

# Alpha-1 antitrypsin inhibits TMPRSS2 protease activity and SARS-CoV-2 infection

Lukas Wettstein <sup>1,16</sup>, Tatjana Weil <sup>1,16</sup>, Carina Conzelmann<sup>1,16</sup>, Janis A. Müller <sup>1</sup>, Rüdiger Groß <sup>1</sup>, Maximilian Hirschenberger<sup>1</sup>, Alina Seidel<sup>1</sup>, Susanne Klute<sup>1</sup>, Fabian Zech<sup>1</sup>, Caterina Prelli Bozzo<sup>1</sup>, Nico Preising<sup>2</sup>, Giorgio Fois <sup>3</sup>, Robin Lochbaum<sup>3</sup>, Philip Maximilian Knaff<sup>4,5</sup>, Volker Mailänder<sup>4,5</sup>, Ludger Ständker <sup>2</sup>, Dietmar Rudolf Thal <sup>6,7</sup>, Christian Schumann <sup>8</sup>, Steffen Stenger<sup>9</sup>, Alexander Kleger <sup>10</sup>, Günter Lochnit <sup>11</sup>, Benjamin Mayer<sup>12</sup>, Yasser B. Ruiz-Blanco<sup>13</sup>, Markus Hoffmann <sup>14,15</sup>, Konstantin M. J. Sparrer <sup>1</sup>, Stefan Pöhlmann <sup>14</sup>, Elsa Sanchez-Garcia <sup>13</sup>, Frank Kirchhoff <sup>1</sup>, Manfred Frick <sup>3</sup> & Jan Münch <sup>1,2</sup> 

SARS-CoV-2 is a respiratory pathogen and primarily infects the airway epithelium. As our knowledge about innate immune factors of the respiratory tract against SARS-CoV-2 is limited, we generated and screened a peptide/protein library derived from bronchoalveolar lavage for inhibitors of SARS-CoV-2 spike-driven entry. Analysis of antiviral fractions revealed the presence of  $\alpha_1$ -antitrypsin ( $\alpha_1$ AT), a highly abundant circulating serine protease inhibitor. Here, we report that  $\alpha_1$ AT inhibits SARS-CoV-2 entry at physiological concentrations and suppresses viral replication in cell lines and primary cells including human airway epithelial cultures. We further demonstrate that  $\alpha_1$ AT binds and inactivates the serine protease TMPRSS2, which enzymatically primes the SARS-CoV-2 spike protein for membrane fusion. Thus, the acute phase protein  $\alpha_1$ AT is an inhibitor of TMPRSS2 and SARS-CoV-2 entry, and may play an important role in the innate immune defense against the novel coronavirus. Our findings suggest that repurposing of  $\alpha_1$ AT-containing drugs has prospects for the therapy of COVID-19.

<sup>1</sup>Institute of Molecular Virology, Ulm University Medical Center, Ulm, Germany. <sup>2</sup>Core Facility Functional Peptidomics, Ulm University Medical Center, Ulm, Germany. <sup>3</sup>Institute of General Physiology, Ulm University, Ulm, Germany. <sup>4</sup>Dermatology Clinic, University Medicine Mainz, Mainz, Germany. <sup>5</sup>Max-Planck-Institute for Polymer Research, Mainz, Germany. <sup>6</sup>Laboratory of Neuropathology, Department of Imaging and Pathology, KU-Leuven and Department of Pathology, UZ-Leuven, Leuven, Belgium. <sup>7</sup>Laboratory of Neuropathology, Institute of Pathology, Ulm University, Ulm, Germany. <sup>8</sup>Pneumology, Thoracic Oncology, Sleep and Respiratory Critical Care Medicine, Clinics Allgäu, Kempten and Immenstadt, Germany. <sup>9</sup>Institute for Microbiology and Hygiene, Ulm University Medical Center, Ulm, Germany. <sup>10</sup>Department of Internal Medicine 1, Ulm University Hospital, Ulm, Germany. <sup>11</sup>Institute of Biochemistry, Justus-Liebig University Giessen, Giessen, Germany. <sup>12</sup>Institute for Epidemiology and Medical Biometry, Ulm University, Ulm, Germany. <sup>13</sup>Computational Biochemistry, Center of Medical Biotechnology, University of Duisburg-Essen, Essen, Germany. <sup>14</sup>Infection Biology Unit, German Primate Center- Leibniz institute for Primate Research, Göttingen, Germany. <sup>15</sup>Faculty of Biology and Psychology, Georg-August-University, Göttingen, Germany. <sup>16</sup>These authors contributed equally: Lukas Wettstein, Tatjana Weil, Carina Conzelmann. ✉email: [jan.muench@uni-ulm.de](mailto:jan.muench@uni-ulm.de)

SARS-CoV-2 is mainly transmitted through inhalation of droplets and aerosols and subsequent infection of cells of the respiratory tract<sup>1</sup>. In many cases, infection is limited to the upper airways resulting in no or mild symptoms. Severe disease is caused by viral dissemination to the lungs ultimately resulting in acute respiratory distress syndrome, cytokine storm, multi-organ failure, septic shock, and death<sup>2,3</sup>. The airway epithelium acts as a frontline defense against respiratory pathogens via the mucociliary clearance and its immunological functions<sup>4</sup>. The epithelial lining fluid is rich in innate immunity peptides as well as proteins with antibacterial and antiviral activity, such as lysozyme, lactoferrin or defensins<sup>5</sup>. Currently, our knowledge about innate immune defense mechanisms against SARS-CoV-2 in the respiratory tract is limited.

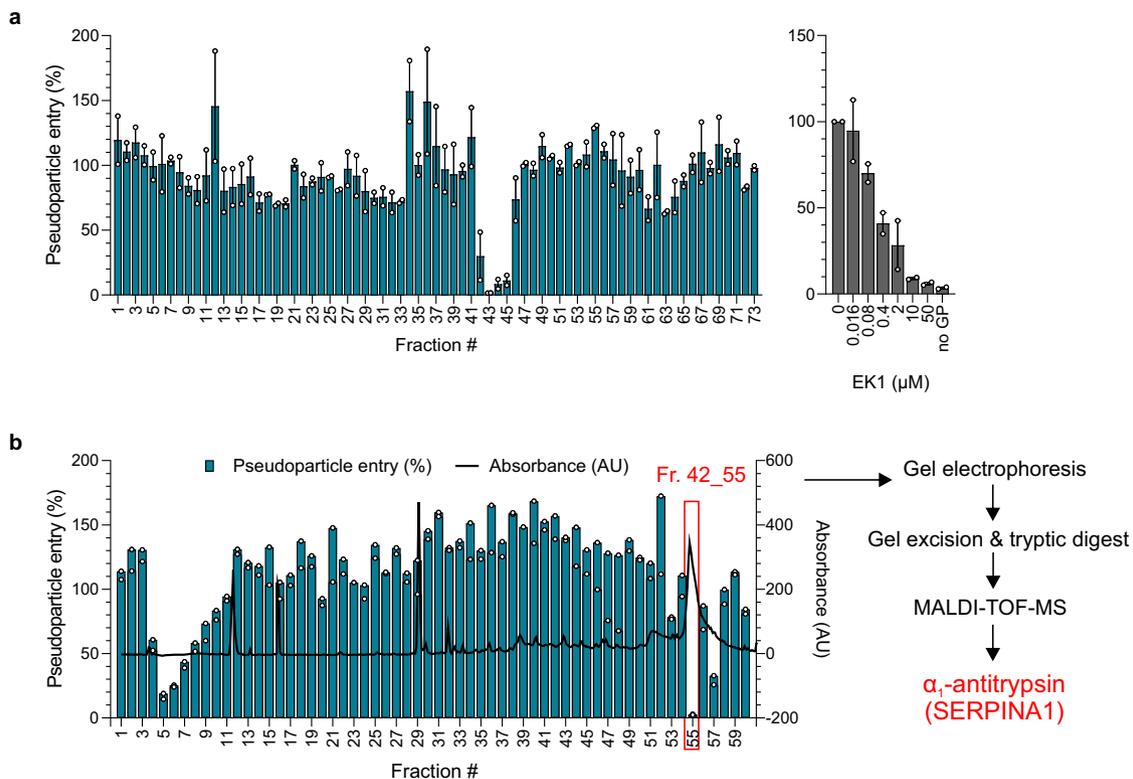
To identify endogenous antiviral peptides and proteins, we previously generated peptide/protein libraries from body fluids and tissues and screened the resulting fractions for antiviral factors<sup>6</sup>. This approach allowed to identify novel modulators of HIV-1<sup>7</sup>, CMV<sup>8</sup>, and HSV-2<sup>9</sup> infection, with prospects for clinical development as antiviral drugs<sup>10</sup>. In this work, we set out to identify factors of the respiratory tract that block SARS-CoV-2 infection. Screening a peptide/protein library derived from bronchoalveolar lavage allowed to identify  $\alpha_1$ -antitrypsin ( $\alpha_1$ AT), a highly abundant circulating serine protease inhibitor, as SARS-CoV-2 entry inhibitor. We show that  $\alpha_1$ AT suppresses viral replication in cell lines and human airway epithelial cultures and that it binds and inactivates the serine protease TMPRSS2, which enzymatically primes the SARS-CoV-2 spike protein for membrane fusion. Hence,  $\alpha_1$ AT

blocks SARS-CoV-2 entry by inhibiting TMPRSS2, and may play an important role in the innate immune defense against the novel coronavirus. Our findings suggest that repurposing of  $\alpha_1$ AT-containing drugs has prospects for treatment of COVID-19.

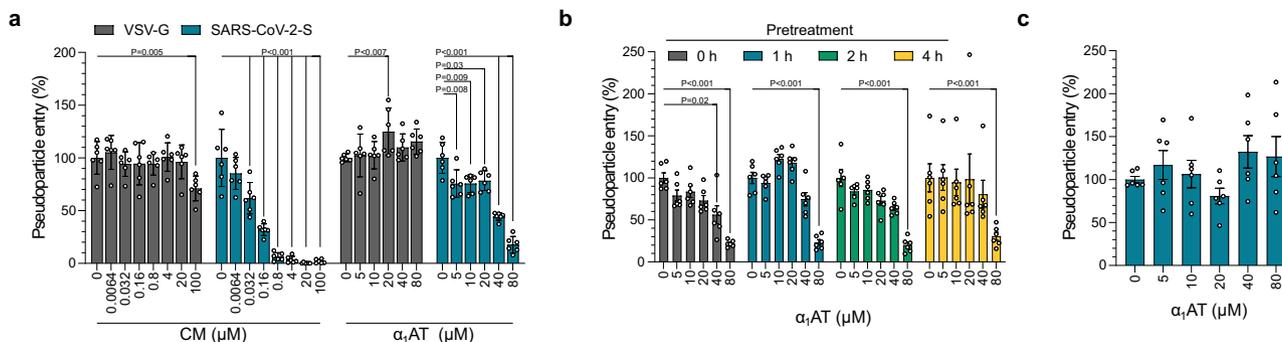
## Results

**Identification of  $\alpha_1$ AT as SARS-CoV-2 inhibitor.** To discover endogenous antiviral peptides and proteins, we extracted polypeptides from 6.5 kg of homogenized human lung or 20 liters of pooled bronchoalveolar lavage (BAL), and separated them by chromatographic means. The corresponding fractions were added to human epithelial colorectal carcinoma (Caco2) cells and the cells were inoculated with luciferase expressing lentiviral pseudoparticles carrying the SARS-CoV-2 spike protein<sup>11</sup>. None of the fractions of the lung library suppressed infection (Supplementary Fig. 1a). In contrast, fractions 42–45 of the BAL library prevented SARS-CoV-2 spike driven entry with an efficiency comparable to that of 10  $\mu$ M EK1, a coronavirus spike-specific peptide fusion inhibitor<sup>12</sup> (Fig. 1a). Titration of BAL fractions 42–45 onto Caco2 cells confirmed dose-dependent inhibition of SARS-CoV-2 spike pseudoparticles (Supplementary Fig. 1b).

To isolate the antiviral factor responsible for blocking spike-driven entry, the BAL mother fraction 42 was further separated chromatographically and the resulting sub-fractions analyzed for antiviral activity. As shown in Fig. 1b, sub-fractions 42\_3 to 42\_8 and 42\_57 reduced, and sub-fraction 42\_55 almost completely prevented host cell entry of SARS-CoV-2 spike pseudoparticles.



**Fig. 1 Identification of  $\alpha_1$ AT as SARS-CoV-2 inhibitor.** **a** Caco2 cells were treated with peptide/protein containing fractions of the bronchoalveolar lavage library (or EK1 peptide as inhibitor control) and transduced with luciferase-encoding lentiviral SARS-CoV-2 spike pseudoparticles. **b** Caco2 cells were treated with subfractions of mother fraction 42 (see **a**) of the bronchoalveolar lavage library and transduced with lentiviral SARS-CoV-2 spike pseudoparticles. Blue columns represent pseudoparticle entry and black line absorbance at 280 nm of the corresponding fraction. Transduction rates in **a** and **b** were determined 2 days after the addition of pseudoparticles by measuring luciferase activities in cell lysates. The means  $\pm$  SEM from  $n = 2$  (**a**) or individual datapoints from  $n = 1$  (**b**) independent experiments are shown, each performed in biological duplicates. The active fraction 42\_55 (red box) was analyzed via gel electrophoresis, gel excision, tryptic digest, and MALDI-TOF-MS to identify  $\alpha_1$ AT, a serine protease inhibitor (serpin). Source data are provided as a Source data file.



**Fig. 2**  $\alpha_1$ AT inhibits SARS-CoV-2 spike but not VSV-G mediated pseudovirus entry. **a** Prolastin ( $\alpha_1$ AT) and the control small molecule inhibitor camostat mesylate (CM) were added to Caco2 cells 1 h prior to transduction of cells with rhabdoviral SARS-CoV-2 spike (blue) or VSV-G pseudoparticles (gray). **b** Caco2 cells were treated for the indicated hours with Prolastin ( $\alpha_1$ AT) and were then transduced with rhabdoviral SARS-CoV-2 spike pseudoparticles. **c** Prolastin ( $\alpha_1$ AT) was added 2 h post transduction of Caco2 cells with rhabdoviral SARS-CoV-2 spike pseudoparticles. Transduction rates in **a–c** were determined at 16 h after addition of pseudoparticles by measuring luciferase activities in cell lysates. The mean  $\pm$  SEM from  $n = 2$  experiments in biological triplicates are shown (2-way ANOVA with Dunett’s multiple comparison test). Source data are provided as a Source data file.

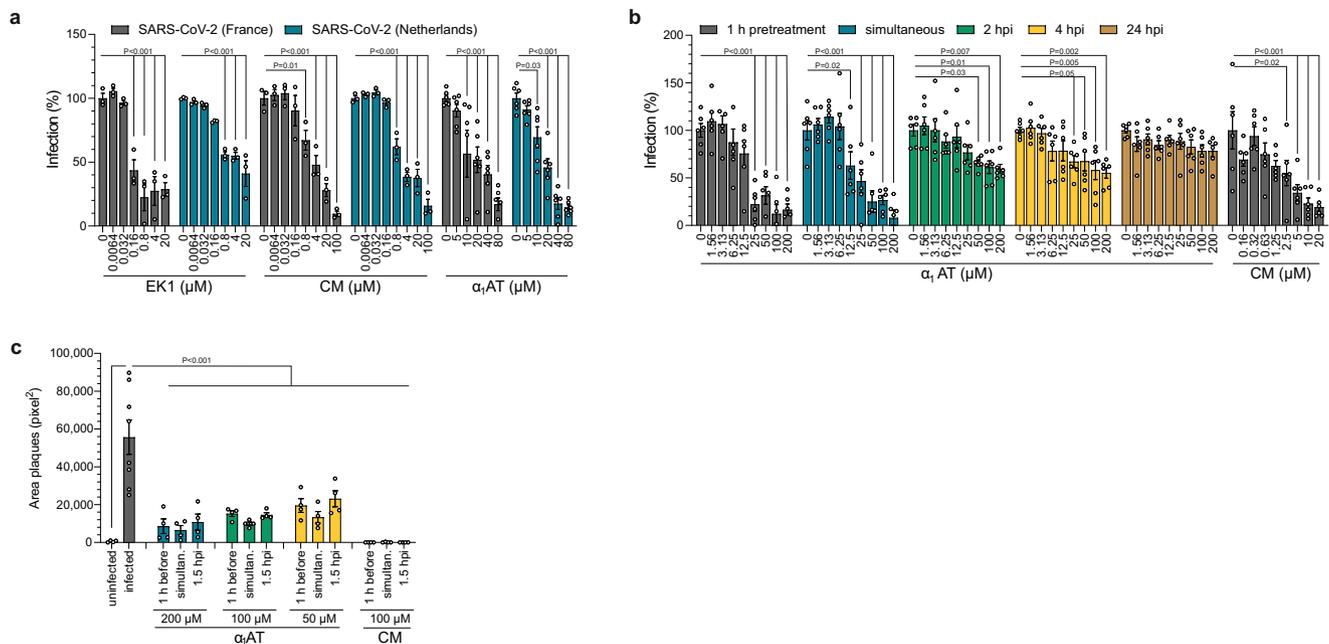
Analysis of these inhibitory fractions by gel electrophoresis revealed distinct protein bands in sub-fractions 42\_55 and 42\_57, whereas no protein was detectable in sub-fractions 42\_5 to 42\_7 (Supplementary Fig. 2a). The most active sub-fraction 42\_55 contained a prominent band at  $\sim 52$  kDa, which was also present in other active fractions but hardly found in neighboring fractions showing no antiviral activity (e.g., 42\_49) (Supplementary Fig. 2a). This band was excised from the gel, digested with trypsin and subjected to MALDI-TOF-MS revealing a 100 % sequence identity to  $\alpha_1$ -antitrypsin (SERPINA1) (Supplementary Fig. 2b and Supplementary Table 1), a 52 kDa protease inhibitor<sup>13</sup>. The presence of  $\alpha_1$ -antitrypsin ( $\alpha_1$ AT) in inhibitory fractions 42\_55 and 42\_57 was confirmed by western blot analysis with an  $\alpha_1$ AT-specific antibody (Supplementary Fig. 2c).  $\alpha_1$ AT belongs to the serine protease inhibitor (serpin) superfamily and protects lung tissue from digestive enzymes released by immune cells, in particular neutrophil elastase<sup>14</sup>. The serpin has a reference range in blood of 0.9–2 mg/ml (corresponding to  $\sim 17$ –38  $\mu$ M), but the concentration can rise 4- to 5-fold upon acute inflammation<sup>15</sup>.  $\alpha_1$ AT purified from donor blood is also available as pharmaceutical product (e.g., Prolastin) for intravenous substitution therapy of  $\alpha_1$ AT deficiency, a hereditary disorder that leads to chronic uncontrolled tissue breakdown in the lower respiratory tract<sup>16</sup>.

#### $\alpha_1$ AT inhibits SARS-CoV-2 spike mediated pseudovirus entry.

To test whether the serpin indeed inhibits SARS-CoV-2, Caco2 cells were exposed to Prolastin, a pharmaceutical preparation of  $\alpha_1$ AT, or camostat mesylate (CM), a small molecule inhibitor of the SARS-CoV-2 spike priming protease TMPRSS2<sup>11,17</sup>.  $\alpha_1$ AT and CM both suppressed SARS-CoV-2 spike pseudoparticle entry with half-maximal inhibitory concentrations ( $IC_{50}$ ) of  $\sim 38.5$   $\mu$ M for  $\alpha_1$ AT, and  $\sim 0.05$   $\mu$ M for CM, respectively (Fig. 2a). Cell viability assays showed that  $\alpha_1$ AT displayed no cytotoxic effects at concentrations of up to 160  $\mu$ M (8.3 mg/ml), whereas CM reduced cell viability at concentrations of 200  $\mu$ M, due to DMSO in the stock (Supplementary Fig. 3). The antiviral activity of  $\alpha_1$ AT and CM was specific for the coronavirus spike because entry of pseudoparticles carrying the G-protein of VSV was not affected (Fig. 2a). In fact, time-of-addition experiments demonstrated that  $\alpha_1$ AT prevented single-round SARS-CoV-2 spike pseudovirus entry only if added prior (1–4 h) to or during infection (Fig. 2b) but not if added 2 h post infection (Fig. 2c). Taken together, these data show that  $\alpha_1$ AT specifically targets SARS-CoV-2 spike-driven entry.

**$\alpha_1$ AT inhibits SARS-CoV-2 infection and replication.** To determine whether  $\alpha_1$ AT inhibits not only spike pseudoparticles but also wild-type SARS-CoV-2, we examined its activity against two SARS-CoV-2 isolates from France (bearing the spike variant D614) and the Netherlands (bearing the fitter spike variant G614)<sup>18,19</sup>. For this, we assessed survival rates of TMPRSS2-expressing Vero E6 cells infected in the absence or presence of EK1, CM or  $\alpha_1$ AT by MTS assay. In the absence of drugs, infection by both SARS-CoV-2 isolates resulted in virus-induced cytopathic effects (CPE) and reduced cell viability by  $\sim 80$  % (Fig. 3a and Supplementary Fig. 4a, b). Microscopic evaluation revealed the absence of CPE in the presence of high concentrations of EK1, CM or  $\alpha_1$ AT, and MTS assay confirmed a concentration-dependent inhibition of cell death and viral replication by EK1 and CM (Supplementary Fig. 4a, b and Fig. 3a) with average  $IC_{50}$  values against both SARS-CoV-2 isolates of 2.8  $\mu$ M for EK1, and 3.6  $\mu$ M for CM, respectively (Fig. 3a).  $\alpha_1$ AT inhibited the French SARS-CoV-2 isolate with an  $IC_{50}$  of 21.2  $\mu$ M (1.1 mg/ml), and the Dutch strain with an  $IC_{50}$  of 17.3  $\mu$ M (0.9 mg/ml) (Fig. 3a). Almost complete rescue of cell viability was observed at  $\alpha_1$ AT concentrations of 40–80  $\mu$ M (Supplementary Fig. 4a, b and Fig. 3a). A similar antiviral activity of  $\alpha_1$ AT against both SARS-CoV-2 isolates was determined in Caco2 cells (Supplementary Fig. 4c).

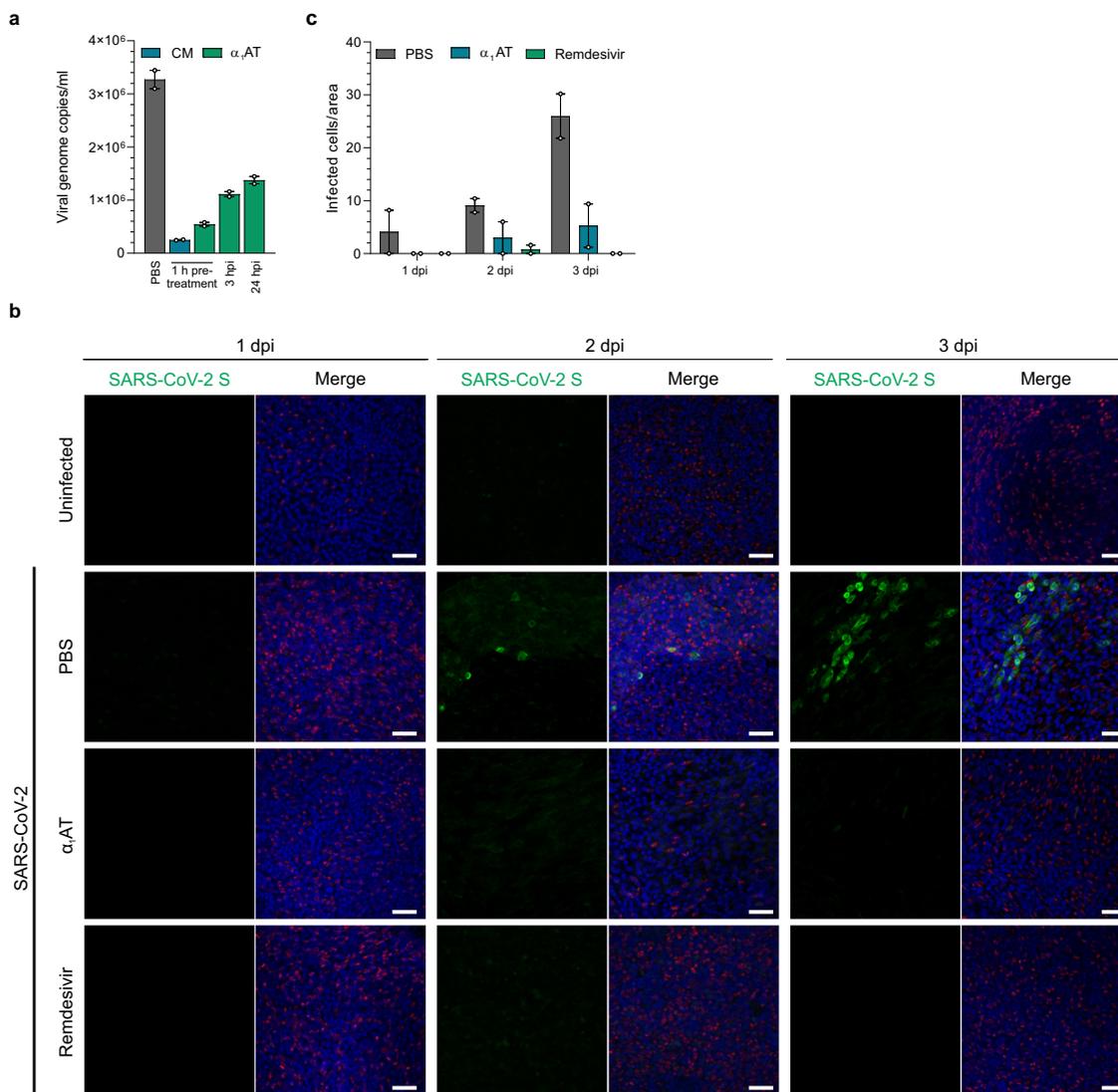
We next performed time-of-addition experiments and found that  $\alpha_1$ AT most effectively inhibited SARS-CoV-2 replication when added 1 h prior to or simultaneously with infection (Fig. 3b, Supplementary Fig. 5a, b), confirming pseudovirus data (Fig. 2b).  $\alpha_1$ AT also suppressed SARS-CoV-2 if added 2 or 4 h post infection (Fig. 3b and Supplementary Fig. 5a and b). Considering the 10 h time span of the viral life cycle<sup>20</sup>, these data suggest that the serpin inhibits spreading virus infection, i.e., infection of uninfected target cells by progeny virus. Accordingly, no antiviral effect of  $\alpha_1$ AT was observed when added 24 h post infection, when already two replication cycles have been completed (Fig. 3b and Supplementary Fig. 5a,b). To investigate whether  $\alpha_1$ AT may inhibit cell-to-cell spread of SARS-CoV-2, a plaque assay with cellulose containing medium (which prevents cell-free viral replication) was performed (Supplementary Fig. 6a,b) and plaque sizes were quantified (Fig. 3c).  $\alpha_1$ AT concentrations of 50 to 200  $\mu$ M blocked viral spread most effectively when the serpin was added 1 h prior to infection (54–87 % reduction) or simultaneously with infection (75–89 % reduction). When adding  $\alpha_1$ AT 1.5 h post infection, spread was still inhibited by 49–83 %, demonstrating effective inhibition of SARS-CoV-2 infection and cell-to-cell viral spread by the serpin.



**Fig. 3**  $\alpha_1$ AT inhibits SARS-CoV-2 infection and replication. **a** TMPRSS2-expressing Vero E6 cells were treated with Prolastin ( $\alpha_1$ AT), EK1 or camostat mesylate (CM) for 1 h, and infected with SARS-CoV-2 isolates either from France (gray) or the Netherlands (blue) at a MOI of 0.001. Virus-induced cytopathic effects were assessed at 2 days post infection by MTS assay (see Supplementary Fig. 4 a, b for raw and cell viability data). **b** TMPRSS2-expressing Vero E6 cells were treated with Prolastin ( $\alpha_1$ AT) at indicated timepoints prior to, simultaneously with or post infection with SARS-CoV-2 at a MOI of 0.001. Camostat mesylate (CM) control was added 1 h prior to infection. Infection rates were assessed at 2 days post infection by MTS assay (see also Supplementary Fig. 5a, b for raw and cell viability data). **c** TMPRSS2-expressing Vero E6 cells were treated with Prolastin ( $\alpha_1$ AT) or CM 1 h prior to, simultaneously with or 1.5 h post infection with SARS-CoV-2. At 1.5 h post infection cellulose overlay was performed. At 2 days post infection, cells were stained with crystal violet (see Supplementary Fig. 6a, b) and plaque areas were quantified. The mean  $\pm$  SEM from  $n = 1$  (**a**, EK1 and CM) or  $n = 2$  independent experiments in biological triplicates (**a**,  $\alpha_1$ AT, **b**) or duplicates (**c**) and quadruplicates (**c**, infected) are shown. (2-way ANOVA with Dunett's multiple comparison test (**a**, **b**), ordinary one-way ANOVA with Dunett's multiple comparison test (**c**)). Source data are provided as a Source data file.

**$\alpha_1$ AT inhibits SARS-CoV-2 replication in primary human airway cells.** To corroborate the antiviral activity of  $\alpha_1$ AT in human primary target cells, we used small airway epithelial cells (SAECs) that support low level SARS-CoV-2 replication. For this, SAECs pretreated with buffer only (PBS), 80  $\mu$ M of  $\alpha_1$ AT, or 100  $\mu$ M of CM were infected with a high dose (MOI of 1) of SARS-CoV-2. Furthermore,  $\alpha_1$ AT was added 3 and 24 h post infection to PBS pretreated cells. Cells were cultivated in the presence of the inhibitors for 6 days, and then viral genome copies in supernatants were quantified by RT-qPCR. As shown in Fig. 4a, CM and  $\alpha_1$ AT present during infection reduced viral titers by  $\sim$ 92 and 83 %, respectively.  $\alpha_1$ AT that was added 3 and 24 h post infection also inhibited viral replication, albeit to a lesser extent (67 and 58 % reduction, respectively). We next analyzed antiviral activity of  $\alpha_1$ AT in fully differentiated primary human airway epithelial cells (HAECs) grown at the air-liquid interface. HAECs derived from two donors were treated with 10  $\mu$ M (0.5 mg/ml) of  $\alpha_1$ AT and then exposed to SARS-CoV-2. As control, we used 5  $\mu$ M of remdesivir, which has previously been shown to suppress coronavirus replication in HAECs<sup>21</sup>. At days 1, 2, and 3 post infection cells were fixed and stained with antibodies against SARS-CoV-2 spike<sup>22</sup>, and  $\alpha$ -tubulin as marker for ciliated cells at the apical surface<sup>23</sup> (Fig. 4b). In infected, PBS-treated HAECs, SARS-CoV-2 spike expression was readily detectable, mostly in neighboring ciliated cells, and increased between day 2 and 3 (Fig. 4b, c and Supplementary Fig. 7), demonstrating productive infection and viral spread in the epithelia. Spike expression levels in  $\alpha_1$ AT and remdesivir treated cultures were greatly reduced at days 2 and 3 in both donors (Fig. 4b, c, Supplementary Fig. 7). Thus,  $\alpha_1$ AT suppresses SARS-CoV-2 infection of HAECs.

**$\alpha_1$ AT binds and inhibits TMPRSS2 protease activity.** Finally, we set out to explore the mechanism underlying SARS-CoV-2 inhibition by  $\alpha_1$ AT. A recent preprint publication suggests that  $\alpha_1$ AT may suppress TMPRSS2, the spike priming protease, similar to camostat mesylate<sup>24</sup>. First, we established a computational model of the Michaelis complex of  $\alpha_1$ AT and TMPRSS2 using protein-protein docking calculations complemented by structural refinement (Supplementary Methods) and observed a calculated binding free energy of  $-10.2 \pm 2.2$  kcal/mol (Fig. 5a, b). We applied this modeling to the complex of  $\alpha_1$ AT (Pittsburgh variant, M358R) and trypsin (S195A) and observed a calculated binding free energy of  $-10 \pm 1.6$  kcal/mol (Supplementary Fig. 9). This calculated binding free energy is close to the experimentally determined binding free energy of the  $\alpha_1$ AT (Pittsburgh variant, M358R) and trypsin (S195A) complex<sup>25</sup>, therefore verifying our modeling approach. The computed structure of the Michaelis complex suggests that this initial step of the reaction mechanism is favored by the interaction of Leu353 and Ala355 of  $\alpha_1$ AT (numbers according to PDB ID 3cwm29), with a hydrophobic patch of TMPRSS2 formed by Tyr416, Leu419, and Trp461, which is located next to the catalytic triad (His296, Asp345, Ser441) (Fig. 5b). Thus, the anchoring site, defined by the hydrophobic patch, can effectively position and restrain the reactive center loop prior to the cleavage. These structural insights together with the calculated affinity of this complex, support a favorable interaction between TMPRSS2 and  $\alpha_1$ AT. To verify a direct interaction of  $\alpha_1$ AT with TMPRSS2, surface plasmon resonance analysis was performed. To this end, recombinant TMPRSS2 was immobilized on a metal surface (Supplementary Fig. 8a) and subjected to increasing doses (0–10  $\mu$ M) of  $\alpha_1$ AT (Supplementary Fig. 8b). A  $K_d$  of  $941 \pm 297$  nM was measured by performing equilibrium analysis which was reached within 5 min



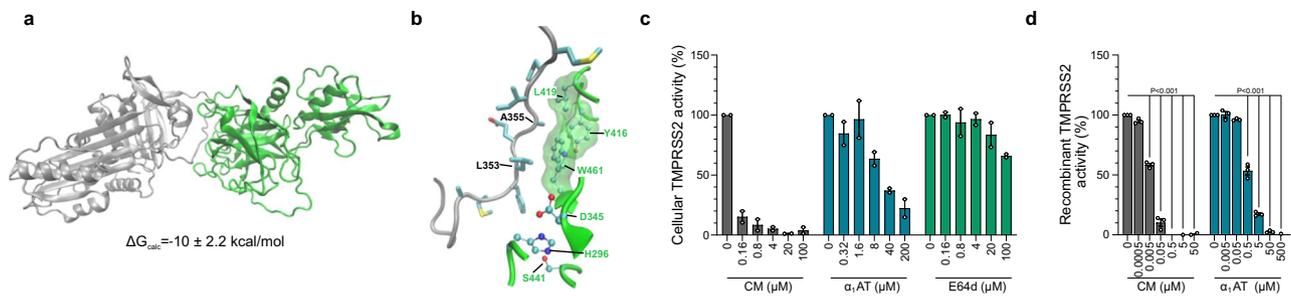
**Fig. 4**  $\alpha_1$ AT inhibits SARS-CoV-2 replication in primary human airway cells. **a** Small airway epithelial cells (SAECs) were treated with 100  $\mu$ M of Prolastin ( $\alpha_1$ AT, green) 1 h prior to, or 3 and 24 h post infection (hpi) with SARS-CoV-2 at a MOI of 1. 100  $\mu$ M CM (blue) added 1 h prior to infection served as control. Immediately after the inoculum was removed (0 dpi) and at day 6 post infection, supernatants were harvested and subjected to RT-qPCR specific for SARS-CoV-2 ORF1b nsp14. Virus titers from day 0 were subtracted from titers at day 6. The means of technical duplicates of one experiment performed in triplicates are shown. **b** The apical and basal site of human airway epithelial cells (HAEC) grown at air-liquid interface was exposed to PBS,  $\alpha_1$ AT (10  $\mu$ M or 0.5 mg/ml) and remdesivir (5  $\mu$ M) and then inoculated with SARS-CoV-2 ( $9.25 \times 10^2$  PFU) for 2 h. Cells were fixed at day 1, 2, and 3 post infection, stained with DAPI (cell nuclei, blue), a SARS-CoV-2 specific spike antibody (SARS-CoV-2 S, green) and an  $\alpha$ -tubulin-specific antibody (red). Images shown are derived from one donor and represent maximum projections of serial sections along the basolateral to apical cell axis. Scale bar: 50  $\mu$ m. **c** Number of infected cells per area in mock- (PBS, gray),  $\alpha_1$ AT- (blue) or remdesivir-treated (green), SARS-CoV-2 infected HAECs. Values represent the mean number of infected HAECs from 2 donors at 5 random spots per culture, treatment and day  $\pm$  SEM. Source data are provided as a Source Data file. For more images, see Supplementary Fig. 7.

interaction time of  $\alpha_1$ AT with TMPRSS2 (Supplementary Fig. 8c), proving a physical interaction of the protease inhibitor with the protease. To investigate whether  $\alpha_1$ AT not only binds but also inhibits proteolytic activity of TMPRSS2, the protease was over-expressed in HEK293T cells and incubated with  $\alpha_1$ AT, CM, or the cysteine protease inhibitor E-64d<sup>11</sup>. Enzymatic activity was assessed by adding a specific substrate that emits fluorescence after proteolytic cleavage. CM but not E-64d suppressed TMPRSS2 activity at sub-micromolar concentrations (Fig. 5c and Supplementary Fig. 8d, f). Interestingly, 100 and 200  $\mu$ M of  $\alpha_1$ AT effectively inhibited cell-associated TMPRSS2 activity (Fig. 5c and Supplementary Fig. 8e). Experiments performed with recombinant TMPRSS2 enzyme confirmed effective and dose-dependent inhibition of TMPRSS2 activity by  $\alpha_1$ AT at physiologically relevant concentrations of 5–50

$\mu$ M (Fig. 5d). These data demonstrate that the abundant serpin  $\alpha_1$ AT is an endogenous inhibitor of TMPRSS2 proteolytic activity.

### Discussion

This study demonstrates that  $\alpha_1$ AT is a potent inhibitor of SARS-CoV-2 which blocks viral replication in cell lines, primary small airway epithelial cells and fully differentiated airway epithelium cultures.  $\alpha_1$ AT binds and inhibits TMPRSS2, a transmembrane serine protease that cleaves the SARS-CoV-2 spike protein to enable viral fusion.  $\alpha_1$ AT is the most abundant serine protease inhibitor (serpin) in the circulation (0.9–2 mg/ml, 17–38  $\mu$ M), and levels of this acute phase protein further increase during acute inflammation<sup>15</sup>. We show that plasma-derived  $\alpha_1$ AT blocks SARS-CoV-2 infection with IC<sub>50</sub> values of 10–20  $\mu$ M, which are



**Fig. 5**  $\alpha_1$ AT binds the extracellular region of TMPRSS2 and inhibits TMPRSS2 protease activity. **a** Protein–protein docking analysis of a homology model for the TMPRSS2 extracellular fragment (green, PDB 1z8g) and  $\alpha_1$ AT (gray, PDB 3cwm) and computationally calculated binding free energy ( $\Delta G_{\text{calc}}$ ) of the complex. **b** Detailed view on  $\alpha_1$ AT–TMPRSS2 binding interface. The sidechains of  $\alpha_1$ AT (gray) residues are represented with sticks, while sidechains of TMPRSS2 (green) are shown with balls and sticks. Hydrogen atoms are omitted for clarity, carbon, oxygen, nitrogen or sulfur atoms of amino acid side chains depicted in light blue, red, dark blue or yellow, respectively. The hydrophobic patch near the TMPRSS2 catalytic triad is highlighted with a green transparent surface. **c**  $\alpha_1$ AT inhibits cell-associated TMPRSS2 activity. HEK293T cells were transfected with a TMPRSS2 expression plasmid and treated with Prolastin ( $\alpha_1$ AT, blue), camostat mesylate (CM, gray) or E-64d (green) followed by incubation with the fluorogenic TMPRSS2 protease substrate BOC–Gln–Ala–Arg–AMC. Graph shows the relative area under the curve analysis of fluorescence intensities over 2 h that were corrected by values for mock-transfected HEK293T cells. **d**  $\alpha_1$ AT inhibits recombinant TMPRSS2 enzyme activity. Recombinant human TMPRSS2 was mixed with Prolastin ( $\alpha_1$ AT, blue) or CM (gray) prior to addition of fluorogenic TMPRSS2 protease substrate BOC–Gln–Ala–Arg–AMC, graph shows relative fluorescence intensities after 3 h of incubation. The mean  $\pm$  SEM of  $n = 2$  (**c**) or  $n = 3$  (**d**) independent experiments in biological duplicates are shown (ordinary one-way ANOVA with Dunnett’s multiple comparison test). Source data are provided as a Source data file.

well within the physiological range, suggesting that plasma itself may exert anti-SARS-CoV-2 activity. In fact, a recent preprint demonstrates that naïve serum exhibits inhibition of SARS-CoV-2 entry and suggested  $\alpha_1$ AT and to a lesser degree  $\alpha$ -2-macroglobulin as potential antiviral factors<sup>26</sup>. Thus,  $\alpha_1$ AT may serve as natural inhibitor of the novel coronavirus, in particular during acute SARS-CoV-2 infection, when blood  $\alpha_1$ AT concentrations increase, as recently shown in a cohort of 40 COVID-19 patients<sup>27</sup>. Further studies to clarify the physiological relevance of  $\alpha_1$ AT in SARS-CoV-2 infection and whether inter-individual differences in  $\alpha_1$ AT concentrations in blood, lungs, and other organs correlate with viral loads and disease progression are thus highly warranted.

$\alpha_1$ AT was isolated as inhibitor of SARS-CoV-2 from a complex peptide/protein library that was generated from pooled bronchoalveolar lavage (BAL). The corresponding fractions presumably contain all soluble peptides and small proteins present in lung fluid and should allow identification of those innate immune factors that are most relevant for controlling viral infection in vivo, at least in the absence of a specific antiviral host immune response.  $\alpha_1$ AT-containing fractions 42–45 were more potent in inhibiting SARS-CoV-2 entry than all remaining BAL- and lung-derived fractions, suggesting a relevant role of the serpin in vivo. In normal individuals,  $\alpha_1$ AT levels range between 10–40  $\mu\text{M}$  in alveolar interstitial fluid, and 2–5  $\mu\text{M}$  in alveolar extracellular lining fluid<sup>28–30</sup>. Assuming that  $\alpha_1$ AT concentrations further increase during the acute phase response, the serpin may act as relevant innate immune factor against SARS-CoV-2 in the respiratory tract.

Biomolecular modeling, surface plasmon resonance, and biochemical analysis established  $\alpha_1$ AT as inhibitor of TMPRSS2, a cell surface serine protease that is involved in cell–cell and cell–matrix interactions, and in prostate cancer metastasis<sup>31–33</sup>. TMPRSS2 not only primes the spike protein of SARS-CoV-2 but is also utilized as entry factor by other coronaviruses such as SARS-CoV, MERS-CoV or the common cold corona virus 229E<sup>34–36</sup>. Furthermore, TMPRSS2 is also the major hemagglutinin-activating protease of influenza A virus in human airways<sup>37–39</sup>. It is, therefore, not only of great interest to clarify whether  $\alpha_1$ AT may act as broad-spectrum inhibitor against respiratory viral pathogens, but also to evaluate its role in prostate cancer.

$\alpha_1$ AT is an approved drug for treatment of  $\alpha_1$ AT deficiency implying repurposing for the therapy of COVID-19.  $\alpha_1$ AT deficiency is a hereditary disorder that results in reduced circulating concentrations of  $\alpha_1$ AT and consequently a chronic uninhibited breakdown of tissue in the lungs, mainly mediated by neutrophil elastase<sup>40</sup>. Several products containing  $\alpha_1$ AT purified from human plasma (such as Prolastin used herein) are approved for decades for intravenous augmentation therapy in patients suffering from  $\alpha_1$ AT deficiency. Of note,  $\alpha_1$ AT may not only be beneficial in COVID-19 therapy because of its direct antiviral effect by targeting TMPRSS2-mediated SARS-CoV-2 entry, but also by inhibiting neutrophil elastase, which has been proposed to act as alternative spike priming protease<sup>41,42</sup> and might contribute to pulmonary inflammation in COVID-19<sup>43,44</sup>. Moreover, a recent study revealed that the proinflammatory IL-6 to  $\alpha_1$ AT ratio in patients with severe COVID-19 was more than two times higher compared to a pneumonia control cohort<sup>27</sup>. Of note,  $\alpha_1$ AT can also be administered via inhalation<sup>45</sup> and at substantially higher doses than those in routine  $\alpha_1$ AT deficiency. In some studies,  $\alpha_1$ AT has been administered at doses of 120 mg/kg<sup>46,47</sup> and even 250 mg/kg without causing side effects, resulting in a 5-fold increase of serpin concentration in lung epithelial lining fluid of  $\alpha_1$ AT deficient patients<sup>48</sup>. However, whether  $\alpha_1$ AT infusion or inhalation allows to reach local concentrations of the serpin that are sufficient to block SARS-CoV-2 in lungs or other organs without causing severe side effects remains to be addressed in clinical studies. Nevertheless, we and others suggest that  $\alpha_1$ AT supplementation may have a therapeutic benefit because of its antiviral and anti-inflammatory properties. Thus, rapid evaluation of  $\alpha_1$ AT for the treatment of severe COVID-19 disease is highly warranted<sup>26,27,49,50</sup> and four clinical trials (NCT04495101, NCT04385836, NCT04547140, EudraCT: 2020-001391-15) have been initiated to evaluate the therapeutic potential of  $\alpha_1$ AT in hospitalized COVID-19 patients.

## Methods

**Ethical statements.** Ethical approval for the generation of peptide libraries from lungs and BAL was obtained from the Ethics Committee of Ulm University (application numbers 274/12 and 324/12). The collection of tissue and generation of human airway epithelia cell cultures for research from these primary cells has been approved by the ethics committee at the University of Ulm (nasal brushings,

application number 126/19) and Medical School Hannover (airway tissue, application number 2699-2015). Tissues were obtained from donors who gave informed consent.

**Reagents.** Camostat mesylate (SML0057) and E-64d (E8640) were obtained from Merck.  $\alpha$ 1-antitrypsin (52 kDa) was obtained from Merck (A9024) or GRIFOLS (Prolastin®) and solubilized freshly in dH<sub>2</sub>O prior to use, Boc-Gln-Ala-Arg-AMC peptide was obtained from PEPTIDE INSTITUTE (3135-v) or Bachem (4017019.0005), EK1 peptide (H<sub>2</sub>N-SLDQINVTFLDLEYEMKKLEEAIAKKLEE-SYIDLKEL-COOH) was synthesized by solid-phase synthesis.

**Generation of a peptide/protein library from lungs.** 6.5 kg of human lung were obtained from deceased individuals without known diseases from pathology of Ulm University. The organs were frozen at  $-20^{\circ}\text{C}$  and freeze-dried. Lung material was homogenized and peptides and small proteins were extracted using ice-cold acetic acid. Then the mixture was centrifuged at  $3500 \times g$  and the supernatant was applied to a  $0.45 \mu\text{m}$  filter. Thereafter, the obtained peptides and proteins were separated by a 30 kDa molecular weight cut-off ultrafiltration step. The filtrate was separated by reversed-phase (RP) chromatography with a Sepax Poly RP300  $50 \times 300$  mm column (Sepax Technologies, Newark DE, USA 260300-30025) at a flow rate of 100 ml/min and a gradient ranging from 0.1 % TFA (Merck, 1082621000) in ultrapure water (buffer A) to 0.1 % TFA in acetonitrile (J.T.Baker, JT9012-3, buffer B). Reversed-phase chromatographic fractions of 50 ml were collected to constitute the lung peptide bank, from which 1 ml-aliquots (2 %) were lyophilized and used for antiviral testing.

**Generation of a peptide/protein library from BAL.** Clinical samples of bronchoalveolar lavage (BAL) comprising a total of 201 were collected and immediately frozen for further processing. Peptide/protein extraction was done by acidification with acetic acid to pH 3, followed by centrifugation at  $3500 \times g$  and filtration ( $0.45 \mu\text{m}$ ) of the supernatant. Further, the filtered BAL was subjected to a 30 kDa molecular weight cut-off. Chromatographic fractionation of the filtrate was performed by using a reversed-phase (PS/DVB) Sepax Poly RP300 HPLC  $30 \times 250$  mm column (Sepax Technologies, Newark DE, USA 260300-30025), at a flow rate of 55 ml/min and a gradient ranging from 0.1 % TFA in ultrapure water (buffer A) to 0.1 % TFA in acetonitrile. Seventy-three reversed-phase chromatographic fractions of 50 ml were collected to constitute the BAL peptide bank, from which 1 ml aliquots (2 %) were lyophilized and used for antiviral testing. For further purification of active fractions, a second reversed-phase C18 HPLC  $4.6 \times 250$  mm column (Phenomenex 00G-4605-EO) was used at a flow rate of 0.8 ml/min and a gradient ranging from 0.1 % TFA in ultrapure water (buffer A) to 0.1 % TFA in acetonitrile, followed again by lyophilization prior to antiviral testing.

**Cell culture.** Unless stated otherwise, HEK293T cells (ATCC, CRL-3216) were cultivated in DMEM supplemented with 10 % fetal calf serum (FCS), 2 mM L-glutamine, 100 U/ml penicillin, and 100 mg/ml streptomycin. Caco2 cells (ATCC, HTB-37) were cultivated in DMEM supplemented with 10 % FCS, 2 mM glutamine, 100 U/ml penicillin and 100 mg/ml streptomycin,  $1 \times$  non-essential amino acids (NEAA) and 1 mM sodium pyruvate. TMPRSS2-expressing Vero E6 cells (kindly provided by the National Institute for Biological Standards and Control (NIBSC), #100978) were cultivated in DMEM supplemented with 10 % fetal calf serum (FCS), 2 mM L-glutamine, 100 U/ml penicillin, 100 mg/ml streptomycin and 1 mg/ml geneticin. Human Small Airway Epithelial Cells (Lonza, CC-2547, batch: 18TL082942, donor: 68 years, female) were cultivated in SAGM™ Small Airway Epithelial Cell Growth Medium (Lonza, CC-3118).

**Generation of lentiviral pseudotypes.** For generation of lentiviral SARS-CoV-2 pseudoparticles (LV(Luc)-CoV-2-S) 900,000 HEK293T cells were seeded in 2 ml HEK293T medium. The next day, medium was replaced and cells were transfected with a total of 1  $\mu\text{g}$  DNA using polyethylenimine (PEI). To this end, 2 % of pCG1-SARS-2-S (encoding the spike protein of SARS-CoV-2 isolate Wuhan-Hu-1, NCBI reference Sequence YP\_009724390.1) were mixed with 98 % of pCMVdr8\_91 (encoding a replication-deficient lentivirus) and pSEW-Luc2 (encoding a luciferase reporter gene, both kindly provided by Christian Buchholz) in a 1:1 ratio in OptiMEM. Plasmid DNA was mixed with PEI at a DNA:PEI ratio of 1:3 (3  $\mu\text{g}$  PEI per 1  $\mu\text{g}$  DNA), incubated for 20 min at RT and added to cells dropwise. At 8 h post transfection, medium was removed, cells were washed with 2 ml of PBS and 2 ml of HEK293T medium with 2.5 % FCS were added. At 48 h post transfection, pseudoparticles containing supernatants were harvested and clarified by centrifugation for 5 min at  $450 \times g$ .

**Generation of VSV-based pseudotypes.** For generation of VSV-based SARS-CoV-2 pseudoparticles (VSV(Luc\_eGFP)-CoV-2-S), HEK293T were seeded in 30 ml HEK293T medium in a T175 cell culture flask. The next day, cells were transfected with a total 44  $\mu\text{g}$  pCG1-SARS-2-S using PEI. Plasmid DNA and PEI were mixed in 4.5 ml of OptiMEM at a 2:1 ratio (2  $\mu\text{g}$  PEI per 1  $\mu\text{g}$  DNA), incubated for 20 min at RT and added to cells dropwise. 24 h post transfection, medium was replaced and cells were transduced with VSV-G-protein pseudotyped VSV

encoding luciferase and GFP reporter gene (kindly provided by Gert Zimmer, Institute of Virology and Immunology, Mittelhäusern/Switzerland<sup>51</sup>). At 2 h post transduction, cells were washed three times with PBS and cultivated for 16 h in HEPES-buffered HEK293T medium. Virus containing supernatants were then harvested and clarified by centrifugation for 5 min at  $450 \times g$ , residual pseudo-particles harboring VSV-G-protein were blocked by addition of anti-VSV-G hybridoma supernatant at 1/10 volume ratio (1I, mouse hybridoma supernatant from CRL-2700; ATCC). Virus stocks were concentrated 10-fold using a 100 kDa Amicon molecular weight cutoff and stored at  $-80^{\circ}\text{C}$  until use.

**SARS-CoV-2 strains and propagation.** Viral isolate BetaCoV/France/IDF0372/2020 (#014V-03890) and BetaCoV/Netherlands/01/ NL/2020 (#010V-03903) were obtained from the European Virus Archive global and propagated on Vero E6 or Caco2 cells. To this end, 70–90 % confluent cells in 75 cm<sup>2</sup> cell culture flasks were inoculated with SARS-CoV-2 isolate (multiplicity of infection (MOI) of 0.03–0.1) in 3.5 ml serum-free medium. Cells were incubated for 2 h at  $37^{\circ}\text{C}$ , before adding 20 ml medium containing 15 mM HEPES. Cells were incubated at  $37^{\circ}\text{C}$  and supernatant harvested when a cytopathic effect (CPE) was visible. Supernatants were centrifuged for 5 min at  $1000 \times g$  to remove cellular debris, and then aliquoted and stored at  $-80^{\circ}\text{C}$  as virus stocks. Infectious virus titer was determined as plaque-forming units (PFU) on Vero E6 cells, which was used to calculate MOI.

**Screening lung and BAL library for inhibitors of SARS-CoV-2 pseudoparticle entry.** 10,000 Caco2 cells were seeded in 100  $\mu\text{l}$  respective medium in a 96-well flat-bottom plate. The next day, medium was replaced by 40  $\mu\text{l}$  of serum-free medium. For screening peptide containing fractions, 10  $\mu\text{l}$  of the solubilized fraction (in dH<sub>2</sub>O) were added to cells. Cells were inoculated with 50  $\mu\text{l}$  of infectivity normalized LV(Luc)-CoV2 (or LV(Luc)-no GP control). Transduction rates were assessed by measuring luciferase activity in cell lysates at 48 hours post transduction with a commercially available kit (Promega). Values for untreated controls were set to 100 % transduction.

**Gel electrophoresis and western blotting.** Gel electrophoresis of active fractions was performed on a 4–12 % Bis-Tris protein gel (NuPAGE™). Prior to electrophoresis, samples were reduced with 50 mM  $\beta$ -mercaptoethanol and heated for 10 min at  $90^{\circ}\text{C}$ . The gel was either directly stained with Coomassie G-250 (GelCode™ Blue Stain) or blotted on PVDF membranes. Blotted membranes were stained with a polyclonal anti- $\alpha_1$ AT antibody (1:1000; Proteintech 16382-1-AP). After incubation with IRDye anti-rabbit secondary antibodies (1:20,000; LiCor), the staining was visualized using an Odyssey Infrared Imager (Licor).

**Tryptic in-gel digestion of proteins.** Bands of interest were excised and the proteins were digested with trypsin. Tryptic peptides were eluted from the gel slices with 1 % trifluoroacetic acid.

**Matrix-assisted laser-desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS).** MALDI-TOF-MS was performed on an Ultraflex TOF/TOF mass spectrometer (Bruker Daltonics, Bremen) equipped with a nitrogen laser and a LIFT-MS/MS facility. The instrument was operated in the positive-ion reflectron mode using 2,5-dihydroxybenzoic acid and methylendiphosphonic acid as matrix. Sum spectra consisting of 200–400 single spectra were acquired. For data processing and instrument control the Compass 1.4 software package consisting of FlexControl 4.4, FlexAnalysis 3.4 4, Sequence Editor and BioTools 3.2 and ProteinScape 3.1. were used. External calibration was performed with a peptide standard (Bruker Daltonics).

**Database search.** Proteins were identified by MASCOT peptide mass fingerprint search (<http://www.matrixscience.com>) using the Uniprot Human database (version 20200226, 210438 sequence entries;  $p < 0.05$ ). For the search, a mass tolerance of 75 ppm was allowed and oxidation of methionine as variable modification was used.

**Pseudoparticle inhibition experiments.** One day prior to transduction, 10,000 Caco2 cells were seeded in 100  $\mu\text{l}$  respective medium in a 96-well plate. For the addition of  $\alpha_1$ AT prior to infection, medium was replaced by 80  $\mu\text{l}$  of serum-free medium and cells were incubated with serial dilutions of Prolastin ( $\alpha_1$ AT) for 0, 1, 2, and 4 h at  $37^{\circ}\text{C}$  followed by infection with 20  $\mu\text{l}$  of infectivity normalized VSV (Luc)-CoV-2-S pseudoparticles. To investigate whether  $\alpha_1$ AT acts post viral entry, cells were inoculated with 20  $\mu\text{l}$  of infectivity normalized VSV(Luc)-CoV-2 pseudoparticles. After 2 h, cells were washed with 100  $\mu\text{l}$  PBS and 100  $\mu\text{l}$  serum-free medium as well as 20  $\mu\text{l}$  of serially diluted  $\alpha_1$ AT were added. Transduction rates were assessed by measuring luciferase activity in cell lysates at 16 h post transduction with a commercially available kit (Promega). Values for untreated controls were set to 100 % transduction.

**TMPRSS2 activity measurement.** 20,000 HEK 293T cells were seeded in 100  $\mu\text{l}$  of the respective medium in a 96-well flat-bottom plate. The next day, cells were

transfected with 100 ng of TMPRSS2 (addgene 53887, kindly provided by Roger Reeves, Johns Hopkins University, Baltimore, United States) per well using PEI transfection reagent. Briefly, plasmid DNA was mixed with PEI in a 3:1 ratio in serum-free medium, incubated for 20 min at RT and added to cells dropwise. At 12 h post transfection, medium was removed and 60  $\mu$ l of PBS were added, followed by serial dilutions of Prolastin or inhibitor control. After incubation for 15 min at 37 °C, 20  $\mu$ l of protease substrate BOC-Gln-Ala-Arg-AMC were added. Fluorescence intensity was recorded at an excitation wavelength of 380 nm and emission wavelength of 460 nm in 1 min intervals for 2 h at 37 °C in a Synergy<sup>™</sup> H1 microplate reader (BioTek, USA) with Gen 5 3.04 software. For assessing the activity of recombinant human TMPRSS2, 25  $\mu$ l of serially diluted Prolastin or inhibitor control in assay buffer (50 mM Tris-HCL, 0.154 mM NaCl pH 8.0) were incubated with 25  $\mu$ l of recombinant TMPRSS2 enzyme for 10 min at 37 °C, followed by addition of 20  $\mu$ M BOC-Gln-Ala-Arg-AMC protease substrate. Fluorescence intensity was measured after 3 h at an excitation wavelength of 360 nm and emission wavelength of 465 nm in a Tecan Genios with Magellan V6.4 software.

**Binding free energy calculations.** The Central Limit Free Energy Perturbation (CL-FEP) approach<sup>52</sup> was employed to estimate the binding affinities of the complexes of the enzymes TMPRSS2 and trypsin with  $\alpha_1$ AT. The structure of the trypsin- $\alpha_1$ AT complex reported with PDB ID 1OPH<sup>53</sup> was used for the calculations, while the complex TMPRSS2- $\alpha_1$ AT was modeled via docking and subjected to further structural refinement, see computational details in supplementary materials. CL-FEP allows an unbiased end-state calculation of free energy changes directly from explicit solvent simulations. The sampling was performed with molecular dynamics simulations of the individual proteins, the enzyme- $\alpha_1$ AT complexes and the bulk solvent. The proteins were sampled under wall-type restraints on their bound-state conformations, which permits to focus the sampling on the most relevant states to the binding energy. The force constant for these restraints was 100 kcalmol<sup>-1</sup> Å<sup>-2</sup> and a restraint-free range of 5 Å was allowed with respect to the initial conformation. 100 ns of sampling simulations were collected for each simulation box, with the energy sampled every 5 ps. The simulation boxes and molecular dynamics setup were obtained using the CLFEP-GUI web server (<https://clfep.zmb.uni-due.de/>). The CL-FEP analyses were performed with ten checkpoints containing increasing fractions of the total energy samples, an oversampling ratio of  $osr = 20$  was fixed to bring the free energy variance until the level of (kT)<sup>2</sup>, and the second-order cumulant estimator (C2) was evaluated to determine the free energy changes at each checkpoint. The final estimate corresponds to the average among the converged checkpoints. The error is obtained as the standard deviation among the individual estimations. The convergence analyses<sup>52</sup> for all the checkpoints are summarized in Supplementary Tables 2 and 3.

**Surface plasmon resonance.** SPR assay was carried out on a Reichert SR7000 system using SPRAutolink 1.1.9 software. To measure the binding affinity between TMPRSS2 and  $\alpha_1$ AT, TMPRSS2 (Creative BioMart) was immobilized on a 11-mercaptopentadecanoic acid (MUA) functionalized gold sensor chip through covalent coupling of amine groups. The surface was activated using 0.4 M N-hydroxysuccinimide (NHS) and 0.1 M 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) coupling reagents and the TMPRSS2 protein was coupled at a concentration of 12.5  $\mu$ g/ml in 10 mM acetate buffer at pH = 5.5. Residual activated carboxylic groups were blocked using a 1 M ethanolamine solution at pH = 8.5.  $\alpha_1$ AT was dissolved at different concentrations (0, 0.2, 0.5, 1, 2, 5, 10, and 20  $\mu$ M) in HEPES buffer (10 mM HEPES, 150 mM NaCl, 0.05 % Tween-20; pH = 7.5) and introduced with a flow rate of 10  $\mu$ l/min. The data were analyzed with Prism 7 using non-linear regression fit and the value for the equilibrium dissociation constant  $K_d$  was obtained by fitting the response at equilibrium against the concentration. Successful immobilization of TMPRSS2 was verified using a TMPRSS2 antibody (rabbit anti-human IgG, ThermoFischer Scientific, PA5-14264) dissolved in HEPES buffer to a concentration of 0.2 mg/ml (1:10 of 2 mg/ml stock). As a negative control, TMPRSS2 antibody solution was passed through a channel with no TMPRSS2 protein immobilization.

**Cytotoxicity assay.** To assess cytotoxicity of  $\alpha_1$ AT and camostat mesylate (CM), 10,000 Caco2 cells were seeded in 100  $\mu$ l medium in a 96-well flat-bottom plate. The next day, medium was replaced by 80  $\mu$ l of serum-free Caco2 medium and cells were treated with serial dilutions of Prolastin, CM or DMSO as solvent control for CM. After 48 h, cell viability was assessed by measuring ATP levels in cells lysates with a commercially available kit (CellTiter-Glo<sup>®</sup>, Promega).

**Virus-induced cytopathic effect analysis.** To quantify SARS-CoV-2 wildtype infection, virus-induced cell death was inferred from remaining cell viability determined by MTS (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium) assay. To this end, 20,000 TMPRSS2-expressing Vero E6 or 30,000 Caco2 cells were seeded in 96-well plates in 100  $\mu$ l respective medium. The next day, medium was replaced with serum-free medium and the respective compound of interest was added. After incubation for 1 h at 37 °C the cells were infected with a MOI of 0.001 of the SARS-CoV-2 BetaCoV/France/IDF0372/2020 or BetaCoV/Netherlands/01/NL/2020, in a total volume of 180  $\mu$ l.

2 (TMPRSS2-expressing Vero E6) or 3 (Caco2) days post infection, infection was quantified by detecting remaining metabolic activity. To this end, 36  $\mu$ l of CellTiter 96<sup>®</sup> Aqueous One Solution Reagent (Promega G3580) were added to the medium and incubated for 3 h at 37 °C. Then, optical density (OD) was recorded at 620 nm using an Asys Expert 96 UV microplate reader (Biochrom) with DigiRead 1.26 software. To determine infection rates, sample values were subtracted from untreated control and untreated control set to 100 %. For time of addition experiments, 18,000 TMPRSS2-expressing Vero E6 cells were seeded in 100  $\mu$ l respective medium. The next day, respective compounds were added 1 h prior to, simultaneously with, or 2, 4 and 24 h post infection with SARS-CoV-2 (BetaCoV/France/IDF0372/2020) at a MOI of 0.01. Viral inoculum was removed 2 h post infection and cells were washed twice with PBS. Cell viability was assessed at 2 days post infection by MTS assay as described above.

**Plaque assay.** 700,000 TMPRSS2-expressing Vero E6 cells were seeded in 12-well plates in 1 ml of the respective medium. The next day, medium was refreshed and cells were treated with respective compound 1 h prior to, simultaneously with or 1.5 h post infection with SARS-CoV-2 (BetaCoV/Netherlands/01/NL/2020). After infection, cells were overlaid with respective medium supplemented with cellulose. At 2 days post infection, cells were fixed in 4 % paraformaldehyde in PBS for 45 min, washed one with PBS and stained with 0.5 % crystal violet in 0.1 % triton in H<sub>2</sub>O for 30 min. Staining solution was removed and cells were washed three times with water. Virus-induced plaques were quantified in ImageJ 1.53c by assessing number of pixels occupied by plaques.

**Infection of small airway epithelial cells.** 30,000 SAEC were seeded in 100  $\mu$ l of respective medium in 96-well flat bottom plate. The next day, medium was refreshed and Prolastin was added 1 h prior to or 3 h and 24 h post infection with SARS-CoV-2 (BetaCoV/France/IDF0372/2020) at a MOI of 1 for 3 h. Cells were washed three times with PBS and fresh medium was added. Immediately after the wash step (0 dpi) and at 6 dpi, supernatant samples were collected for qPCR analysis. To analyze virus replication by RT-qPCR, RNA was isolated using the QIAamp Viral RNA Mini Kit (Qiagen) according to the manufacturer's instructions. For RT-qPCR, samples were thawed and 5  $\mu$ l of lysate used as sample in a 20  $\mu$ l reaction using Fast Virus 1-Step Mastermix (Thermo Fisher, # 4444436), 0.5  $\mu$ M Taqman primers targeting SARS-CoV-2-ORF1b-*nsp14* and 0.25  $\mu$ M probe (for primer sequences see Supplementary Table 4). Cycling conditions in were as follows: 1 cycle of reverse transcription (50 °C, 300 s) and RT-inactivation (95 °C, 20 s); 40 cycles of denaturation (95 °C, 5 s) and extension (60 °C, 30 s) in a Step One Plus qPCR cycler (AppliedBiosystems) with Step One Software 2.3. RNA isolated from virus stocks of BetaCoV/France/IDF0372/2020 with copy numbers previously estimated using synthetic SARS-CoV-2 RNA standard (Twist Bioscience, #102024) was used to determine genome copies from Ct values.

**Generation of human airway epithelial cells.** Differentiated air-liquid interface cultures of human airway epithelial cells (HAECs) were generated from primary human basal cells isolated from airway epithelia. Cells were expanded in a T75 flask (Sarstedt) in Airway Epithelial Cell Basal Medium supplemented with Airway Epithelial Cell Growth Medium SupplementPack (both Promocell). Growth medium was replaced every two days. Upon reaching 90 % confluence, HAECs were detached using DetachKIT (Promocell) and seeded into 6.5 mm Transwell filters (Corning Costar). Filters were precoated with Collagen Solution (StemCell Technologies) overnight and irradiated with UV light for 30 min before cell seeding for collagen crosslinking and sterilization.  $3.5 \times 10^4$  cells in 200  $\mu$ l growth medium were added to the apical side of each filter, and an additional 600  $\mu$ l of growth medium was added basolaterally. The apical medium was replaced after 48 h. After 72–96 h, when cells reached confluence, the apical medium was removed and basolateral medium was switched to differentiation medium. Differentiation medium consisted of a 1:1 mixture of DMEM-H and LHC Basal (Thermo Fisher) supplemented with Airway Epithelial Cell Growth Medium SupplementPack and was replaced every 2 days. Air-lifting (removal of apical medium) defined day 0 of air-liquid interface (ALI) culture, and cells were grown at ALI conditions until experiments were performed at day 25–28. To avoid mucus accumulation on the apical side, HAEC cultures were washed apically with PBS for 30 min every 3 days from day 14 onwards.

**SARS-CoV-2 infection of HAECs.** Immediately before infection, the apical surface of HAECs grown on Transwell filters were washed three times with 200  $\mu$ l PBS to remove accumulated mucus. Then, 10  $\mu$ M of  $\alpha_1$ AT or 5  $\mu$ M remdesivir were added into the basal medium and onto the apical surface. Cells were infected with  $9.25 \times 10^2$  plaque-forming units (PFU) of SARS-CoV-2 (BetaCoV/France/IDF0372/2020). After incubation for 2 h at 37 °C, viral inoculum was removed and cells were washed three times with 200  $\mu$ l PBS and again cultured at the air-liquid interface. At 1, 2, and 3 days post infection, cells were fixed for 30 min in 4 % paraformaldehyde in PBS, permeabilized for 10 min with 0.2 % saponin and 10 % FCS in PBS, washed twice with PBS and stained with anti-SARS-CoV-2 spike (ab252690, Abcam) and anti-alpha-tubulin (MA1-8007, Thermo Scientific) diluted 1:300 to 1:500, respectively, in PBS, 0.2 % saponin and 10 % FCS over night at 4 °C.

Subsequently, cells were washed twice with PBS and incubated for 1 h at room temperature in PBS, 0.2 % saponin and 10 % FCS containing AlexaFluor 488-labeled anti-rabbit and AlexaFluor 647-labeled anti-rat secondary antibody, respectively (all 1:500; Thermo Scientific) and DAPI + phalloidin AF 405 (1:5000; Thermo Scientific). Images were taken on an inverted confocal microscope (Leica TCS SP5, Leica Microsystems, Leica application suite version 2.7.3.9723) using a  $\times 40$  lens (Leica HC PL APO CS2 40  $\times$  1.25 OIL). Images for the blue (DAPI), green (AlexaFluor 488) and far-red (AlexaFluor 647) channels were taken in sequential mode using appropriate excitation and emission settings that were kept constant for all the acquisitions. For quantification, randomly chosen locations in each filter were selected and z-stacks were acquired. A maximum z projection was performed and anti-SARS-CoV-2 positive cells per area (0.15 mm<sup>2</sup>) were visually identified and counted.

**Non-linear regression and statistics.** Unless stated otherwise, analysis was performed using GraphPad Prism version 8.4.2. Calculation of IC<sub>50</sub> values via non-linear regression was performed using normalized response-variable slope equation. For statistical analysis, either ordinary one-way ANOVA or 2-way ANOVA with Dunnett's multiple comparison test was used as indicated in respective figure legends.

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

### Data availability

Proteins were identified by MASCOT peptide mass fingerprint search (<http://www.matrixscience.com>) using the Uniprot Human database (version 20200226, 210438 sequence entries;  $p < 0.05$ ), Human Uniprot database, P01009 (A1AT\_HUMAN). Crystal structures were obtained from SWISSMODEL repository (<https://swissmodel.expasy.org/repository/uniprot/O15393?csm=C05B5531C8A311C7>) or Protein Data Bank with accession codes PDB ID: 1OPH, PDB ID: 1Z8G, PDB ID: 3CWM. Source data are provided with this paper.

Received: 16 July 2020; Accepted: 18 February 2021;

Published online: 19 March 2021

### References

- Hou, Y. J. et al. SARS-CoV-2 reverse genetics reveals a variable infection gradient in the respiratory tract. *Cell* **182**, e14 (2020).
- Tay, M. Z., Poh, C. M., Rénia, L., MacAry, P. A. & Ng, L. F. P. The trinity of COVID-19: immunity, inflammation and intervention. *Nat. Rev. Immunol.* **20**, 363–374 (2020).
- Gupta, A. et al. Extrapulmonary manifestations of COVID-19. *Nat. Med.* **26**, 1017–1032 (2020).
- Vareille, M., Kieninger, E., Edwards, M. R. & Regamey, N. The airway epithelium: soldier in the fight against respiratory viruses. *Clin. Microbiol. Rev.* **24**, 210–229 (2011).
- Rogan, M. P. et al. Antimicrobial proteins and polypeptides in pulmonary innate defence. *Respir. Res.* **7**, 1–11 (2006).
- Münch, J., Ständker, L., Forssmann, W. G. & Kirchhoff, F. Discovery of modulators of HIV-1 infection from the human peptidome. *Nat. Rev. Microbiol.* **12**, 715–722 (2014).
- Münch, J. et al. Discovery and optimization of a natural HIV-1 entry inhibitor targeting the gp41 fusion peptide. *Cell* **129**, 263–275 (2007).
- Borst, E. M. et al. A peptide inhibitor of cytomegalovirus infection from human hemofiltrate. *Antimicrob. Agents Chemother.* **57**, 4751–4760 (2013).
- Groß, R. et al. A placenta derived C-terminal fragment of  $\beta$ -hemoglobin with combined antibacterial and antiviral activity. *Front. Microbiol.* **11**, 1–17 (2020).
- Forssmann, W. G. et al. Short-term monotherapy in HIV-infected patients with a virus entry inhibitor against the gp41 fusion peptide. *Sci. Transl. Med.* **2**, 63re3 (2010).
- Hoffmann, M. et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **181**, e8 (2020).
- Xia, S. et al. A pan-coronavirus fusion inhibitor targeting the HR1 domain of human coronavirus spike. *Sci. Adv.* **5**, eaav4580 (2019).
- Carrell, R. W. et al. Structure and variation of human  $\alpha$ 1-antitrypsin. *Nature* **298**, 329–334 (1982).
- Travis, J. & Salvesen, G. S. Human plasma proteinase inhibitors. *Annu. Rev. Biochemistry* **52**, 655–709 (1983).
- Janciauskiene, S. & Welte, T. Well-known and less well-known functions of Alpha-1 antitrypsin: Its role in chronic obstructive pulmonary disease and other disease developments. *Ann. Am. Thorac. Soc.* **13**, S280–S288 (2016).
- Gadek, J. E., Klein, H. G., Holland, P. V. & Crystal, R. G. Replacement therapy of alpha 1-antitrypsin deficiency. Reversal of protease-antiprotease imbalance within the alveolar structures of PiZ subjects. *J. Clin. Invest.* **68**, 1158–1165 (1981).
- Kawase, M., Shirato, K., van der Hoek, L., Taguchi, F. & Matsuyama, S. Simultaneous treatment of human bronchial epithelial cells with serine and cysteine protease inhibitors prevents severe acute respiratory syndrome coronavirus entry. *J. Virol.* **86**, 6537–6545 (2012).
- Baric, R. S. Emergence of a Highly Fit SARS-CoV-2 Variant. *N. Engl. J. Med.* **383**, 2684–2686 (2020).
- Korber, B. et al. Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. *Cell* **182**, e19 (2020).
- Bar-On, Y. M., Flamholz, A., Phillips, R. & Milo, R. Sars-cov-2 (Covid-19) by the numbers. *Elife* **9**, e57309 (2020).
- Pizzorno, A. et al. Characterization and Treatment of SARS-CoV-2 in Nasal and Bronchial Human Airway Epithelia. *Cell Reports Med.* **1**, 100059 (2020).
- Krüger, J. et al. Drug Inhibition of SARS-CoV-2 Replication in Human Pluripotent Stem Cell-Derived Intestinal Organoids. *Cell. Mol. Gastroenterol. Hepatol.* **11**, 935–948 (2020).
- Park, T. J., Mitchell, B. J., Abitua, P. B., Kintner, C. & Wallingford, J. B. Dishevelled controls apical docking and planar polarization of basal bodies in ciliated epithelial cells. *Nat. Genet.* **40**, 871–879 (2008).
- Azouz, N. P., Klingler, A. M. & Rothenberg, M. E. Alpha 1 antitrypsin is an inhibitor of the SARS-CoV2-priming protease TMPRSS2. *bioRxiv* **2**, 2020.05.04.077826. Preprint at <https://doi.org/10.1101/2020.05.04.077826> (2020).
- Stratikos, E. & Gettins, P. G. W. Major proteinase movement upon stable serpin-proteinase complex formation. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 453–458 (1997).
- Oguntuyo, K. Y. et al. In plain sight: The role of alpha-1-antitrypsin in COVID-19 pathogenesis and therapeutics. *bioRxiv*. Preprint at <https://doi.org/10.1101/2020.08.14.248880>.
- McElvaney, O. J. et al. Characterization of the inflammatory response to severe COVID-19 illness. *Am. J. Respir. Crit. Care Med.* <https://doi.org/10.1164/rccm.202005-1583oc> (2020).
- Wewers, M. D., Casolaro, M. A. & Crystal, R. G. Comparison of alpha-1-antitrypsin levels and antineutrophil elastase capacity of blood and lung in a patient with the alpha-1-antitrypsin phenotype null-null before and during alpha-1-antitrypsin augmentation therapy. *Am. Rev. Respir. Dis.* **135**, 539–543 (1987).
- Crystal, R. G.  $\alpha$ 1-Antitrypsin deficiency, emphysema, and liver disease: genetic basis and strategies for therapy. *J. Clin. Invest.* **85**, 1343–1352 (1990).
- Hubbard, R. C. & Crystal, R. G. Alpha-1-antitrypsin augmentation therapy for alpha-1-antitrypsin deficiency. *Am. J. Med.* **84**, 52–62 (1988).
- Wilson, S. et al. The membrane-anchored serine protease, TMPRSS2, activates PAR-2 in prostate cancer cells. *Biochem. J.* **388**, 967–972 (2005).
- Lucas, J. M. et al. The androgen-regulated protease TMPRSS2 activates a proteolytic cascade involving components of the tumor microenvironment and promotes prostate cancer metastasis. *Cancer Disco.* **4**, 1310–1325 (2014).
- Ko, C. J. et al. Androgen-induced TMPRSS2 activates matriptase and promotes extracellular matrix degradation, prostate cancer cell invasion, tumor growth, and metastasis. *Cancer Res.* **75**, 2949–2960 (2015).
- Gierer, S. et al. 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 (2013).
- Bertram, S. et al. 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 (2013).
- Matsuyama, S. et al. Efficient activation of the severe acute respiratory syndrome coronavirus spike protein by the transmembrane protease TMPRSS2. *J. Virol.* **84**, 12658–12664 (2010).
- Hatesuer, B. et al. Tmprss2 is essential for influenza H1N1 virus pathogenesis in mice. *PLoS Pathog.* **9**, 1–10 (2013).
- Böttcher, E. et al. Proteolytic activation of influenza viruses by serine proteases TMPRSS2 and HAT from human airway epithelium. *J. Virol.* **80**, 9896–9898 (2006).
- Limburg, H. et al. TMPRSS2 is the major activating protease of influenza A virus in primary human airway cells and influenza B virus in human type II pneumocytes. *J. Virol.* **93**, 649–668 (2019).
- Greene, C. M. et al. Alpha-1 antitrypsin deficiency: a conformational disease associated with lung and liver manifestations. *J. Inherit. Metab. Dis.* **31**, 21–34 (2008).
- Hu, J. et al. The D614G mutation of SARS-CoV-2 spike protein enhances viral infectivity. *bioRxiv* 2020.06.20.161323. Preprint at <https://doi.org/10.1101/2020.06.20.161323>.
- Bhattacharyya, C. et al. Global Spread of SARS-CoV-2 Subtype with Spike Protein Mutation D614G is Shaped by Human Genomic Variations that Regulate Expression of TMPRSS2 and MX1 Genes. *bioRxiv*

- 2020.05.04.075911. Preprint at <https://doi.org/10.1101/2020.05.04.075911> (2020).
43. Barnes, B. J. et al. Targeting potential drivers of COVID-19: neutrophil extracellular traps. *J. Exp. Med.* <https://doi.org/10.1084/jem.20200652> (2020).
  44. Leppkes, M. et al. Vascular occlusion by neutrophil extracellular traps in COVID-19. *EBioMedicine* **58**, 102925 (2020).
  45. Hubbard, R. C. et al. Recombinant DNA-produced  $\alpha$ 1-antitrypsin administered by aerosol augments lower respiratory tract antineutrophil elastase defenses in individuals with  $\alpha$ 1-antitrypsin deficiency. *J. Clin. Invest.* **84**, 1349–1354 (1989).
  46. Campos, M. A. et al. Safety and pharmacokinetics of 120 mg/kg versus 60 mg/kg weekly intravenous infusions of alpha-1 proteinase inhibitor in alpha-1 antitrypsin deficiency: a multicenter, randomized, double-blind, crossover study (SPARK). *COPD* **10**, 687–695 (2013).
  47. Campos, M. A. et al. The biological effects of double-dose alpha-1 antitrypsin augmentation therapy a pilot clinical trial. *Am. J. Respir. Crit. Care Med.* **200**, 318–326 (2019).
  48. Hubbard, R. C., Sellers, S., Czernski, D., Stephens, L. & Crystal, R. G. Biochemical efficacy and safety of monthly augmentation therapy for  $\alpha$ 1-antitrypsin deficiency. *JAMA* **260**, 1259–1264 (1988).
  49. de Loyola, M. B. et al. Alpha-1-antitrypsin: a possible host protective factor against Covid-19. *Rev. Med. Virol.* <https://doi.org/10.1002/rmv.2157> (2020).
  50. Bai, X. et al. Hypothesis: Alpha-1-antitrypsin is a promising treatment option for COVID-19. *Med. Hypotheses* <https://doi.org/10.1016/j.mehy.2020.110394> (2020).
  51. Rentsch, M. B. & Zimmer, G. A vesicular stomatitis virus replicon-based bioassay for the rapid and sensitive determination of multi-species type I interferon. *PLoS ONE* **6**, e25858 (2011).
  52. Ruiz-Blanco, Y. B. & Sanchez-Garcia, E. CL-FEP: an end-state free energy perturbation approach. *J. Chem. Theory Comput.* **16**, 1396–1410 (2020).
  53. Dementiev, A., Simonovic, M., Volz, K. & Gettins, P. G. W. Canonical inhibitor-like interactions explain reactivity of 1-proteinase inhibitor Pittsburgh and antithrombin with proteinases. *J. Biol. Chem.* <https://doi.org/10.1074/jbc.M305195200>. (2003).

## Acknowledgements

L.W., C.C., T.W., R.G., M.H., and A.S. are part of and R.G. is funded by a scholarship from the International Graduate School in Molecular Medicine Ulm. This work was supported by the German Research Foundation (DFG) through “Fokus-Förderung COVID-19” grants to J.M. (MU 3115-13), A.K. (KL 2544/8-1) and M.F. (FR 2974/3-1), the CRC 1279 to S.S., F.K., E.S.-G., and J.M., SPP1923 to F.K. and K.S., and SP1600/4-1 to K.S. J.M. also acknowledges funding by the EU’s Horizon 2020 research and innovation programme (Fight-nCoV, 101003555 to J.M.). J.A.M. is indebted to the Baden-Württemberg Stiftung for the financial support by the Eliteprogramme for Postdocs and receives funding by the DFG. S.P. was supported by BMBF (RAPID Consortium, 01KI1723D and 01KI2006D; RENACO, 01KI20328A, 01KI20396). E.S.-G. further acknowledges the Boehringer Ingelheim Foundation (Plus-3 Grant) and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Federal and State Excellence Strategy EXC- 2033 Projektnummer 390677874 and the Collaborative Research Center CRC 1093. J.A.M., J.M., F.K., M.F. and A.K. further acknowledge funding by the Sonderfördermassnahme COVID-19 of the MWK Baden-Württemberg, Germany, and A.K. by the Heisenberg-Programm (KL 2544/6-1). K.S. was supported by the BMBF (IMMUNOMOD, 01KI2014).

## Author contributions

L.W. generated LV pseudotypes, performed screening, gel electrophoresis,  $\alpha$ 1AT inhibition studies with pseudotypes and in-cell TMPRSS2 activity measurements; C.C., T.W., and J.A.M. generated SARS-CoV-2 stocks and performed all infection experiments in the BSL-3; F.Z., C.P.B., R.G., and A.S. generated VSV pseudotypes; T.W. supported L.W. in most experiments; M.H. performed western blots; C.C. and S.K. supported L.W. in screening; N.P. synthesized EK1, generated peptide libraries and purification; G.F. and R.L. generated HAEC and did staining; P.K. and V.M. performed SPR and provided reagents and controls; Y.B.R.B. and E.S.G. designed, performed and analyzed the results of the computational modeling; S.P. and M.H. performed recombinant TMPRSS2 activity analysis; L.S. supervised the generation of peptide libraries; D.R.T. performed autopsies and collected lungs; C.S. collected BAL; S.S. supervised BSL-3 work; L.W. arranged the figures and L.W., C.C., T.W., J.A.M., and R.G. edited the manuscript; A.K. advised and wrote the manuscript; G.L. performed mass spectrometry; K.S. advised and edited the manuscript; F.K. supervised work and wrote the manuscript; B.M. supervised statistics; M.F. supervised work with HAEC; J.M. is responsible for the study, supervised all work and wrote the manuscript.

## Funding

Open Access funding enabled and organized by Projekt DEAL.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-021-21972-0>.

**Correspondence** and requests for materials should be addressed to J.M.

**Peer review information** *Nature Communications* thanks David Curiel, Brett Hurst and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021