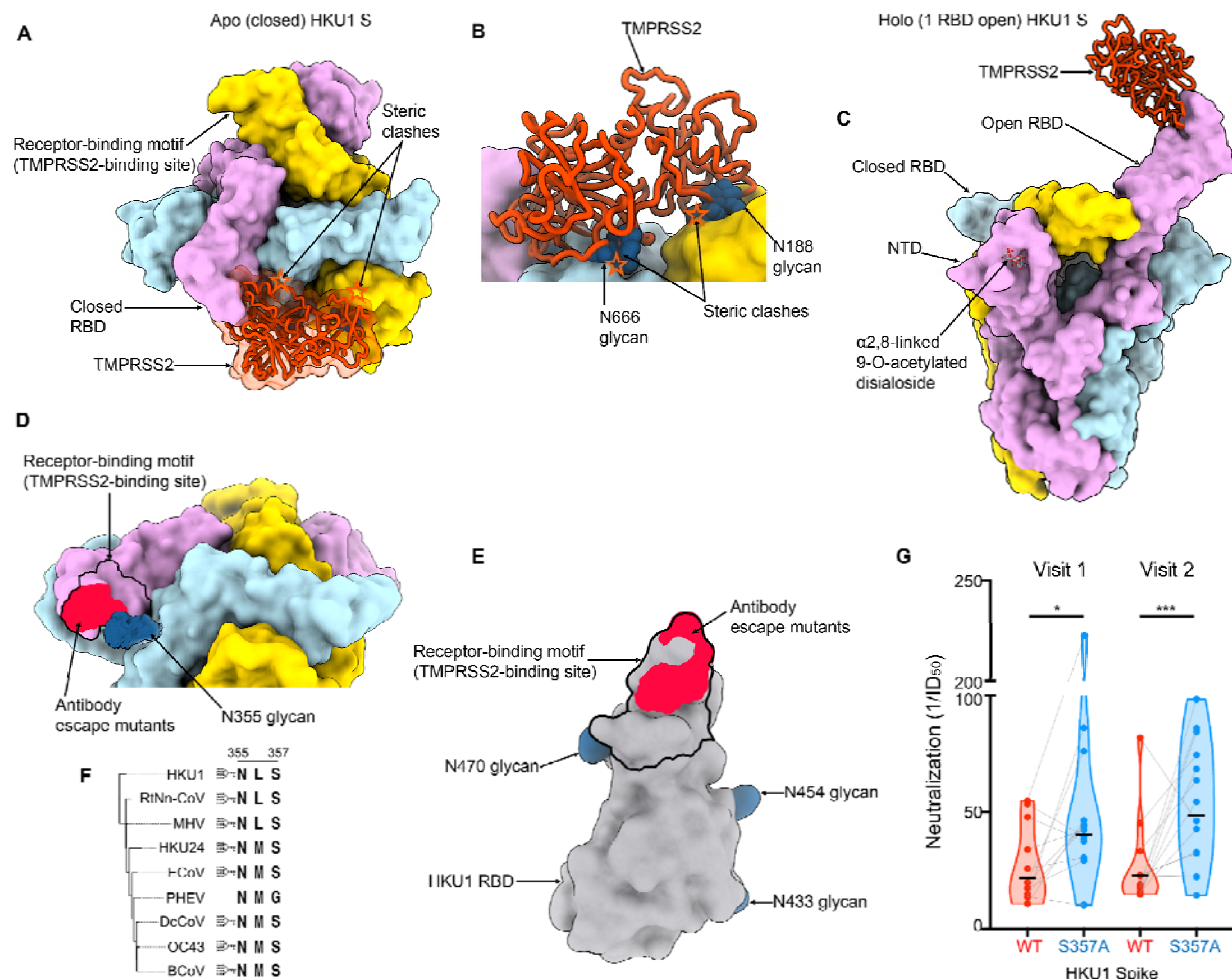
Although the HKU1 RBM protrudes out from the closed S trimer, leading to markedly increased solvent exposure relative to sarbecovirus RBMs<sup>12,49</sup>, TMPRSS2 binding would be precluded due to steric hindrance involving both proteinaceous and oligosaccharide moieties from neighboring protomers (**Figure 5A-B**). Sialoglycan receptor binding, which is known to promote RBD opening<sup>30</sup>, would relieve these steric constraints and enable TMPRSS2 attachment to an RBD in the open conformation (**Figure 5C**). Similar to other coronaviruses, this cascade of events is expected to lead to fusion of the viral and host membranes to promote genome delivery to a target cell and initiation of infection<sup>49-51</sup>.

Strikingly, the HKU1 glycan at position N355 from a neighboring RBD masks the RBM (**Figure 5D**), as observed in an apo closed HKU1 S structure<sup>30</sup>, sterically limiting RBM accessibility in this conformation. As a result, residues D511, H512, W515 and R517, which map to the RBM and contribute to the epitopes recognized by the previously described mHKUS-2 and -3 neutralizing monoclonal antibodies<sup>45</sup>, are also partially masked by the HKU1 N355 glycan which would reduce targeting by this class of antibodies (**Figure 5D-E**). Given the overlap between the epitopes of the mHKUS-2 and -3 antibodies and the RBM, the primary mechanism of

neutralization for these antibodies is most likely through direct competition with HKU1 attachment to the TMPRSS2 host receptor. Given the conservation of this glycan in most embecovirus S glycoproteins (**Figure 5F**), similar masking is anticipated to occur, as confirmed in the prefusion OC43 S structure<sup>26</sup> although the role of the topological equivalent of the RBD (domain B) remains unclear for many members of this subgenus.

To investigate the relationship between antibody specificity and viral inhibition, we assessed neutralizing antibody titers against an HKU1 isolate N1 S VSV using human sera collected after PCR-confirmed HKU1 infection. These serum samples were collected in four European countries between 2007 and 2010 at the time of PCR-positive testing (Visit 1) and 1-2 months later (Visit 2)<sup>52,53</sup>. Neutralizing antibody titers ranged from half-maximum inhibition dilution of 1/10 to 1/80 for a geometric mean of 1/23 for visit 1 and 1/49 for visit 2 (**Figure 5G and Figure S6**). Serum neutralizing activity was enhanced upon removal of the N355 glycan (via S357A mutation), relative to the wildtype HKU1 S, concurring with the partial masking of the RBM by this oligosaccharide (**Figure 5G and Figure S6**). These findings indicate that RBM-directed serum antibodies are a key component of neutralization, in line with prior studies of potent sarbecovirus and merbecovirus antibodies<sup>49,54–59</sup>. Collectively, our data reveal that HKU1 relies on conformational masking and glycan shielding to balance conflicting requirements for immune evasion and receptor engagement, ultimately modulating viral fitness.





**Figure 5. Glycan shielding and conformational masking mediate HKU1 immune evasion**

(A-B) Composite model obtained by superimposing the TMPRSS2 (orange)-bound HKU1 RBD cryoEM structure onto the apo closed HKU1 S trimer structure (PDB 8OHN)<sup>30</sup> rendered as a surface with each protomer colored distinctly viewed in two orthogonal orientations. In the closed trimer conformation, TMPRSS2 binding would be hindered by steric clashes (orange stars) with neighboring protomers, including with N-linked glycans N188, N355, N666 (shown as blue spheres). (C) Composite model obtained by superimposing the TMPRSS2 (orange)-bound HKU1 RBD cryoEM structure onto the  $\alpha$ 2,8-linked 9-O-acetylated disialoside-bound HKU1 S trimer with one open RBD (PDB 8OPN)<sup>30</sup> engaging the TMPRSS2 receptor. (D) Zoomed-in view of the apo closed HKU1 S trimer structure (PDB 8OHN)<sup>30</sup> showing the unsharpened cryoEM density for the N355 glycan partially masking the receptor-binding motif (black outline) and the site of vulnerability to the mHKUS-2 and -3 neutralizing antibodies<sup>45</sup> (neon red). The extent of the N355 glycan-mediated glycan shielding is expected to be even greater than shown as only the first three monosaccharides are resolved in the map. (E) HKU1 RBD rendered as a grey surface with the TMPRSS2 footprint delineated with a black outline and the escape mutants (D511, H512, W515, R517) identified for the neutralizing antibodies mHKUS-2 and -3 shown in neon red. The HKU1 RBD N-linked glycans are rendered as blue surfaces. (F) Neighbor-Joining phylogram of embecovirus S glycoproteins focused on the HKU1 N355 glycosylation site. Included S sequences are HKU1 (AZS52618.1), RtNn-CoV (RtNn-CoV), MHV, HKU24, HCoV, PHEV, DcCoV, OC43, and BCuV. (G) Neutralization ( $1/ID_{50}$ ) for HKU1 Spike. Visit 1 and Visit 2. WT and S357A. Significance: \* p < 0.05, \*\*\* p < 0.001.

CoV/SAX2015, ATP66762.1), mouse hepatitis virus (MHV, C0KYY9), HKU24 (A0A866W1F1), Equine coronavirus (ECoV, E5RPZ2), Porcine hemagglutinating encephalomyelitis virus (PHEV, A0A1Z2WUW0), Dromedary Camel coronavirus (DcCoV, HKU23, A0A291L0R6), OC43 (P36334) and bovine coronavirus (BCoV, P15777). (G) Neutralization of VSV pseudotyped with wildtype isolate N1 HKU1 S (WT) or the N355 glycan knockout (S357A) mutant by serum antibodies elicited by HKU1 infection. Visits 1 and 2 correspond to blood draws at the time of PCR-positive testing and 1-2 months later, respectively<sup>52,53</sup>. Neutralization values were compared by two-way ANOVA with the Geisser-Greenhouse correction and Tukey's multiple comparison test; \* corresponds to a P value of 0.02, while \*\*\* corresponds to a P value of 0.0007.

## Discussion

HKU1 genomes sequences are divided into three genotypes designated A, B and C with evidence of recombination among them<sup>5,9</sup>. The S glycoproteins of genotypes A and B share 85% amino acid sequence identity whereas that of genotypes B and C are almost identical<sup>29</sup>. Furthermore, HKU1 RBDs are highly divergent with ~74% amino acid sequence identity between the HKU1 isolate N1 RBD (genotype A) and the isolate N2 RBD (genotype B) or the isolate N5 RBD (genotype C). This magnitude of sequence divergence is comparable to that observed among the RBDs of distinct sarbecovirus clades<sup>60,61</sup> and may participate in explaining that protective immunity to seasonal coronaviruses is of short duration due to antigenic drift<sup>62,63</sup>. Furthermore, HKU1 infections elicit modest neutralizing antibody titers (at least in the cohort analyzed here), as compared to SARS-CoV-2 infection<sup>64</sup>, which may further compound the durability and breadth of humoral immune responses. The enhanced neutralization of an HKU1 S mutant lacking the N355 glycan, relative to wildtype, with a panel of human sera collected after HKU1 infection show that RBM-directed antibodies are making a key contribution to neutralizing activity. These findings may explain the observed polymorphism of TMPRSS2-interacting residues among HKU1 isolates and are reminiscent of observations made for SARS-CoV-2<sup>55,65,66</sup> and 229E<sup>63,67</sup>. Our structural data therefore provide a molecular blueprint to follow evolution of RBM residues and their impact on receptor binding and immune evasion. The identification of the RBM as a key site of vulnerability further suggests that removal of the N355 glycan could be a suitable vaccine design strategy to improve the potency of antibody responses elicited by HKU1 S and for most viruses from the embecovirus subgenus. Although the OC43 domain B is not known to be a functional RBD, potent neutralizing antibodies recognizing this region were described<sup>68</sup>, underscoring the possible key role of this conserved embecovirus glycan for immune evasion. Since HKU1 recognizes 9-O-acetylated sialosides using the S NTD<sup>27,30,69,70</sup>, similar to OC43 and bovine coronavirus<sup>19,26</sup>, the isolation of a neutralizing antibody competitively inhibiting sialoside binding to the OC43 NTD<sup>68</sup> suggests that this mechanism of action is likely conserved against many, if not all, lineage A  $\beta$ -coronaviruses attaching to cell-surface carbohydrates. Future studies will elucidate the importance and possible synergy of these different types of antibodies, along with that of non-neutralizing antibodies, for in vitro neutralization and in vivo protection against these viruses.

Rapid antigenic evolution, masking of epitopes and exposure of immune-dominant epitopes targeted by non-neutralizing antibodies are immune-evasion strategies widely used by viruses. Coronavirus S trimers can adopt at least two distinct conformations of their RBDs enabling to mask or expose the RBM in the closed and open states, respectively. Although the SARS-CoV-2, SARS-CoV-1 and MERS-CoV S trimers appear to populate this conformational landscape spontaneously<sup>12,49,71,72</sup>, most coronavirus S trimers appear to preferentially adopt a closed conformation<sup>18,26,51,73–79</sup>. Recent work showed that  $\alpha$ 2,8-linked 9-O-acetylated disialosides binding to the HKU1 S NTD induces conformational changes leading to RBD opening<sup>30</sup>, which our data reveal to expose the TMPRSS2-binding site. This elegant mechanism would allow to ensure proper spatial and temporal coordination of these conformational changes, ensuring that RBD opening occur upon attachment to a target cell, thereby balancing the conflicting requirements for receptor binding to initiate infection (which requires RBD opening) and immune evasion through masking of a key site of vulnerability (RBM) recognized by polyclonal serum neutralizing antibodies (which requires RBD closing).

Although HKU1 has been proposed to have originated in rodents, due to phylogenetic clustering with rodent coronaviruses<sup>80</sup>, evidence supporting or invalidating this hypothesis is still lacking. We identified the molecular determinants of HKU1 utilization of TMPRSS2 and revealed that rat, hamster and camel, but not mouse, TMPRSS2 orthologs support HKU1 S-mediated entry into cells. These data are compatible with a possible rodent origin of HKU1 and the putative involvement of the species identified here as reservoir or intermediate hosts. We note, however, that host tropism is a complex process involving additional factors than receptor recognition, including proteolytic spike activation and innate immune antagonism, for a successful infection to occur. Considering the wide use of golden syrian hamsters as small animal models<sup>59,81,82</sup>, our data suggest a possible path to develop an HKU1 challenge model to study pathogenicity, correlates of protection and the effectiveness of countermeasures.

## Acknowledgements

This study was supported by the National Institute of Allergy and Infectious Diseases (P01AI167966, DP1AI158186 and 75N93022C00036 to D.V.), a Pew Biomedical Scholars Award (D.V.), an Investigators in the Pathogenesis of Infectious Disease Awards from the Burroughs Wellcome Fund (D.V.), the University of Washington Arnold and Mabel Beckman cryoEM center and the National Institute of Health grant S10OD032290 (to D.V.). D.V. is an Investigator of the Howard Hughes Medical Institute and the Hans Neurath Endowed Chair in Biochemistry at the University of Washington. We gratefully acknowledge the Biobank Antwerp, Belgium (ID: BE 71030031000) for providing the human sera collected in the GRACE study ([www.grace-lrti.org](http://www.grace-lrti.org)) supported by the 6th Framework Program of the European Commission, contract no. LSHM-CT-2005-518226).

## Author Contributions

MM, YJP, and DV conceived the study and designed the experiments; MM designed the engineered TMPRSS2 constructs and carried out activity assays. JB performed the S2P cleavage assay. MM designed and cloned protein constructs. MM, CS and JB recombinantly expressed and purified glycoproteins. CS performed binding assays. Y.J.P. carried out cryoEM

specimen preparation, data collection and processing. Y.J.P. and D.V built and refined atomic models. EW and AT carried out the single nucleotide polymorphism analysis. G.I. provided unique reagents. MM, YJP, and D.V. analyzed the data and wrote the manuscript with input from all authors. DV supervised the project.

### **Competing Interests**

M.M. and D.V. are named as inventors on a patent describing the designed TMPRSS2 constructs and D.V. is named as inventor on patents for coronavirus vaccines filed by the University of Washington. E.W. and A.T. are employees of Vir Biotechnology and may hold shares in Vir Biotechnology. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## **MATERIALS AND METHODS**

### **Cells**

Cell lines used in this study were obtained from ATCC (HEK293T and VeroE6) or Thermo Fisher Scientific (ExpiCHO-S cells and Expi293F cells).

### **Constructs**

The membrane-anchored HKU1 isolate N1 S (ref. seq. YP\_173238, genotype A) with a 21 amino acid C-terminal deletion and six residues flexible linker followed by C-terminal HA tag were codon-optimized, synthesized, and inserted the pcDNA3.1(+) vector by Genscript. The HKU1 RBD constructs encoding S residues 320-614 of the wildtype isolate N1 (ref. seq. YP\_173238, genotype A) and the H488A, E505A, V509A, L510A, W515A, R517A and Y528A HKU1 RBD interface mutants, containing N-terminal mu-phosphatase secretion signal peptide, with or without an avi tag followed by three residues flexible linker and C-terminal octa-histidine tag were codon optimized, synthesized, and inserted into the pcDNA3.1(+) vector by Genscript. The membrane-anchored TMPRSS2 and TMPRSS2 ectodomain constructs were codon-optimized, synthesized, and inserted into a pCMVR vector. TMPRSS2 ectodomain constructs were further mutated by In-Fusion Assembly (Takara Bio). All TMPRSS2 sequences are listed in **Table S1**.

The SARS-CoV-2 S 2P (Wuhan-Hu-1) ectodomain construct<sup>12</sup> harbors an N-terminal mu-phosphatase secretion signal peptide, S1/S2 cleavage site R682S/R683G/R685G mutations, K986P/V987P mutations, one residue linker followed by C-terminal hexa-histidine tag and foldon trimerization domain was codon-optimized, synthesized, and inserted the pCMV vector by Genscript.

### **Generation of pseudoviruses**

To produce pseudoviruses for entry and neutralization assays, HEK293T cells were seeded in DMEM enriched with 10% FBS (Hyclone), 1% Penstrep at the appropriate density to yield 80% confluency in polylysine-coated 100 mm cell culture dishes and placed in an incubator at 37°C

with 5% CO<sub>2</sub>. After 18-22 hr incubation, cells were washed with OPTI-MEM (Life Technologies). 24µg of either the wildtype HKU1 isolate N1 (genotype A) S or the S357A mutant S plasmids were prepared in 1.5mL of OPTI-MEM and combined with 60µL of Lipofectamine 2000 (Life Technologies) diluted in 1.5mL of OPTI-MEM and incubated at room temperature for 15-20 min. The mixture was added to the HEK293T cells which were placed for 2 hours in an incubator at 37°C with 5% CO<sub>2</sub> after which 2mL of DMEM enriched with 20% FBS and 2% PenStrep was added to the transfected cells and incubated overnight. The following day, cells were washed with DMEM and transduced with VSVΔG/Fluc and incubated for 2 hours at 37°C with 5% CO<sub>2</sub>. After washing with DMEM, medium supplemented with anti-VSV-G antibody (I1-mouse hybridoma supernatant diluted to 1:25 from CRL-2700, ATCC) was added to the cells to reduce background from the parental virus and an additional incubation at 37°C with 5% CO<sub>2</sub> was performed overnight. The next day, the supernatant from the cells was harvested from the 100 mm dishes, further clarified by centrifugation at 3,000xg for 10 minutes, filtered (0.45µm), and concentrated 10 times by using centrifugal devices with 30 kDa cutoff membranes. Pseudoviruses were then aliquoted and frozen at -80°C until used.

### **Pseudovirus neutralization and entry assays**

HEK293T cells were cultured in DMEM with 10% FBS (Hyclone), 1% PenStrep and placed in cell-culture grade, polylysine-coated 100 mm plates overnight in an incubator at 37°C with 5% CO<sub>2</sub>. The following day, when cells reached 90% confluency, they were transfected with 4 µg of TMPRSS2 and incubated for 5 hours. The cells were subsequently trypsinized, counted and reseeded at ~40,000 cells per well in poly-lysine-coated 96 well plates and placed in an incubator at 37°C with 5% CO<sub>2</sub> overnight. The next day, a half-area 96-well plate was prepared with a 1:3 serial dilution of sera in DMEM in 22 µL final volume. 22 µL of pseudovirus was then added to each well and incubated at room temperature for 30-45 min. The media was removed from transfected HEK293T cells and 40 µL of the sera/pseudovirus mixture was added to the cells and incubated for 2 h at 37°C with 5% CO<sub>2</sub> before adding 40 µL of 20% FBS and 2% PenStrep containing DMEM. No sera were used for assessment of entry with the panel of TMPRSS2 orthologs. Following 18-22 hour incubation, 40 µL of One-GloEX (Promega) was added to the cells and incubated in the dark for 5 min prior to reading on an Agilent BioTek Neo2 plate reader. Relative luciferase units were plotted and normalized in Prism (GraphPad) using a zero value of cells alone and a 100% value of 1:2 virus alone. Nonlinear regression of log(inhibitor) vs. normalized response was used to determine ID<sub>50</sub> values from curve fits. At least two biological replicates with two distinct batches of pseudovirus were conducted for each serum sample.

### **Recombinant glycoprotein production**

The SARS-CoV-2 S 2P ectodomain, HKU1 RBDs or TMPRSS2 ectodomains were produced and purified using Expi293F or ExpiCHO-S cells. Expi293F cells were grown to a density of 3 x 10<sup>6</sup> cells/mL and transfected using the ExpiFectamine 293 Transfection Kit (ThermoFisher Scientific) and expression carried out for 3 to 5 days post-transfection at 37°C with 8% CO<sub>2</sub>. ExpiCHO-S cells were grown to a density of 6 x 10<sup>6</sup> cells/mL and transfected using the ExpiFectamine CHO Transfection Kit (ThermoFisher Scientific) and expression carried out for 7-



10 days post-transfection at 37°C with 8% CO<sub>2</sub>. The SARS-CoV-2 S 2P ectodomain and HKU1 RBDs were purified from clarified supernatants using HisTrap HP affinity columns (Cytiva) and washed with ten column volumes of 10 mM imidazole, 25 mM sodium phosphate pH 8.0, and 300 mM NaCl before elution with two column volumes of 300 mM imidazole, 25 mM sodium phosphate pH 8.0, and 300 mM NaCl. The purified HKU1 RBD proteins were buffer exchanged into 20 mM sodium phosphate pH 8 and 100 mM NaCl or 50 mM Tris pH 8.0 and 150 mM NaCl. The HKU1 RBDs were biotinylated using the BirA biotin ligase reaction kit (Avidity) following the manufacturer's protocol. Biotinylation was carried out at room temperature for 30 minutes followed by incubation for 10 hours at 4°C. TMPRSS2 constructs were purified using TALON metal affinity resin (Takara Bio) and washed with 100 column volumes of 5 mM imidazole, 50mM HEPES pH 7.5, and 150mM NaCl prior to elution with 600 mM imidazole, 50mM HEPES pH 7.5, and 150mM NaCl. Purified TMPRSS2s were diluted to 0.3 mg/ml and then digested with a 1/500 dilution of enterokinase (EKmax, ThermoFisher) for 16 hours at room temperature or 1 hour at 37 °C to convert the zymogen in active enzyme, remove the SUMO tag and the His tag. Digested TMPRSS2 was then passed through fresh TALON metal affinity resin to remove enterokinase and the cleaved His tag. Active TMPRSS2 with the DS mutation, SUMO tag, and restored N249 glycan activated itself during expression and purification, and did not require enterokinase activation. Purified HKU1 RBD and TMPRSS2 – except for TMPRSS2 used for Figure 4D – were then further purified by size exclusion chromatography using a Superdex 200 Increase 10/300 GL column (Cytiva) and concentrated using centrifugal filters (Amicon Ultra) before being flash frozen.

### **Biolayer interferometry**

Biotinylated HKU1 wildtype (isolate N1, genotype A) RBD and H488A, E505A V509A, L510A, W515A, R517A and Y528A HKU1 RBD interface alanine mutants were diluted to concentrations of 0.002 mg/mL in modified kinetics buffer (1% BSA, 0.06% Tween20, 1x Cold Spring Harbor PBS pH 7.4) and loaded onto pre-hydrated streptavidin biosensors to a 1 nm total shift. The loaded tips were dipped into 100 nM TMPRSS2 (harboring the S441A residue substitution and the C379-T447C disulfide bond) for 200 or 300 seconds followed by dissociation in modified kinetics buffer for 500 seconds. Analysis of TMPRSS2 interface alanine mutants was carried out similarly but with constructs harboring S441 and the C379-T447C disulfide bond. Binding curves were plotted using Graphpad Prism 10.1.2. The area under the curves were determined by summation of the binding response values for each reading, with a limit of binding detection of 0.05 nm. Sums were normalized to WT, as was the limit of detection which corresponded to 2 ± 0.5 % (the practical limit of binding detection was set to two standard deviations above this value, i.e. 3.1 % of WT binding).

### **TMPRSS2 enzymatic activity assay**

Peptidase assays were performed with Boc-QAR-AMC substrate (GlpBio) in black half area 96 well plates (Greiner Bio-One Fluotrac) in 100 µl reaction volumes in a Agilent BioTek Microplate Reader at 22 °C, monitoring fluorescence every two minutes over 30 minutes at 341 nm excitation and 441 nm emission. The slope of the fluorescence curve over the first 10 minutes (corresponding to <5 % substrate conversion) was used for velocity calculations. AMC concentration was calculated using standard curves at each substrate and HKU1 RBD

concentration, to correct for the inner filter effect. For determination of TMPRSS2 enzyme kinetics, serial dilutions of Boc-Gln-Ala-Arg-AMC substrate were reacted with 6.8 nM TMPRSS2. For determination of the  $K_i$  of the HKU1 RBD, 5 nM TMPRSS2 was reacted with serial dilutions of Boc-QAR-AMC substrate in the presence of 215, 72, or 24 nM HKU1-RBD (TMPRSS2 was incubated for 2 minutes with the HKU1 RBD at 22 °C prior to the addition of substrate). Velocities were plotted and fit using GraphPad Prism (v10.1.1).

### **Spike Cleavage Assay**

Spike cleavage assays were performed with SARS-CoV-2 S-2P ectodomain in 20mM phosphate pH 8.0 and 100mM NaCl incubated with 500nM TMPRSS2 in 50 mM HEPES pH 7.5 150mM NaCl. Samples were diluted in 50 mM HEPES pH 7.5 150mM NaCl, mixed, and incubated at 37°C with 300 rpm shaking. Reactions were stopped at various time points by mixing with 4x gel loading buffer containing 200mM Tris pH 6.8, 8% SDS, 20% β-mercaptoethanol, 40% Glycerol, 0.2% Bromophenol blue. The HKU1 RBD inhibition assays were performed by incubating the HKU1 RBD (without avi tag) with TMPRSS2 (harboring the N-terminal SUMO fusion and the C379-T447C disulfide bond) in 50 mM HEPES pH 7.5 150mM NaCl for 5 minutes at 37°C prior to adding SARS-CoV-2 S 2P and incubating at 37°C with shaking at 600 rpm for 15 minutes prior to stopping the reaction by mixing with 4x gel loading buffer containing 200mM Tris pH 6.8, 8% SDS, 20% β-mercaptoethanol, 40% Glycerol, 0.2% Bromophenol blue. Samples were then analyzed by SDS-PAGE.

### **Human serum sample donors**

The patients/participants provided their written informed consent to participate in this study which was reviewed and approved by Antwerp University Hospital. Human sera were collected after PCR-confirmed HKU1 infection in four European countries between 2007 and 2010 at the time of PCR-positive testing (Visit 1) and 1-2 months later (Visit 2). Demographic data can be found in **Data S3**.

### **Single nucleotide polymorphism (SNP) analysis**

The SNP information was extracted from two human population allele frequency database, [GnomAD \(version 4\)](#) and [Regeneron's Million Exome Variant database](#). For GnomAD v4, we utilized the allele frequencies of short genetic variations from the exome sequencing data of 730,947 individuals. These genetic variations have passed the QC (i.e. "FILTER" column in the VCF = "PASS"). GnomAD v4 provided the functional consequences of the genetic variations from the [VEP tool](#). We obtained the predicted deleteriousness of mutations from AlphaMissense<sup>83</sup>. For Regeneron's Million Exome Variant data, we utilized the allele frequencies of short genetic variations from the exome sequencing data of 983,578 individuals. We observed a total of 49 short genetic variations in the GnomAD v4 and the Regeneron's Million Exome Variant databases.

### **CryoEM sample preparation, data collection and data processing**

Complex formation was performed by mixing a 1:1.2 molar ratio of the human TMPRSS2 ectodomain (harboring the S441A residue substitution and a C379-T447C disulfide bond) with the HKU1 isolate N1 RBD (residues 320-614 with C-terminal octa-histidine tag and no avi tag)

before incubation for 1 hour at 4°C. CryoEM grids of the complex were prepared using three separate methods and data were combined during data processing. For the first dataset, 3  $\mu$ L of 1.1 mg/mL or 0.5 mg/mL of complex were loaded onto freshly glow discharged R 2/2 UltrAuFoil grids<sup>84</sup> prior to plunge freezing using a Vitrobot Mark IV (ThermoFisher Scientific) with a blot force of 0 and 6 sec blot time at 100% humidity and 22°C. 6,681 movies from UltrAuFoil grids with 1.1 mg/mL protein complex were collected with a defocus range comprised between -0.2 and -3  $\mu$ m and stage tilt angle of 0° and 20°<sup>85</sup>. 3,820 movies from UltrAuFoil grids with 0.5 mg/mL protein complex were collected with a defocus range comprised between -0.2 and -3  $\mu$ m and stage tilt angle of 30° and 45°. For the second dataset, 3  $\mu$ L of 1 mg/mL complex was added to the glow discharged side of R 2/2 UltrAuFoil grids and 1  $\mu$ L was added to the back side before plunging into liquid ethane using a GP2 (Leica) with 6 sec blot time. 9,883 movies were collected with a defocus range comprised between -0.2 and -3  $\mu$ m and stage tilt angle of 0° and 30°. For the third dataset, 3  $\mu$ L of 8 mg/mL complex with 3 mM CHAPSO or 0.01% fluorinated octyl-maltoside (FOM) (Anatrace) were applied onto freshly glow discharged R 2/2 UltrAuFoil grids prior to plunge freezing using a vitrobot MarkIV (ThermoFisher Scientific) with a blot force of 0 and 5 sec blot time at 100% humidity and 22°C. 6,822 and 1,859 movies were collected from UltrAuFoil grids with CHAPSO and FOM detergents, respectively, with a defocus range comprised between -0.2 and -3.5  $\mu$ m. The data were acquired using an FEI Titan Krios transmission electron microscope operated at 300kV and equipped with a Gatan K3 direct detector and Gatan Quantum GIF energy filter, operated in zero-loss mode with a slit width of 20 eV. Automated data collection was carried out using Leginon<sup>86</sup> at a nominal magnification of 105,000 $\times$  with a pixel size of 0.835 Å. The dose rate was adjusted to 15 counts/pixel/s, and each movie was acquired in super-resolution mode fractionated in 75 frames of 40 ms. Movie frame alignment, estimation of the microscope contrast-transfer function parameters, particle picking, and extraction were carried out using Warp<sup>87</sup>. Particles were extracted with a box size of 168 pixels with a pixel size of 1.67Å. Two rounds of reference-free 2D classification were performed using CryoSPARC<sup>88</sup> to select well-defined particle images from each dataset, which were subsequently combined. Initial model generation was done using ab-initio reconstruction in cryoSPARC and used as references for a heterogeneous 3D refinement in cryoSPARC. Particles belonging to classes with the best resolved HKU1 RBD and TMPRSS2 density were selected. To improve particle picking further, we trained the Topaz<sup>89</sup> picker on Warp-picked particle sets belonging to the selected classes after heterogeneous 3D refinement. The particles picked using Topaz were extracted and subjected to 2D classification using cryoSPARC, which improved the number of unique 2D views. The two different particle sets picked from Warp and Topaz were merged and duplicate particle picks were removed using a minimum distance cutoff of 60Å. After two rounds of heterogeneous refinements and removal of junk particles, 3D refinement was carried out using non-uniform refinement with per-particle defocus refinement in cryoSPARC<sup>90</sup> and the particles were transferred from cryoSPARC to Relion using pyem (<https://github.com/asarnow/pyem>) to be subjected to the Bayesian polishing procedure implemented in Relion<sup>91</sup> during which particles were re-extracted with a box size of 280 pixels and a pixel size of 1.0 Å. Subsequent 3D refinement used non-uniform refinement along with per-particle defocus refinement in cryoSPARC to yield the final reconstruction at 2.9 Å resolution comprising 810,357 particles. Reported resolutions are based on the 0.143 gold-

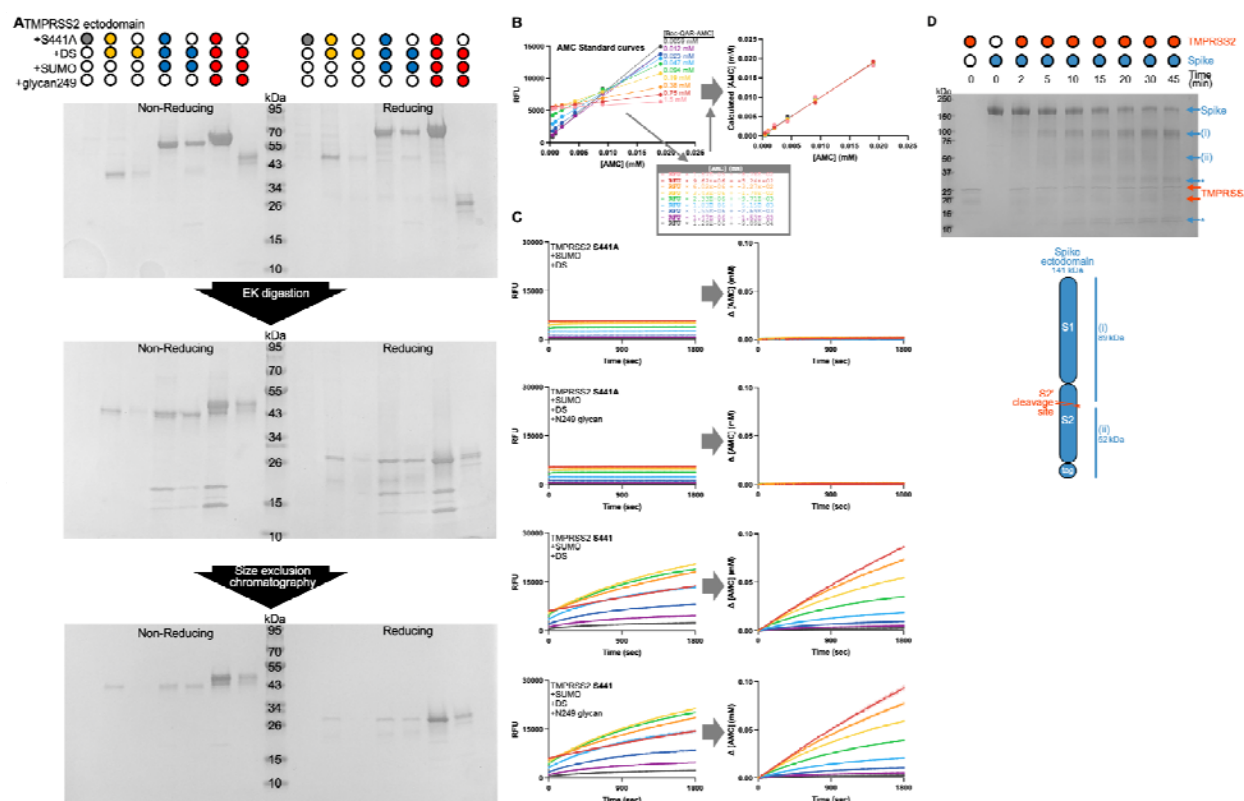
standard Fourier shell correlation (FSC) criterion<sup>92,93</sup>. Local resolution estimation, filtering, and sharpening were carried out using cryoSPARC.

## Model building and refinement

UCSF Chimera<sup>94</sup> was used to rigid-body dock models into the sharpened cryoEM map and adjustments and refinement were carried out with Coot<sup>95</sup> and Rosetta<sup>96,97</sup> using sharpened and unsharpened maps. Validation used Molprobity<sup>98</sup>, Phenix<sup>99</sup> and Privateer<sup>100</sup>.

## Data Availability

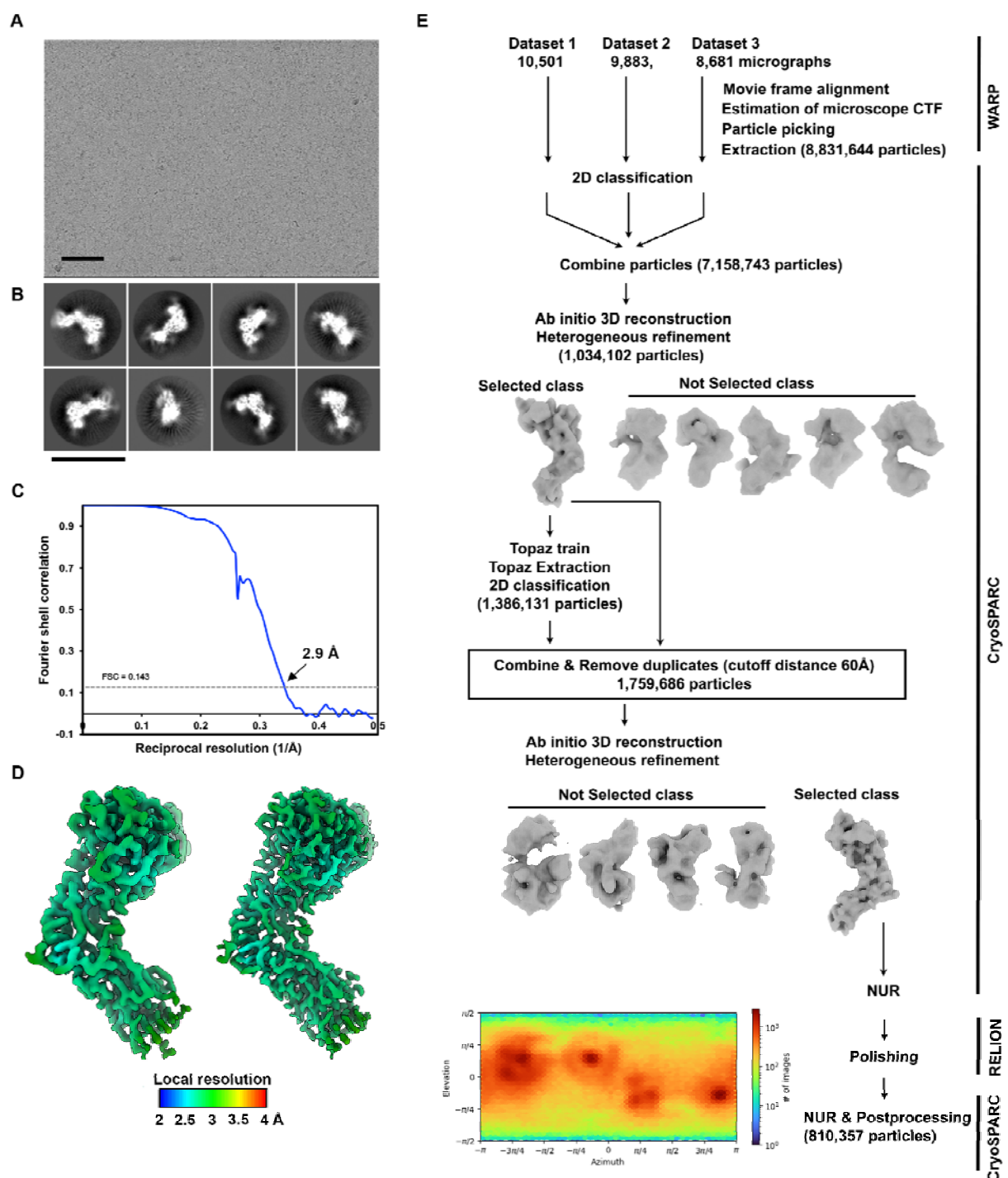
The cryoEM map and atomic model have been deposited to the EMDDB and PDB with accession codes EMD-43224 and PDB 8VGT. All datasets generated and information presented in the study are available from the corresponding authors on reasonable request. Materials generated in this study can be available on request and may require a material transfer agreement.



**Figure S1. Design and functional characterization of TMPRSS2 ectodomain constructs.** (A) Reducing and non-reducing SDS-PAGE analysis of the samples shown in Figure 1B, including the intermediate step between enterokinase (EK) digestion and size-exclusion chromatography (SEC) purification. (B) Summary of AMC standard curves used to calculate AMC release from the Boc-QAR-AMC peptide substrate accounting for the inner filter effect.

The left panel shows the relationship between AMC concentration and measured fluorescence (RFU). The lower inset shows the numerical equations of the linear relationships shown in the top panel. The right panel shows the calculated AMC concentration for each known AMC concentration shown in the left panel, using the equations in the lower inset at the indicated Boc-QAR-AMC concentrations. (C) Raw data and calculations used for Figure 1C; the color key is identical to Figure S1B. The left panels show the raw RFU over time for TMPRSS2 with and without the S441A mutation or the restored N249 glycan, while the right panel shows the change in calculated AMC concentration over time, using the equations from Figure S1B. (D) Reducing SDS-PAGE used for analysis shown in Figure 1D. Expected products (i) and (ii), based on TMPRSS2 cleavage at the S<sub>2</sub>' site, are labeled, along with additional cleavage products (asterisks). Reaction progress was monitored by densitometry of the S peak only.

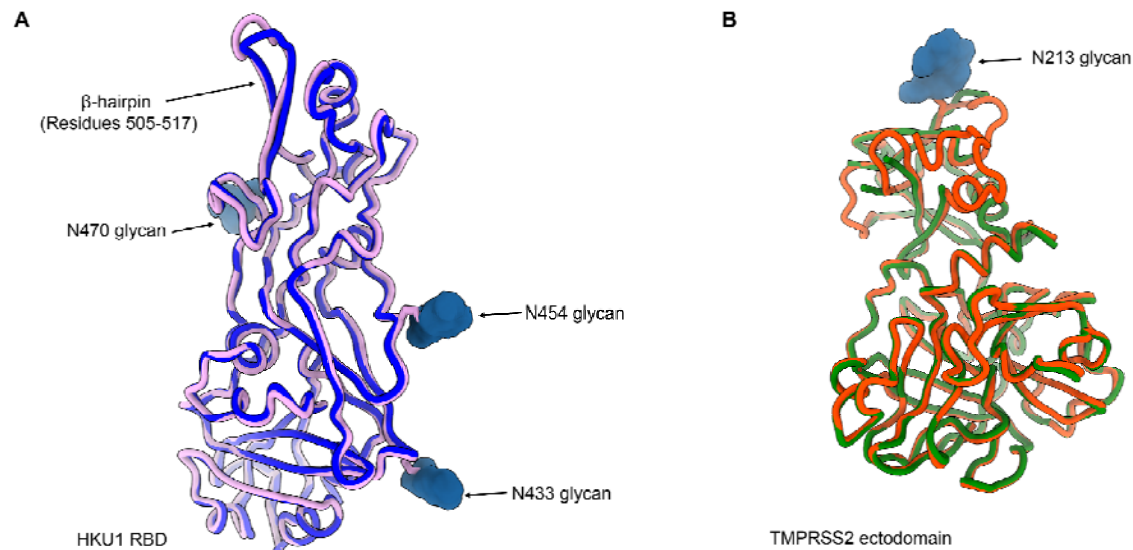




**Figure S2. CryoEM data processing of the TMPRSS2-bound HKU1 RBD dataset**

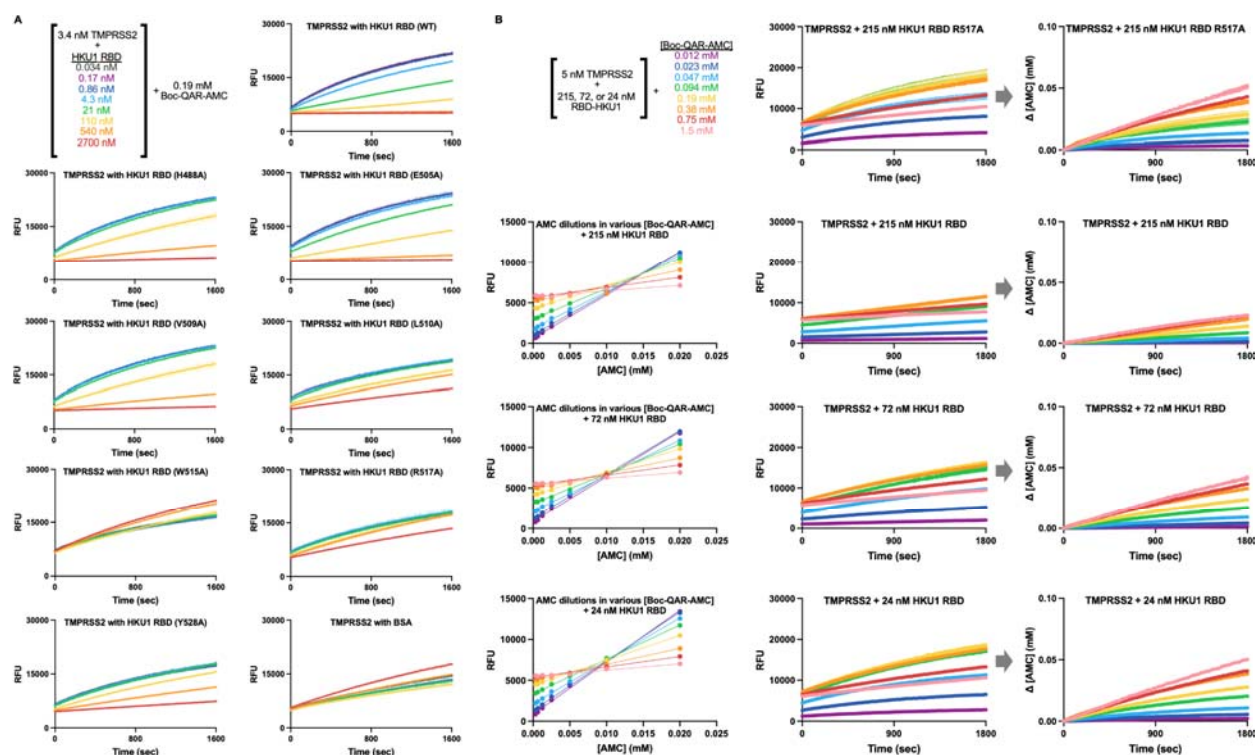
(A-B) Representative electron micrograph and 2D class averages of the TMPRSS2-bound HKU1 RBD complex embedded in vitreous ice. The scale bars represent 100 nm and 150 Å respectively. (C) Gold-standard Fourier shell correlation curve. The 0.143 cutoff is indicated by a horizontal dashed line. (D) Local resolution estimation calculated using cryoSPARC and plotted on the unsharpened (left) and sharpened (right) maps. (E) Data processing flowchart. CTF: contrast transfer function; NUR: non-uniform refinement. The angular distribution calculated in

cryoSPARC for particle projections is shown. The heat map shows the number of particles for each viewing angle.

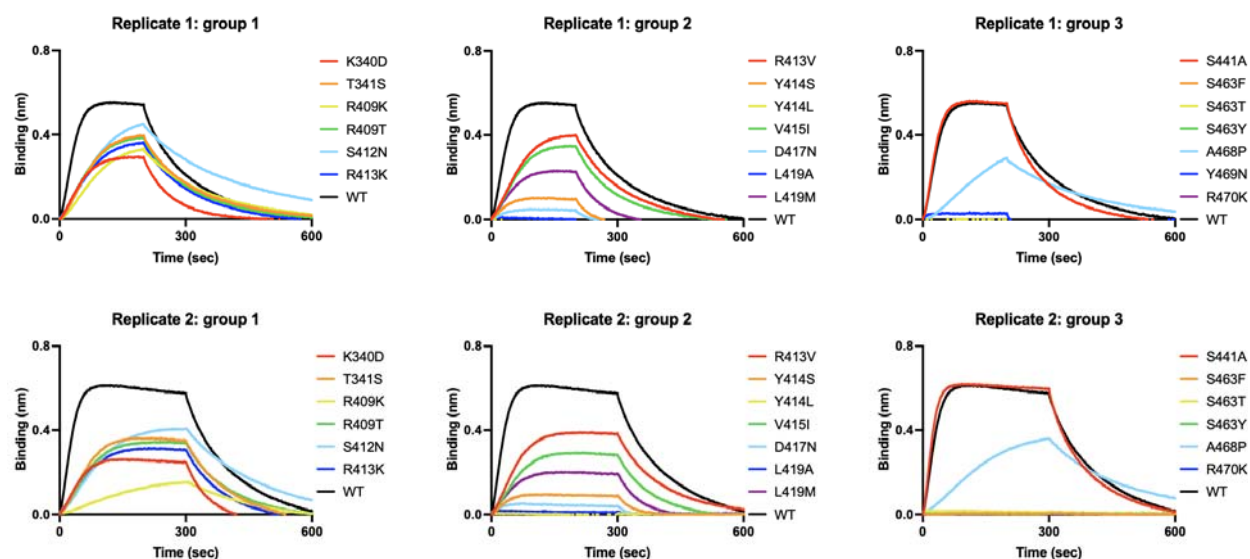


**Figure S3. Comparison of the HKU1 RBD and TMPRSS2 structures to the TMPRSS2-bound HKU1 RBD cryoEM structure.**

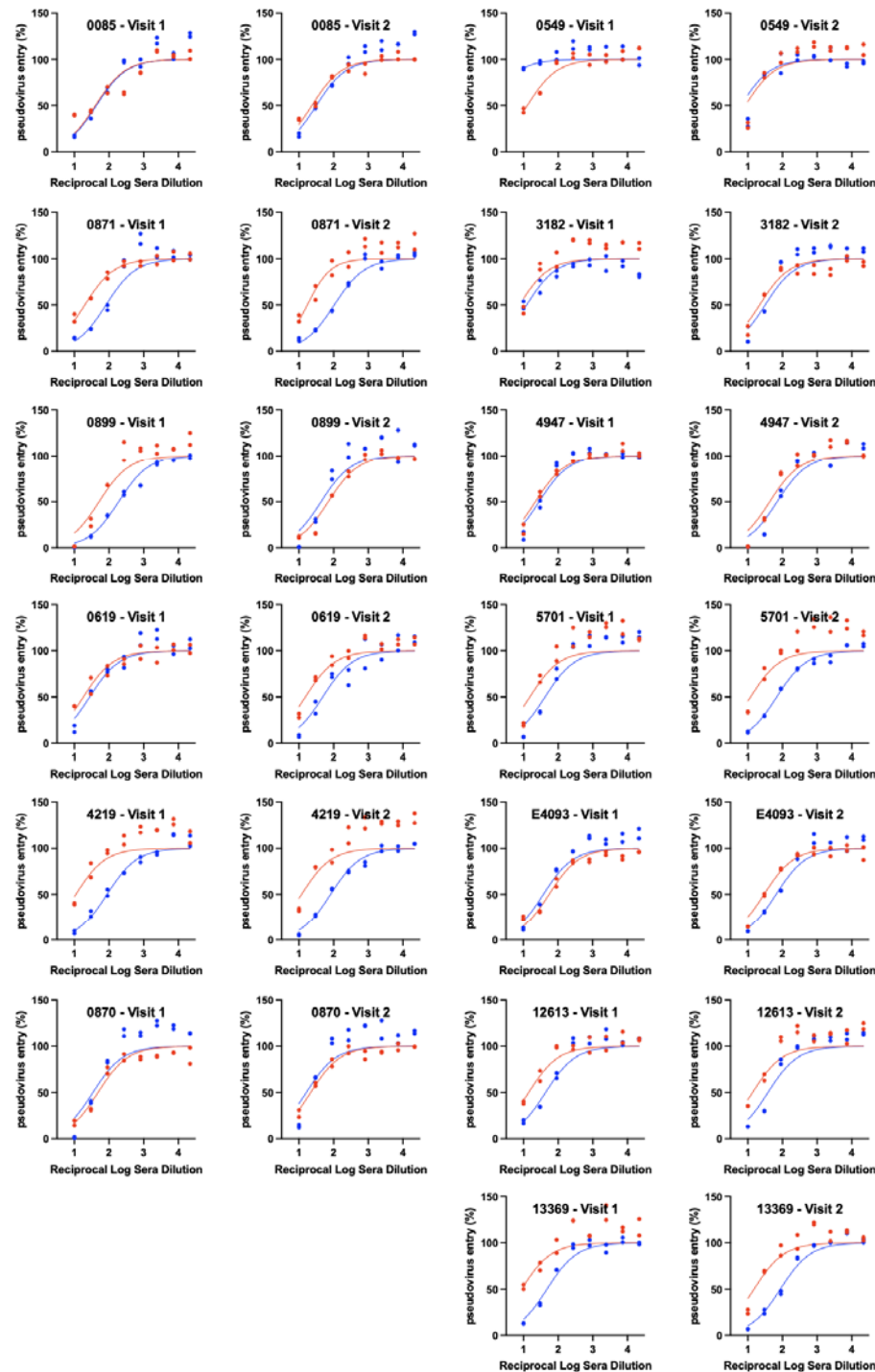
(A) Ribbon diagram of the cryoEM structure of the HKU1 RBD (purple) bound to the human TMPRSS2 ectodomain superimposed to the crystal structure of the apo HKU1 RBD (blue, PDB 5KWB). The TMPRSS2 ectodomain is omitted for clarity. (B) Ribbon diagram of the cryoEM structure of the human TMPRSS2 ectodomain (red) bound to the HKU1 RBD superimposed to the crystal structure of the nafamostat-bound TMPRSS2 (green, PDB 7MEQ). The HKU1 RBD is omitted for clarity.



**Figure S4. Enzymatic characterization of TMPRSS2 inhibition by the wildtype isolate N1 and alanine interface mutant HKU1 RBDs.** (A) Raw data used for Figure 3B; the upper left panel shows the color key used for the other panels. Data show measured fluorescence (RFU) over time with 3.4 nM TMPRSS2 (harboring the N-terminal SUMO fusion, the C379-T447C disulfide bond and the N249 glycan) incubated with 0.19 mM Boc-QAR-AMC at 22°C in the presence of various concentrations of HKU1 isolate N1 RBD and alanine interface mutant RBDs. (B) Raw data and calculations used for Figure 3C-D; the upper left panel shows the color key used for the other panels. Lower left panels show a summary of AMC standard curves used to calculate AMC release from the Boc-QAR-AMC peptide substrate with 215, 72, and 24 nM HKU1 RBD, accounting for the inner filter effect. The middle panels show measured fluorescence (RFU) over time with 5 nM TMPRSS2 and 215, 72, or 24 nM WT HKU1 RBD, or 215 nM HKU1 R517A RBD, incubated with various concentrations of Boc-QAR-AMC at 22 °C. The right panel shows the change in calculated AMC concentration over time.



**Figure S5. Binding of TMPRSS2 point mutants to the biotinylated HKU1 RBD immobilized on biolayer interferometry SA biosensors.** The top and bottom rows show baseline-subtracted response curves for the first and second biological replicates. For the first replicate, biotinylated HKU1 RBD-loaded SA tips were dipped into 100 nM TMPRSS2 for 200 seconds followed by dissociation for 500 seconds. For the second replicate, biotinylated HKU1 isolate N1 RBD-loaded SA tips were dipped into 100 nM TMPRSS2 for 300 seconds followed by dissociation for 500 seconds. To facilitate visualization, the binding curves from point mutants are split into three groups and include wildtype (WT) TMPRSS2 for reference.



**Figure S6. Evaluation of HKU1 infection-elicited serum neutralizing activity in humans.** Dose-response curves for neutralization of VSV pseudotyped with wildtype isolate N1 HKU1 S (WT, red) or the N355 glycan knockout (S357A, blue) mutant by human serum antibodies, shown in red and blue, respectively. Visits 1 and 2 correspond to blood draws at the time of PCR-positive testing and 30 days later, respectively<sup>52,53</sup>.



**Table S1. Amino acid sequences of designed TMPRSS2 constructs.**

<p><b>Full length TMPRSS2 (S441A) <i>Mus musculus</i> Q3UKE3:</b></p> <p>MALNSGSPPGIGPCYENHGYQSEHICPPRPVAPNGYNLYPAQYYPSPVPQYAPRITTQASTSVIHTHPKSSGALCTSKSKKSLCL  ALALGTVLTGAAVAALLWRFWDSNCSTSEMECGSSGTCISSSLWCDGVAHCPNGEDENRCVRLYGQSFILQVYSSQRKAWYP  VCQDDWSESYGRAACKDMGYKNNFYSSQGIPDQSGATSFMKLNVSSGNVDLYKKLYHSDSCSSRMVSLRCIECGVRSVKRQS  RIVGGLNASPGDWPWQVSLHVQGVHVCGGSITPEWIVTAAHCVEEPLSSPRYWTAFAFILRQSLMFYGSRHQVEKVISHPNYDS  KTKNNDIALMKLQTLPLAFNDLVKPVCLPSPGMMLDLQECWISGWGATYEKGKTSVDLNAAMVPLIEPSKCNISKYIYNNLITPAMIC  AGFLQGSVDSCQGDSSGGLVTLKNGIWWLIGDTSWGS GCAKALRPGVYGNVTVFTDWIYQMRANSSGDYKDHDGDYKDHDID  YKDDDDK</p>
<p><b>Full length TMPRSS2 (S441A) <i>Ratus Norvegicus</i> Q6P7D7:</b></p> <p>MALNSGSPPGIGPYENHGYQSEHVYSPRPVSPSGYNLYPAQSCSPVPQYAPRVTTQASTPAIHQPRSSGLTCTSKSKKSMML  VALALGTVLAGAAVAAGLLWKFWDSCSSSEMECGSSGTCISSSLWCDGVSQCPNGEDENRCVRLYGTSFTLQVYSSQRKAWY  PVCQDDWNESYGRAACKDMGYKNSFYSSQGIPDQSGATSFMKLNVSAGNVDLYKKLYHSDSCSSRMVSLRCIECGVRSVRRQ  SRIVGGSTASPGDWPWQVSLHVQGIHVC GGSITPEWIVTAAHCVEEPLSSPRYWTAFAFILRQSLMFYGSRHQVEKVISHPNYDS  KTKNNDIALMKLQTLPLAFNDLVKPVCLPNPGMMLDLAQECWISGWGATYEKGKTSVDLNAAMVPLIEPSKCNISKYIYNNLITPAMI  CAGFLQGSVDSCQGDSSGGLVTLKNEIWWLIGDTSWGS GCAKAYRPGVYGNVTVFTDWIYQMRANSSGDYKDHDGDYKDHDID  DYKDDDDK</p>
<p><b>Full length TMPRSS2 (S441A) <i>Rhinolophus ferrumequinum</i> A0A671DXW4:</b></p> <p>MALNSGSPPGVGPYYENHGYQPESLYPPRPVAPSAVSYPARYYTPVVPQYTPRVMTHSTPAIHTQPKSPSGTLCTSKTKKAL  CITFSLGAVLVGAVAALLWKFMEKNCVSGIECGSSGTCVSASHWCDGILHCPSGEDENRCVRLYGNFILQVYSSQRKSWHP  VCQDDWSENYGRAACQDMGYRNSFYSSQGIVDDSGATSFMKLNVSAHGTDLYKKLYHSDVCSSTVSLRCIECGVNTKMSRQS  RIVGGTSAALGDWPWQVSLHVQGVHVC GGSITPEWIVTAAHCVEEPLNNPRYWTAFAFILRQSLMFYGSRYRVEKVISHPNYDS  KSKNNDIALMKLQTLPLTFNDKVKPVCLPNPGLMLEPTQSCWISGWGATYEKGKTSEVLNAVMPVPLIEPWQCNSKYVYNNLITPAMI  CAGYLQGTIDSCQGDSSGGLVTLKSSVWWLIGDTSWGS GCAKANRPGVYGNVTVFTDWIYRQMRANSSGDYKDHDGDYKDHDID  DYKDDDDK</p>
<p><b>Full length TMPRSS2 (S441A) <i>Chlorocebus sabaeus</i> green A0A0D9R7A1:</b></p> <p>MALNSGSPPGVGPYYENHGYQPENYPYPAQPTVAPNVYEVHPAQYYPSPVPQYIPSVLTHASNPAVRTQPKSPSGTVCTSKTKKA  LCVTMTLGAVLVGAALAAGLLWKFMSGKSDSGIECDSSGTCISSSNWCDGVSHCPSGEDENRCVRLYGNFILQVYSSQRKSW  HPVCRDDWNESYGRAACQDMGYQNRFYSSQGIVDDSGATSFMKLNVSAHGTDLYKKLYHSDVCSSTVSLRCIECGVNTKMSRQS  QSRIVGGQNALPGAWPWQVSLHVQNIHVC GGSITPEWIVTAAHCVKPLNSPWQWTAFAGILTQSSMFYVKGHRVEKVISHPNY  DSKTKNNDIALMKLHTPLTFNELVKPVCLPNPGMMLPEQHCWISGWGATQEKGKTSDDLNAAMVPLIEPRRCNSKRVYDGLVT  PAMICAGFLQGTVDSCQGDSSGGLVTLKNDVWWLIGDTSWGS GCAQANRPGVYGNVTVFTDWIYRQMRADSSGDYKDHDGDY  KDHDIDYKDDDDK</p>
<p><b>Full length TMPRSS2 (S441A) <i>Mesocricetus auratus</i> hamster A0A1U8C7X1:</b></p> <p>MALNSGSPPGIGPYENHGFQSEHIYPPRPVAPDVYNYPYPQNYPPVPQYFPRVTTQASTTVTHTQPHSSGKLTCTSTSKTKKS  LCFALSGLGIVLVGAAVAALLWKFPLPGCSTSEMECMSSGTCISSSLWCDGTHCPNGEDENRCVRLYGPSFTLQVYSSQRKAWY  PVCQDDWNDSYGRAACKDMGYKNNFYTTQGIPDSSGATSFMKLNVSAHGTDLYKKLYHSDVCSSTVSLRCIECGVNTKMSRQS  RIVGGSNASPGDWPWQVSLHVQGVHVC GGSITPEWIVTAAHCVEEPLNSPRYWTAFAGILSQSLMFYGSRHQVEKVISHPNYNS  ETKNNDIALMKLQTLPLTFNDLVKPVCLPNPGMMLDPAQECWISGWGSTYEKGKTSDDLNAAMVPLIERKCNKNKYIYNNLITPAMV  CAGFLQGTVDSCQGDSSGGLVTLKNDIWWLIGDTSWGS GCAKALRPGVYGNVTVFTDWIYQMMANSSGDYKDHDGDYKDHDID  DYKDDDDK</p>
<p><b>Full length TMPRSS2 (S441A) <i>Mustela furo ferret</i> A0A8U0SMZ2:</b></p> <p>MALNAGSPPGVGPYYENHGYQPESLYPAPPATVPSVYVAYPAAYYPVAVPQYTPRVLTQASTPAVRTQPKSPSGTACTAKAKKA  LCITISLGAVLAGAAVAVLLWKFMEKNCVSGIECGSSGTCISPSHWCDGVLHCPSGEDENRCVRLYGNFILQVYSAQRKSWH  PVCQDDWNDSYGRAACKDMGYRNSFYSSQGIVDDSGATSFMKLNVSAHGTDLYKKLYHSDVCSSTVSLRCIECGVAGKTMRO  SRIVGGSSASPGDWPWQVSLHVQGVHVC GGSITPEWIVTAAHCVEEPLNNPRYWTVFAGVLRQSLMFYGHGYRVGKVISHPSY  DSKTKNNDIALMKLQTLPLTFSDKVKPVCLPNPGMMLPEPNQSCWISGWGATHEKGKTSDELNAVMPVPLIEPWRCNSKYVYNSLVT  PAMICAGYLRGGTDSCQGDSSGGLVTLKSRIWWLIGDTSWGS GCAKANRPGVYGNVTVFTDWIYRQMRANSSGDYKDHDGDYK  DHDIDYKDDDDK</p>
<p><b>Full length TMPRSS2 (S441A) <i>Camelus dromedarius</i> XP_010993167.1:</b></p> <p>MALNSGSPPGVGPYYENHGYQPESFYPLKPSAASSAYMVYPAQYYPVAVPQYTPRVQTHSTPVIRMQPKPPSETVCTSKTKKV  LCVTGLGAILVGAVALAALLWKFESRCSASEMECGSSGTCISPSQWCDGVLHCPGGEENQCVRLYGNFILQVYSPQRKSW</p>

HPVCQEDWSESFGRAMCQDLGYGNSFYSSQGVVDDSGATSMFKLNISANNIDLYKKLYHSDVCSSKRVSRLRCIECGVSEKTSR  
QSRIVGGSSANLGDWPWQVSLHVQGIHVCGGSIITPEWIVTAAHCVEQPLSNAKIWTAFAGILSQSLMIYNGYQIAKVISHPNYDS  
KTKNNIDIALMKLQPLTFNDRVKPVCLPNPGMMLEATQSCWISGWGATYEKGKTSVDLNAVVMVHLIEPWKCNSKYVYNLITSAM  
ICAGYLQGGVDSCQGDGGPLVTLKNSVWWLLGDTSWGSGCAKAYRPGVYGNMTVFTDWIYRQMRANSSGDYKDHGDYKD  
HDIDYKDDDDK

**Full length TMPRSS2 (S441A) *Gallus domesticus* XP\_015156666.1:**

MTSTVNPPYYENHGFQTEENYSSARPQVGANPYPQYFSTNVPSVPTYIPRVSTHQSSIPVAPPSSSSRMCSSEIKKIVITLSILLVIC  
CAIAAFLIWYFVENRCLGSLIECGSSGVCISPSVWCDGVTDCPNGEDENRCVRLYGPNFILEVYSPVSQTWYPCQDDWTDDFGK  
IACEDMGYNVDYTYYSQGVAAEVFSKSFMKLNTSAGNTDLYKRLQSSDYCASGNVSLRCIECGLPTKSTAVMSRIVGGSMASLG  
QWPWPQVSLHVQDTHVCGGSIITREWLVTAAHCVEGLFSDPYIWSVYAGILSQNEMHSRPGYRVQKIISHPNYDTSKDNDVALMK  
LETPLSFTNTIRPVCLPNPGMMFQPNQQCWISGWGAEYQGGKTANDLNYVMVPLIERSTCNSVYVYDGMVLPMTVCAGYLQGGI  
DSCQGDGGPLVTNKNSVWWLVGDTSWGTCASPNNRPGVYGNMTVFTDWIYKNMQANRSGDYKDHGDYKDHIDYKDDDD  
K

**Full length TMPRSS2 (S441A) *Chamaeleo jacksonii* G1KE28:**

MNSRPPYYENYAYQENIAPPRHAGGYMYPPYPSPPYSSVPHYIPRVSTNQSTPVPTVQPKATPAKCMKTRKAVCLFLAIS  
VLLIGGAIAAVLIWHFVTDSCFGSKIKCGTTGMCVAPSQWCDGIRDCPNNEDETRCVRLFGEFQLEVYSSSEKDWYPVCSDDDWN  
DKHGKTACEDLGYSNTYFYSQTIPLMSASKGFMKLNASAGDIDLKLYKLYNSKSSRSLVSLQCIDCGTRRVNRRNRIVGGTSAS  
LGDWPWQVSLHSSGTHLCGGSITPEWIVTAAHCVEKAFSNPNYWTVFAGILTQPEMISSKGHKVAKVIPHPGYDTSSKTNDDVALM  
KLQSPLVFDEFVRPVCLPNPGMMFQSDQPYWISGWGAWEQKGTSSKLNAAIRLIDSDTCNNRYIYNGILPTMICAGYLNGGID  
SCQGDGGPLVTSKDSLWWLVGDTSWGTCATKYRPGVYGNMTVFTDWIYKNMQANRRSGDYKDHGDYKDHIDYKDDDD  
K

**Full length TMPRSS2 (S441A) *Xenopus laevis* XP\_018104413.1:**

MAGRDYRWKSPSPYFENYGFQDNNNAFSRPVQPNFYEAQPPRPQLSPAPHYIPQVSTIHSVPAINHEKSTWWTTPRRKKIA  
CIVAATSVLIALLIVGAVLCWYFVTMSCQMKCGTSGSCVRSIQWCDGVAQCPEGGEDESYCRMYPGDFQLQAYIPATSSWLSVCN  
ENWGDSQGRSVCQDMGYSTYVKSSVSATSATGYLKLNTSVVNGKLQSRITYKSSFCSTGVVTLRCIEGSSSTKNVENRIVGGS  
QASLGDWWPWQVSLQYNERHVCSSITSNYILTAHCVEGAYSSPYAWTVYVGSISRSTAGIRYVYKSVIGHQKYDTKTKNNDVAL  
MRLKISILFSSVTQPVCLPNAGMPWASGQSCWTSWGWGATYEGGTSSNVLNAAMVPLIDADTCNRPAVYNGAVTSTMICAGYLRG  
GIDSCQGDGGPLVTKNSLWWLVGDTSWGTCANVNKPGVYGNITEFLPWIFLQMPTYGSGDYKDHGDYKDHIDYKDDDD  
K

**Full length TMPRSS2 (S441) *Homo sapiens*:**

MALNSGSPPAIGPYENHGYQENPYPQAQPTVVPTVYEVHPAQYYPSPVPQYAPRVLTQASNPVCTQPKSPSGTVCTSKTKKA  
LCITLTGTLVGAALAGLLWKFMSGKCSNSGIECDSSGTCINPSNWCDGVSHCPGGEDENRCVRLYGPNFILQVYSSQKRKSWH  
PVCQDDWNENYGRAACRDMGYKNNFYSSQGIVDDSGSTSFMKLNTSAGNVDIYKKLYHSDACSSKAVVSLRCIACGVNLNDSRQ  
SRIVGGESALPGAWPWQVSLHVQNVHVCSSITPEWIVTAAHCVEKPLNNPWHWTAFAGILRQSFMYGAGYQVEKVISHPNYD  
SKTKNNDIALMKLQKPLTFNDLVKPVCLPNPGMMLQPEQLCWISGWGATEEKGTSEVLNAAKVLLIETQRCNSRYVYDNLITPAM  
ICAGFLQGNVDSCQGDGGPLVTSKNNIWWLIGDTSWGSGCAKAYRPGVYGNVMVFTDWIYRQMRADG

**TMPRSS2 ectodomain (S441A):**

MTRLTVLALLAGLLASSRASMGSKCSNSGIECDSSGTCINPSNWCDGVSHCPGGEDENRCVRLYGPNFILQVYSSQKRKSWHPVC  
QDDWNENYGRAACRDMGYKNNFYSSQGIVDDSGSTSFMKLNTSAGNVDIYKKLYHSDACSSKAVVSLRCIACGVNLNDDDDKIV  
GGESALPGAWPWQVSLHVQNVHVCSSITPEWIVTAAHCVEKPLNNPWHWTAFAGILRQSFMYGAGYQVEKVISHPNYDSKT  
KNNDIALMKLQKPLTFNDLVKPVCLPNPGMMLQPEQLCWISGWGATEEKGTSEVLNAAKVLLIETQRCNSRYVYDNLITPAMICA  
GFLQGNVDSCQGDAGGPLVTSKNNIWWLIGDTSWGSGCAKAYRPGVYGNVMVFTDWIYRQMRADGDDDDKSGHHHHHHHH

**TMPRSS2 ectodomain (S441A + DS):**

MTRLTVLALLAGLLASSRASMGSKCSNSGIECDSSGTCINPSNWCDGVSHCPGGEDENRCVRLYGPNFILQVYSSQKRKSWHPVC  
QDDWNENYGRAACRDMGYKNNFYSSQGIVDDSGSTSFMKLNTSAGNVDIYKKLYHSDACSSKAVVSLRCIACGVNLNDDDDKIV  
GGESALPGAWPWQVSLHVQNVHVCSSITPEWIVTAAHCVEKPLNNPWHWTAFAGILRQSFMYGAGYQVEKVISHPNYDSKT  
KNNDIALMKLQKPLTFNDLVKPVCLPNPGMMLQPEQLCWISGWGATEEKGTSEVLNAAKVLLIETQRCNSRYVYDNLITPAMICA  
GFLQGNVDSCQGDAGGPLVTSKNNIWWLIGDTSWGSGCAKAYRPGVYGNVMVFTDWIYRQMRADGDDDDKSGHHHHHHHH

**TMPRSS2 ectodomain (S441 + DS):**

MTRLTVLALLAGLLASSRASMGSKCSNSGIECDSSGTCINPSNWCDGVSHCPGGEDENRCVRLYGPNFILQVYSSQKRKSWHPVC  
QDDWNENYGRAACRDMGYKNNFYSSQGIVDDSGSTSFMKLNTSAGNVDIYKKLYHSDACSSKAVVSLRCIACGVNLNDDDDKIV  
GGESALPGAWPWQVSLHVQNVHVCSSITPEWIVTAAHCVEKPLNNPWHWTAFAGILRQSFMYGAGYQVEKVISHPNYDSKT

KNNDIALMKLQKPLTFNDLVKPVCLPNPGMMLQPEQLCWISGWGATEEKGKTSEVLNAAKVLLIETQRCNSRYVYDNLITPAMICA  
GFLQGNVDSCQGDGGPLVCSKNNIWWLIGDTSWGS GCAKAYRPGVYGNVMVFTDWIYRQMRADGDDDDKSGHHHHHHHH

**TMPRSS2 ectodomain (S441A + DS + SUMO):**

MTRLTVLALLAGLLASSRASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFARQKGKEMDSLRFY  
DGIRIQADQTPEDLDMEDNDIEAHRENDDDDKTTGSKCSNSGIECDSSGTCINPSNWCDGVSHCPGGEDENRCVRLYGPNFILQ  
VYSSQRKSWHPVCQDDWNENYGRAACRDMGYKNNFYSSQGIVDDSGSTSFMKLNTSAGNVDIYKKLYHSDACSSKAVVSLRCI  
ACGVNLNDDDDKIVGGESALPGAWPWQVSLHVQNVHVCGGSIITPEWIVTAAHCVEKPLNNPWHWTAFAGILRQSFMYGAGYQ  
VEKVISHPNYDSKTKNNDIALMKLQKPLTFNDLVKPVCLPNPGMMLQPEQLCWISGWGATEEKGKTSEVLNAAKVLLIETQRCNSR  
YVYDNLITPAMICAGFLQGNVDSCQGDAGGPLVCSKNNIWWLIGDTSWGS GCAKAYRPGVYGNVMVFTDWIYRQMRADGDDDD  
KSGHHHHHHHH

**TMPRSS2 ectodomain (S441 + DS + SUMO):**

MTRLTVLALLAGLLASSRASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFARQKGKEMDSLRFY  
DGIRIQADQTPEDLDMEDNDIEAHRENDDDDKTTGSKCSNSGIECDSSGTCINPSNWCDGVSHCPGGEDENRCVRLYGPNFILQ  
VYSSQRKSWHPVCQDDWNENYGRAACRDMGYKNNFYSSQGIVDDSGSTSFMKLNTSAGNVDIYKKLYHSDACSSKAVVSLRCI  
ACGVNLNDDDDKIVGGESALPGAWPWQVSLHVQNVHVCGGSIITPEWIVTAAHCVEKPLNNPWHWTAFAGILRQSFMYGAGYQ  
VEKVISHPNYDSKTKNNDIALMKLQKPLTFNDLVKPVCLPNPGMMLQPEQLCWISGWGATEEKGKTSEVLNAAKVLLIETQRCNSR  
YVYDNLITPAMICAGFLQGNVDSCQGDAGGPLVCSKNNIWWLIGDTSWGS GCAKAYRPGVYGNVMVFTDWIYRQMRADGDDDD  
KSGHHHHHHHH

**TMPRSS2 ectodomain (S441A + DS + SUMO + N249 glycan):**

MTRLTVLALLAGLLASSRASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFARQKGKEMDSLRFY  
DGIRIQADQTPEDLDMEDNDIEAHRENDDDDKTTGSKCSNSGIECDSSGTCINPSNWCDGVSHCPGGEDENRCVRLYGPNFILQ  
VYSSQRKSWHPVCQDDWNENYGRAACRDMGYKNNFYSSQGIVDDSGSTSFMKLNTSAGNVDIYKKLYHSDACSSKAVVSLRCI  
ACGVNLNSSDDDDKIVGGESALPGAWPWQVSLHVQNVHVCGGSIITPEWIVTAAHCVEKPLNNPWHWTAFAGILRQSFMYGAG  
YQVEKVISHPNYDSKTKNNDIALMKLQKPLTFNDLVKPVCLPNPGMMLQPEQLCWISGWGATEEKGKTSEVLNAAKVLLIETQRCN  
SRYVYDNLITPAMICAGFLQGNVDSCQGDAGGPLVCSKNNIWWLIGDTSWGS GCAKAYRPGVYGNVMVFTDWIYRQMRADGDD  
DDKSGHHHHHHHH

**TMPRSS2 ectodomain (S441 + DS + SUMO + N249 glycan):**

MTRLTVLALLAGLLASSRASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFARQKGKEMDSLRFY  
DGIRIQADQTPEDLDMEDNDIEAHRENDDDDKTTGSKCSNSGIECDSSGTCINPSNWCDGVSHCPGGEDENRCVRLYGPNFILQ  
VYSSQRKSWHPVCQDDWNENYGRAACRDMGYKNNFYSSQGIVDDSGSTSFMKLNTSAGNVDIYKKLYHSDACSSKAVVSLRCI  
ACGVNLNSSDDDDKIVGGESALPGAWPWQVSLHVQNVHVCGGSIITPEWIVTAAHCVEKPLNNPWHWTAFAGILRQSFMYGAG  
YQVEKVISHPNYDSKTKNNDIALMKLQKPLTFNDLVKPVCLPNPGMMLQPEQLCWISGWGATEEKGKTSEVLNAAKVLLIETQRCN  
SRYVYDNLITPAMICAGFLQGNVDSCQGDAGGPLVCSKNNIWWLIGDTSWGS GCAKAYRPGVYGNVMVFTDWIYRQMRADGSG  
LNDIFEAQKIEWHEQSGHHHHHHHH

**Table S2. Cryo-EM data collection, refinement and validation statistics**

	HKU1 RBD - TMPRSS2
	PDB 8VGT
	EMD-43224
<b>Data collection and processing</b>	
Magnification	105,000
Voltage (kV)	300
Electron exposure (e <sup>-</sup> /Å <sup>2</sup> )	60
Defocus range (μm)	-0.2 - -3.5
Pixel size (Å)	0.835
Symmetry imposed	C1
Initial particle images (no.)	8,831,644
Final particle images (no.)	810,357
Map resolution (Å)	2.9
FSC threshold	0.143
<b>Refinement</b>	
Model resolution (Å)	3.1
FSC threshold	0.5
Map sharpening <i>B</i> factor (Å <sup>2</sup> )	-130
<b>Model composition</b>	
Non-hydrogen atoms	5029
Protein residues	637
Ligands	5
<b><i>B</i> factors (Å<sup>2</sup>)</b>	
Protein	16.34
Ligand	28.76
<b>R.m.s. deviations</b>	
Bond lengths (Å)	0.011
Bond angles (°)	1.219
<b>Validation</b>	
MolProbity score	1.06
Clashscore	0.82
Poor rotamers (%)	0.71
<b>Ramachandran plot</b>	
Favored (%)	95.56
Allowed (%)	4.12
Disallowed (%)	0.32

**Data S1. Human TMPRSS2 SNPs at the HKU1 RBD-interacting site present in the GnomADv4 database.**

**Data S2. Human TMPRSS2 SNPs at the HKU1 RBD-interacting site present in the Regeneron Million Exome Variant database.**

### Data S3. Demographic data for HKU1 infected serum donors

Manuscript code	Country of isolation	Age	Gender	Collection date V1	Collection date V2
HK-01	Netherlands	47	2	2/3/10	3/5/10
HK-02	Spain	36	2	3/16/10	4/20/10
HK-03	Spain	27	2	3/5/10	4/26/10
HK-04	Spain	38	2	4/8/10	5/12/10
HK-05	Poland	56	1	3/8/10	4/16/10
HK-06	Poland	54	2	1/21/10	2/19/10
HK-07	Poland	48	1	3/3/10	4/2/10
HK-08	Poland	28	2	2/23/10	4/19/10
HK-09	Poland	35	1	3/4/10	4/26/10
HK-10	Poland	47	1	2/10/10	3/15/10
HK-11	Poland	26	2	2/8/10	3/11/10
HK-12	United Kingdom	53	2	12/11/07	1/18/08
HK-13	Spain	27	2	2/16/09	4/16/09

### References

1. Woo, P.C.Y., Lau, S.K.P., Chu, C.-M., Chan, K.-H., Tsoi, H.-W., Huang, Y., Wong, B.H.L., Poon, R.W.S., 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.
2. Góes, L.G., Durigon, E.L., Campos, A.A., Hein, N., Passos, S.D., and Jerez, J.A. (2011). Coronavirus HKU1 in Children, Brazil, 1995. *Emerging Infectious Disease journal* 17, 1147.
3. Esper, F., Weibel, C., Ferguson, D., Landry, M.L., and Kahn, J.S. (2006). Coronavirus HKU1 infection in the United States. *Emerg. Infect. Dis.* 12, 775–779.
4. Ruohola, A., Waris, M., Allander, T., Ziegler, T., Heikkinen, T., and Ruuskanen, O. (2009). Viral etiology of common cold in children, Finland. *Emerg. Infect. Dis.* 15, 344–346.
5. Woo, P.C.Y., Lau, S.K.P., Yip, C.C.Y., Huang, Y., Tsoi, H.-W., Chan, K.-H., and Yuen, K.-Y. (2006). Comparative analysis of 22 coronavirus HKU1 genomes reveals a novel genotype and evidence of natural recombination in coronavirus HKU1. *J. Virol.* 80, 7136–7145.
6. Sloots, T.P., McErlean, P., Speicher, D.J., Arden, K.E., Nissen, M.D., and Mackay, I.M. (2006). Evidence of human coronavirus HKU1 and human bocavirus in Australian children. *J. Clin. Virol.* 35, 99–102.
7. Zeng, Z.-Q., Chen, D.-H., Tan, W.-P., Qiu, S.-Y., Xu, D., Liang, H.-X., Chen, M.-X., Li, X.,



- Lin, Z.-S., Liu, W.-K., et al. (2018). Epidemiology and clinical characteristics of human coronaviruses OC43, 229E, NL63, and HKU1: a study of hospitalized children with acute respiratory tract infection in Guangzhou, China. *Eur. J. Clin. Microbiol. Infect. Dis.* 37, 363–369.
8. Frutos, A.M., Balmaseda, A., Vydiswaran, N., Patel, M., Ojeda, S., Brouwer, A., Tutino, R., Cai, S., Bakker, K., Sanchez, N., et al. (2023). Burden and seasonality of primary and secondary symptomatic common cold coronavirus infections in Nicaraguan children. *Influenza Other Respi. Viruses* 17, e13078.
9. Woo, P.C.Y., Lau, S.K.P., Tsoi, H.-W., Huang, Y., Poon, R.W.S., Chu, C.-M., Lee, R.A., Luk, W.-K., Wong, G.K.M., Wong, B.H.L., et al. (2005). Clinical and Molecular Epidemiological Features of Coronavirus HKU1–Associated Community-Acquired Pneumonia. *J. Infect. Dis.* 192, 1898–1907.
10. Tortorici, M.A., and Veasler, D. (2019). Structural insights into coronavirus entry. *Adv. Virus Res.* 105, 93–116.
11. Li, W., Moore, M.J., Vasilieva, N., Sui, J., Wong, S.K., Berne, M.A., Somasundaran, M., Sullivan, J.L., Luzuriaga, K., Greenough, T.C., et al. (2003). Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* 426, 450–454.
12. Walls, A.C., Park, Y.J., Tortorici, M.A., Wall, A., McGuire, A.T., and Veasler, D. (2020). Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* 181, 281–292.e6.
13. Letko, M., Marzi, A., and Munster, V. (2020). Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat Microbiol* 5, 562–569.
14. Zhou, P., Yang, X.L., Wang, X.G., Hu, B., Zhang, L., Zhang, W., Si, H.R., Zhu, Y., Li, B., Huang, C.L., et al. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 10.1038/s41586-020-2012-7.
15. Hofmann, H., Pyrc, K., van der Hoek, L., Geier, M., Berkhout, B., and Pohlmann, S. (2005). Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry. *Proc. Natl. Acad. Sci. U. S. A.* 102, 7988–7993.
16. Xiong, Q., Cao, L., Ma, C., Tortorici, M.A., Liu, C., Si, J., Liu, P., Gu, M., Walls, A.C., Wang, C., et al. (2022). Close relatives of MERS-CoV in bats use ACE2 as their functional receptors. *Nature* 612, 748–757.
17. Yeager, C.L., Ashmun, R.A., Williams, R.K., Cardellicchio, 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.
18. Tortorici, M.A., Walls, A.C., Joshi, A., Park, Y.-J., Eguia, R.T., Miranda, M.C., Kepl, E., Dosey, A., Stevens-Ayers, T., Boeckh, M.J., et al. (2022). Structure, receptor recognition, and antigenicity of the human coronavirus CCoV-HuPn-2018 spike glycoprotein. *Cell* 185, 2279–2291.e17.
19. Li, W., Hulswit, R.J.G., Kenney, S.P., Widjaja, I., Jung, K., Alhamo, M.A., van Dieren, B.,

- van Kuppeveld, F.J.M., Saif, L.J., and Bosch, B.J. (2018). Broad receptor engagement of an emerging global coronavirus may potentiate its diverse cross-species transmissibility. *Proc. Natl. Acad. Sci. U. S. A.* *115*, E5135–E5143.
20. Delmas, B., Gelfi, J., L'Haridon, R., Vogel, Sjöström, H., Norén, and Laude, H. (1992). Aminopeptidase N is a major receptor for the enteropathogenic coronavirus TGEV. *Nature* *357*, 417–420.
21. Raj, V.S., Mou, H., Smits, S.L., Dekkers, D.H.W., Müller, M.A., Dijkman, R., Muth, D., Demmers, J.A.A., Zaki, A., Fouchier, R.A.M., et al. (2013). Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. *Nature* *495*, 251–254.
22. Yang, Y., Du, L., Liu, C., Wang, L., Ma, C., Tang, J., Baric, R.S., Jiang, S., and Li, F. (2014). Receptor usage and cell entry of bat coronavirus HKU4 provide insight into bat-to-human transmission of MERS coronavirus. *Proc. Natl. Acad. Sci. U. S. A.* *111*, 12516–12521.
23. Wang, Q., Qi, J., Yuan, Y., Xuan, Y., Han, P., Wan, Y., Ji, W., Li, Y., Wu, Y., Wang, J., et al. (2014). Bat origins of MERS-CoV supported by bat coronavirus HKU4 usage of human receptor CD26. *Cell Host Microbe* *16*, 328–337.
24. Chen, J., Yang, X., Si, H., Gong, Q., Que, T., Li, J., Li, Y., Wu, C., Zhang, W., Chen, Y., et al. (2023). A bat MERS-like coronavirus circulates in pangolins and utilizes human DPP4 and host proteases for cell entry. *Cell* *186*, 850–863.e16.
25. Dveksler, G.S., Pensiero, M.N., Cardellichio, C.B., Williams, R.K., Jiang, G.S., Holmes, K.V., and Dieffenbach, C.W. (1991). Cloning of the mouse hepatitis virus (MHV) receptor: expression in human and hamster cell lines confers susceptibility to MHV. *J. Virol.* *65*, 6881–6891.
26. Tortorici, M.A., Walls, A.C., Lang, Y., Wang, C., Li, Z., Koerhuis, D., Boons, G.J., Bosch, B.J., Rey, F.A., de Groot, R.J., et al. (2019). Structural basis for human coronavirus attachment to sialic acid receptors. *Nat. Struct. Mol. Biol.* *26*, 481–489.
27. 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.* *10.1073/pnas.1809667116*.
28. Vlasak, R., Luytjes, W., Spaan, W., and Palese, P. (1988). Human and bovine coronaviruses recognize sialic acid-containing receptors similar to those of influenza C viruses. *Proc. Natl. Acad. Sci. U. S. A.* *85*, 4526–4529.
29. 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*. *10.1038/s41586-023-06761-7*.
30. Pronker, M.F., Creutzmacher, 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*. *10.1038/s41586-023-06599-z*.
31. 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.
32. Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, 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.e8.
33. Zang, R., Gomez Castro, M.F., McCune, B.T., Zeng, Q., Rothlauf, P.W., Sonnek, N.M., Liu, Z., Brulois, K.F., Wang, X., Greenberg, H.B., et al. (2020). TMPRSS2 and TMPRSS4 promote SARS-CoV-2 infection of human small intestinal enterocytes. *Sci Immunol* 5. 10.1126/sciimmunol.abc3582.
34. Glowacka, I., Bertram, S., Müller, 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.
35. Heurich, A., Hofmann-Winkler, H., Gierer, S., Liepold, T., Jahn, O., and Pöhlmann, S. (2014). TMPRSS2 and ADAM17 cleave ACE2 differentially and only proteolysis by TMPRSS2 augments entry driven by the severe acute respiratory syndrome coronavirus spike protein. *J. Virol.* 88, 1293–1307.
36. 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.
37. Gierer, S., Bertram, S., Kaup, F., Wrensch, F., Heurich, A., Krämer-Kühl, A., Welsch, K., Winkler, M., Meyer, B., Drosten, C., et al. (2013). The spike protein of the emerging betacoronavirus EMC uses a novel coronavirus receptor for entry, can be activated by TMPRSS2, and is targeted by neutralizing antibodies. *J. Virol.* 87, 5502–5511.
38. 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.
39. Limburg, H., Harbig, A., Bestle, D., Stein, D.A., Moulton, H.M., Jaeger, J., Janga, H., Hards, 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.
40. Böttcher, E., Matrosovich, T., Beyerle, M., Klenk, H.-D., Garten, W., and Matrosovich, M. (2006). Proteolytic activation of influenza viruses by serine proteases TMPRSS2 and HAT from human airway epithelium. *J. Virol.* 80, 9896–9898.
41. Shapira, T., Monreal, I.A., Dion, S.P., Buchholz, D.W., Imbiakha, B., Olmstead, A.D., Jager, M., Désilets, A., Gao, G., Martins, M., et al. (2022). A TMPRSS2 inhibitor acts as a pan-SARS-CoV-2 prophylactic and therapeutic. *Nature* 605, 340–348.
42. Hamilton, B.S., Chung, C., Cyphers, S.Y., Rinaldi, V.D., Marcano, V.C., and Whittaker, G.R. (2014). Inhibition of influenza virus infection and hemagglutinin cleavage by the

protease inhibitor HAI-2. *Biochem. Biophys. Res. Commun.* **450**, 1070–1075.

43. Mahoney, M., Damalanka, V.C., Tartell, M.A., Chung, D.H., Lourenço, A.L., Pwee, D., Mayer Bridwell, A.E., Hoffmann, M., Voss, J., Karmakar, P., et al. (2021). A novel class of TMPRSS2 inhibitors potently block SARS-CoV-2 and MERS-CoV viral entry and protect human epithelial lung cells. *Proc. Natl. Acad. Sci. U. S. A.* **118**. 10.1073/pnas.2108728118.
44. 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.
45. Ou, X., Guan, H., Qin, B., Mu, Z., Wojdyla, J.A., Wang, M., Dominguez, S.R., Qian, Z., and Cui, S. (2017). Crystal structure of the receptor binding domain of the spike glycoprotein of human betacoronavirus HKU1. *Nat. Commun.* **8**, 15216.
46. 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.
47. Dauparas, J., Anishchenko, I., Bennett, N., Bai, H., Ragotte, R.J., Milles, L.F., Wicky, B.I.M., Courbet, A., de Haas, R.J., Bethel, N., et al. (2022). Robust deep learning-based protein sequence design using ProteinMPNN. *Science* **378**, 49–56.
48. Vernon, R.M., Chong, P.A., Tsang, B., Kim, T.H., Bah, A., Farber, P., Lin, H., and Forman-Kay, J.D. (2018). Pi-Pi contacts are an overlooked protein feature relevant to phase separation. *Elife* **7**. 10.7554/eLife.31486.
49. 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. (2019). Unexpected Receptor Functional Mimicry Elucidates Activation of Coronavirus Fusion. *Cell* **176**, 1026–1039.e15.
50. Walls, A.C., Tortorici, M.A., Snijder, J., Xiong, X., Bosch, B.J., Rey, F.A., and Veasler, D. (2017). Tectonic conformational changes of a coronavirus spike glycoprotein promote membrane fusion. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 11157–11162.
51. Walls, A.C., Tortorici, M.A., Bosch, B.J., Frenz, B., Rottier, P.J.M., DiMaio, F., Rey, F.A., and Veasler, D. (2016). Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer. *Nature* **531**, 114–117.
52. Sechan, F., Grobbs, M., Edridge, A.W.D., Jebbink, M.F., Loens, K., Ieven, M., Goossens, H., van Hemert-Glaubitz, S., van Gils, M.J., and van der Hoek, L. (2022). Atypical Antibody Dynamics During Human Coronavirus HKU1 Infections. *Front. Microbiol.* **13**, 853410.
53. Ieven, M., Coenen, S., Loens, K., Lammens, C., Coenjaerts, F., Vanderstraeten, A., Henriques-Normark, B., Crook, D., Huygen, K., Butler, C.C., et al. (2018). Aetiology of lower respiratory tract infection in adults in primary care: a prospective study in 11 European countries. *Clin. Microbiol. Infect.* **24**, 1158–1163.
54. Tortorici, M.A., Beltramello, M., Lempp, F.A., Pinto, D., Dang, H.V., Rosen, L.E., McCallum, M., Bowen, J., Minola, A., Jaconi, S., et al. (2020). Ultrapotent human antibodies protect against SARS-CoV-2 challenge via multiple mechanisms. *Science* **370**, 950–957.

55. Piccoli, L., Park, Y.J., Tortorici, M.A., Czudnochowski, N., Walls, A.C., Beltramello, M., Silacci-Fregni, C., Pinto, D., Rosen, L.E., Bowen, J.E., et al. (2020). Mapping Neutralizing and Immunodominant Sites on the SARS-CoV-2 Spike Receptor-Binding Domain by Structure-Guided High-Resolution Serology. *Cell* 183, 1024–1042.e21.
56. Park, Y.-J., De Marco, A., Starr, T.N., Liu, Z., Pinto, D., Walls, A.C., Zatta, F., Zepeda, S.K., Bowen, J.E., Sprouse, K.R., et al. (2022). Antibody-mediated broad sarbecovirus neutralization through ACE2 molecular mimicry. *Science*, eabm8143.
57. Corti, D., Zhao, J., Pedotti, M., Simonelli, L., Agnihothram, S., Fett, C., Fernandez-Rodriguez, B., Foglierini, M., Agatic, G., Vanzetta, F., et al. (2015). Prophylactic and postexposure efficacy of a potent human monoclonal antibody against MERS coronavirus. *Proc. Natl. Acad. Sci. U. S. A.* 112, 10473–10478.
58. Traggiai, E., Becker, S., Subbarao, K., Kolesnikova, L., Uematsu, Y., Gismondo, M.R., Murphy, B.R., Rappuoli, R., and Lanzavecchia, A. (2004). An efficient method to make human monoclonal antibodies from memory B cells: potent neutralization of SARS coronavirus. *Nat. Med.* 10, 871–875.
59. Park, Y.-J., Pinto, D., Walls, A.C., Liu, Z., De Marco, A., Benigni, F., Zatta, F., Silacci-Fregni, C., Bassi, J., Sprouse, K.R., et al. (2022). Imprinted antibody responses against SARS-CoV-2 Omicron sublineages. *Science*, eadc9127.
60. Starr, T.N., Czudnochowski, N., Liu, Z., Zatta, F., Park, Y.-J., Addetia, A., Pinto, D., Beltramello, M., Hernandez, P., Greaney, A.J., et al. (2021). SARS-CoV-2 RBD antibodies that maximize breadth and resistance to escape. *Nature*. 10.1038/s41586-021-03807-6.
61. Starr, T.N., Zepeda, S.K., Walls, A.C., Greaney, A.J., Alkhovsky, S., Veessler, D., and Bloom, J.D. (2022). ACE2 binding is an ancestral and evolvable trait of sarbecoviruses. *Nature* 603, 913–918.
62. Edridge, A.W.D., Kaczorowska, J.M., Hoste, A.C.R., Bakker, M., Klein, M., Jebbink, M.F., Matser, A., Kinsella, C., Rueda, P., Prins, M., et al. (2020). Coronavirus protective immunity is short-lasting. *medRxiv*, 2020.05.11.20086439.
63. Eguia, R.T., Crawford, K.H.D., Stevens-Ayers, T., Kelnhofer-Millevolte, L., Greninger, A.L., Englund, J.A., Boeckh, M.J., and Bloom, J.D. (2021). A human coronavirus evolves antigenically to escape antibody immunity. *PLoS Pathog.* 17, e1009453.
64. Bowen, J.E., Park, Y.-J., Stewart, C., Brown, J.T., Sharkey, W.K., Walls, A.C., Joshi, A., Sprouse, K.R., McCallum, M., Tortorici, M.A., et al. (2022). SARS-CoV-2 spike conformation determines plasma neutralizing activity elicited by a wide panel of human vaccines. *Sci Immunol* 7, eadf1421.
65. Starr, T.N., Greaney, A.J., Hilton, S.K., Ellis, D., Crawford, K.H.D., Dingens, A.S., Navarro, M.J., Bowen, J.E., Tortorici, M.A., Walls, A.C., et al. (2020). Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* 182, 1295–1310.e20.
66. McCallum, M., Czudnochowski, N., Rosen, L.E., Zepeda, S.K., Bowen, J.E., Walls, A.C., Hauser, K., Joshi, A., Stewart, C., Dillen, J.R., et al. (2022). Structural basis of SARS-CoV-2 Omicron immune evasion and receptor engagement. *Science*, eabn8652.



67. Wong, A.H.M., Tomlinson, A.C.A., Zhou, D., Satkunarajah, M., Chen, K., Sharon, C., Desforges, M., Talbot, P.J., and Rini, J.M. (2017). Receptor-binding loops in alphacoronavirus adaptation and evolution. *Nat. Commun.* 8, 1735.
68. Wang, C., Hesketh, E.L., Shamorkina, T.M., Li, W., Franken, P.J., Drabek, D., van Haperen, R., Townend, S., van Kuppeveld, F.J.M., Grosveld, F., et al. (2022). Antigenic structure of the human coronavirus OC43 spike reveals exposed and occluded neutralizing epitopes. *Nat. Commun.* 13, 2921.
69. 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.
70. 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.
71. Yuan, Y., Cao, D., Zhang, Y., Ma, J., Qi, J., Wang, Q., Lu, G., Wu, Y., Yan, J., Shi, Y., et al. (2017). Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. *Nat. Commun.* 8, 15092.
72. Pallesen, J., Wang, N., Corbett, K.S., Wrapp, D., Kirchdoerfer, R.N., Turner, H.L., Cottrell, C.A., Becker, M.M., Wang, L., Shi, W., et al. (2017). Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. *Proc. Natl. Acad. Sci. U. S. A.* 114, E7348–E7357.
73. Lee, J., Zepeda, S.K., Park, Y.-J., Taylor, A.L., Quispe, J., Stewart, C., Leaf, E.M., Treichel, C., Corti, D., King, N.P., et al. (2023). Broad receptor tropism and immunogenicity of a clade 3 sarbecovirus. *Cell Host Microbe.* 10.1016/j.chom.2023.10.018.
74. 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.
75. Walls, A.C., Tortorici, M.A., Frenz, B., Snijder, J., Li, W., Rey, F.A., DiMaio, F., Bosch, B.-J., and Veisler, D. (2016). Glycan shield and epitope masking of a coronavirus spike protein observed by cryo-electron microscopy. *Nat. Struct. Mol. Biol.* 23, 899–905.
76. Xiong, X., Tortorici, M.A., Snijder, J., Yoshioka, C., Walls, A.C., Li, W., McGuire, A.T., Rey, F.A., Bosch, B.J., and Veisler, D. (2018). Glycan Shield and Fusion Activation of a Deltacoronavirus Spike Glycoprotein Fine-Tuned for Enteric Infections. *J. Virol.* 92. 10.1128/JVI.01628-17.
77. Shang, J., Zheng, Y., Yang, Y., Liu, C., Geng, Q., Tai, W., Du, L., Zhou, Y., Zhang, W., and Li, F. (2018). Cryo-electron microscopy structure of porcine Deltacoronavirus spike protein in the prefusion state. *J. Virol.* 92. 10.1128/JVI.01556-17.
78. Shang, J., Zheng, Y., Yang, Y., Liu, C., Geng, Q., Luo, C., Zhang, W., and Li, F. (2018). Cryo-EM structure of infectious bronchitis coronavirus spike protein reveals structural and functional evolution of coronavirus spike proteins. *PLoS Pathog.* 14, e1007009.

79. Li, Z., Tomlinson, A.C., Wong, A.H., Zhou, D., Desforages, M., Talbot, P.J., Benlekbi, S., Rubinstein, J.L., and Rini, J.M. (2019). The human coronavirus HCoV-229E S-protein structure and receptor binding. *Elife* 8. 10.7554/eLife.51230.
80. Lau, S.K.P., Woo, P.C.Y., Li, K.S.M., Tsang, A.K.L., Fan, R.Y.Y., Luk, H.K.H., Cai, J.-P., Chan, K.-H., Zheng, B.-J., Wang, M., et al. (2015). Discovery of a novel coronavirus, China Rattus coronavirus HKU24, from Norway rats supports the murine origin of Betacoronavirus 1 and has implications for the ancestor of Betacoronavirus lineage A. *J. Virol.* 89, 3076–3092.
81. Addetia, A., Piccoli, L., Case, J.B., Park, Y.-J., Beltramello, M., Guarino, B., Dang, H., de Melo, G.D., Pinto, D., Sprouse, K., et al. (2023). Neutralization, effector function and immune imprinting of Omicron variants. *Nature* 621, 592–601.
82. McCallum, M., De Marco, A., Lempp, F.A., Tortorici, M.A., Pinto, D., Walls, A.C., Beltramello, M., Chen, A., Liu, Z., Zatta, F., et al. (2021). N-terminal domain antigenic mapping reveals a site of vulnerability for SARS-CoV-2. *Cell* 184, 2332–2347.e16.
83. Cheng, J., Novati, G., Pan, J., Bycroft, C., Žemgulytė, A., Applebaum, T., Pritzel, A., Wong, L.H., Zielinski, M., Sargeant, T., et al. (2023). Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science* 381, eadg7492.
84. Russo, C.J., and Passmore, L.A. (2014). Electron microscopy: Ultrastable gold substrates for electron cryomicroscopy. *Science* 346, 1377–1380.
85. Tan, Y.Z., Baldwin, P.R., Davis, J.H., Williamson, J.R., Potter, C.S., Carragher, B., and Lyumkis, D. (2017). Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat. Methods* 14, 793–796.
86. Suloway, C., Pulokas, J., Fellmann, D., Cheng, A., Guerra, F., Quispe, J., Stagg, S., Potter, C.S., and Carragher, B. (2005). Automated molecular microscopy: the new Legimon system. *J. Struct. Biol.* 151, 41–60.
87. Tegunov, D., and Cramer, P. (2019). Real-time cryo-electron microscopy data preprocessing with Warp. *Nat. Methods* 16, 1146–1152.
88. Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* 14, 290–296.
89. Bepler, T., Morin, A., Rapp, M., Brasch, J., Shapiro, L., Noble, A.J., and Berger, B. (2019). Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat. Methods* 16, 1153–1160.
90. Punjani, A., Zhang, H., and Fleet, D.J. (2020). Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat. Methods* 17, 1214–1221.
91. Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., and Scheres, S.H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* 7. 10.7554/eLife.42166.

92. Rosenthal, P.B., and Henderson, R. (2003). Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* 333, 721–745.
93. Chen, S., McMullan, G., Faruqi, A.R., Murshudov, G.N., Short, J.M., Scheres, S.H., and Henderson, R. (2013). High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy* 135, 24–35.
94. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612.
95. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66, 486–501.
96. Frenz, B., Rämisch, S., Borst, A.J., Walls, A.C., Adolf-Bryfogle, J., Schief, W.R., Veesler, D., and DiMaio, F. (2019). Automatically Fixing Errors in Glycoprotein Structures with Rosetta. *Structure* 27, 134–139.e3.
97. Wang, R.Y., Song, Y., Barad, B.A., Cheng, Y., Fraser, J.S., and DiMaio, F. (2016). Automated structure refinement of macromolecular assemblies from cryo-EM maps using Rosetta. *Elife* 5. 10.7554/eLife.17219.
98. Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 66, 12–21.
99. Liebschner, D., Afonine, P.V., Baker, M.L., Bunkóczi, 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.
100. Agirre, J., Iglesias-Fernández, J., Rovira, C., Davies, G.J., Wilson, K.S., and Cowtan, K.D. (2015). Privateer: software for the conformational validation of carbohydrate structures. *Nat. Struct. Mol. Biol.* 22, 833–834.