A comprehensive study of SARS-CoV-2 main protease (M^{pro}) inhibitor-resistant mutants selected in a VSV-based system

Francesco Costacurta[1†], Andrea Dodaro[2], David Bante[1], Helge Schöppe[3], Bernhard Sprenger[4], Seyed Arad Moghadasi[5], Jakob Fleischmann[6,7], Matteo Pavan[2], Davide Bassani[2], Silvia Menin[2], Stefanie Rauch[1], Laura Krismer[1], Anna Sauerwein[1], Anne Heberle[1], Toni Rabensteiner[1], Joses Ho[10], Reuben S. Harris[8,9], Eduard Stefan[6,7], Rainer Schneider[4], Teresa Kaserer[3], Stefano Moro[2], Dorothee von Laer[1] and Emmanuel Heilmann[1*]

Affiliations:

[1] Institute of Virology, Medical University of Innsbruck, Innsbruck, 6020, Tyrol, Austria

[2] Molecular Modeling Section (MMS), Department of Pharmaceutical and Pharmacological Sciences, University of Padua, Via F. Marzolo 5, 35131, Padova, Italy

[3] Institute of Pharmacy/Pharmaceutical Chemistry, University of Innsbruck, Innsbruck, 6020, Tyrol, Austria

[4] Department of Biochemistry, University of Innsbruck, Innsbruck, 6020, Austria

[5] Department of Biochemistry, Molecular Biology and Biophysics, Institute for Molecular Virology, University of Minnesota, Minneapolis, MN 55455, United States

[6] Institute of Molecular Biology, University of Innsbruck, Innsbruck, 6020, Tyrol, Austria

[7] Tyrolean Cancer Research Institute (TKFI), Innrain 66, Innsbruck, 6020, Tyrol, Austria

[8] Department of Biochemistry and Structural Biology, University of Texas Health San Antonio, San Antonio, TX 78229, United States

[9] Howard Hughes Medical Institute, University of Texas Health San Antonio, San Antonio, TX 78229, United States

[10] Bioinformatics Institute, Agency for Science Technology and Research, Singapore

* Corresponding author.

Email: emmanuel.heilmann@i-med.ac.at

Short title: VSV-based system selects M^{pro} mutants resistant to nirmatrelvir and ensitrelvir.

Abstract

Nirmatrelvir was the first protease inhibitor (PI) specifically developed against the SARS-CoV-2 main protease (3CL^{pro}/M^{pro}) and licensed for clinical use. As SARS-CoV-2 continues to spread, variants resistant to nirmatrelvir and other currently available treatments are likely to arise. This study aimed to identify and characterize mutations that confer resistance to nirmatrelvir. To safely generate M^{pro} resistance mutations, we passaged a previously developed, chimeric vesicular stomatitis virus (VSV-M^{pro}) with increasing, yet suboptimal concentrations of nirmatrelvir. Using Wuhan-1 and Omicron M^{pro} variants, we selected a large set of mutants. Some mutations are frequently present in GISAID, suggesting their relevance in SARS-CoV-2. The resistance phenotype of a subset of mutations was characterized against clinically available PIs (nirmatrelvir and ensitrelvir) with cell-based and biochemical assays. Moreover, we showed the putative molecular mechanism of resistance based on in silico molecular modelling. These findings have implications on the development of future generation M^{pro} inhibitors, will help to understand SARS-CoV-2 protease-inhibitor-resistance mechanisms and show the relevance of specific mutations in the clinic, thereby informing treatment decisions.

Teaser:

Understanding how SARS-CoV-2 could counter the antiviral drug nirmatrelvir and what it means for the future of COVID-19 treatment.

Keywords:

Antimicrobial resistance; ensitrelvir; nirmatrelvir; non-gain-of-function research; Omicron; protease inhibitor resistance mutations; SARS-CoV-2; selection pressure experiments

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the causative agent of COVID-19 (coronavirus disease-19), has established itself as a permanent human and animal pathogen worldwide. Vaccines, alongside monoclonal antibodies, have drastically reduced hospitalization and/or mortality, especially in immunocompromised individuals, the elderly, and people with pre-existing medical conditions (*1*). As vaccines do not confer complete immunity against infection, the virus continues to spread effectively due to easier host-to-host transmission (*2*), and immune-escaping variants, such as Beta, Delta and Omicron (*3–5*).

In November 2021, the FDA granted emergency use authorization to Paxlovid, an orally administered medication that combines nirmatrelvir, the active component, and ritonavir, a pharmacokinetic enhancer. Nirmatrelvir is a highly potent protease inhibitor against the SARS-CoV-2 $3CL^{pro}/M^{pro}$, and is most relevant in the clinical setting, as it has been shown to decrease hospitalization, in-hospital disease progression and death (*6*). One year after the licensing of Paxlovid, another protease inhibitor, ensitrelvir, was approved in Japan through the emergency regulatory approval system under the commercial name Xocova (*7*), and recently received Fast Track designation by the FDA. Recently, leritrelvir (*8*) has been approved by the China National Medical Products Administration. Other oral M^{pro} inhibitors that have entered clinical trials include bofutrelvir (FB2001) (*9*), pomotrelvir (PBI-0451 (*10*), halted (*11*)), simnotrelvir (*12*), EDP-235 (*13*), and HS-10517 (GDDI-4405). SARS-CoV-2 continues to spread, and it is expected that the use of protease inhibitors could eventually lead to the selection of drug-resistant M^{pro} variants, with serious consequences for individuals who cannot benefit from vaccines, for example due to immune defects or pre-existing comorbidities.

To date, several studies have addressed the issue of nirmatrelvir-resistant/escaping variants using different approaches: highlighting mutational hotspots (14, 15), in silico investigation of specific mutations (16), studying resistance phenotypes (17–20), addressing which mutations are the most prone to decrease inhibitor susceptibility (21–23) or more comprehensive studies describing either mutation resistance, fitness costs, or both (24–27). Notably, a few groups have performed gain-of-function selection experiments using the Wuhan-1 strain of SARS-CoV-2. While this represents the most straightforward system to generate and study drug-resistant mutants, it requires government approval, BLS-3 facilities, and demands absolute caution to avoid biosafety breaches and potential release of these protease inhibitor-resistant variants.

Furthermore, owing to safety concerns, most of the studies that employed live SARS-CoV-2 were not performed with the Omicron variant (17–19). This precaution was taken, because protease resistant pre-Omicron variants would likely be unable to compete with the current Omicron strains, if released accidentally. In a preprint by Lan et al. (*28*), the authors validated the antiviral resistance of certain M^{pro} resistant variants in the Omicron-M^{pro} context by using SARS-CoV-2 replicons, circumventing potential safety issues for mutated transmissible variants. It is worth noting that the Omicron signature mutation (P132H) does not confer resistance to common inhibitors (*29–32*) but alters the thermal stability of the protease, as reported by Sacco et al. (*33*). This alteration, as well as other structural and biochemical features caused by the P132H substitution, could potentially affect the mechanism of resistance development and the relevance of mutations selected with Wuhan-1 M^{pro} for Omicron.

Recently, we described a BSL-2 mutation selection system based on a M^{pro}-dependent chimeric vesicular stomatitis virus (VSV-3CL^{pro}/VSV-M^{pro}) (*34*). Replication of the chimeric VSV was effectively inhibited by nirmatrelvir. Using this system, we were able to select a panel of mutants and characterize them through computational and cellular methods. In the present study, we deepened our understanding of the SARS-CoV-2 M^{pro} resistance mutation landscape by selecting mutations in the Omicron M^{pro} and increasing the resistance phenotypes of an already resistant variant, namely VSV-L167F-M^{pro}. By generating and characterizing the mechanism of Omicron-M^{pro} and L167F-derived mutations, individually or in combination, this study aims to advance the understanding of protease-inhibitor-resistant M^{pro} variants and provide valuable information for treatment decision-making.

Results

A safe method to select protease inhibitor resistant M^{pro} variants

In a previous study, we developed a safe alternative to using un-attenuated SARS-CoV-2 for the selection of inhibitor-resistant variants (*17*, *19*, *34*). The technology is based on a chimeric VSV-M^{pro} that encodes an artificial, non-functional polyprotein (G-M^{pro}-L), and relies on M^{pro} activity for its replication (**Fig. 1A,B**). In the absence of an inhibitor, the protease processes the polyprotein, releasing G and L, and the virus can replicate. However, when a protease inhibitor is applied, the protease is inhibited, and the virus can no longer replicate.

Wuhan-1 and Omicron VSV-M^{pro} variants are equally susceptible to nirmatrelvir

To gain a deeper understanding of the potential of the Omicron-M^{pro} to acquire resistance mutations, we introduced the Omicron-M^{pro} signature mutation P132H into chimeric VSV-M^{pro}, generating VSV-Omicron-M^{pro} (VSV-O-M^{pro}) for subsequent selection experiments with the protease inhibitor nirmatrelvir (**Fig. S1A**). First, nirmatrelvir dose response studies were performed with the Omicron and Wuhan-1 VSV-M^{pro} in the presence of nirmatrelvir and it was found to be equally effective against both, as described previously using WT SARS-CoV-2 (*29–32*) (**Fig. S1B,C**).

VSV-M^{pro} supports the main protease evolution of Omicron variant and variants with existing resistance to escape nirmatrelvir

Following-up on our first study (*34*), a resistant variant of VSV-M^{pro} (VSV-L167F-M^{pro}), which was previously identified in selection experiments, was used. This mutation was of particular interest, because it was also found in resistance studies with authentic SARS-CoV-2 (*17–19*). Then, selection experiments using both VSV-O-M^{pro} and VSV-L167F-M^{pro} were performed. By passaging both VSV-L167F-M^{pro} and VSV-O-M^{pro} in the presence of suboptimal concentrations of nirmatrelvir, M^{pro} mutations that were generated by the error-prone VSV polymerase were selected (*35*, *36*). Samples from the first two pilot experiments with VSV-L167F-M^{pro} and VSV-O-M^{pro} were sequenced via Sanger sequencing. Samples from subsequent selection experiments were deep-sequenced (Nanopore) (**Fig. 1C,D**) on the same target region (G_{Cterm}-M^{pro}-L_{Nterm}) (**Table S2**).

We obtained 29 distinct non-synonymous mutations in double-, triple-, and quadruple mutated VSV-L167F-M^{pro} variants and 47 distinct non-synonymous single- and double-mutated VSV-O-M^{pro} variants, as schematically shown by the unrooted phylogenetic tree (**Fig. 1E**).

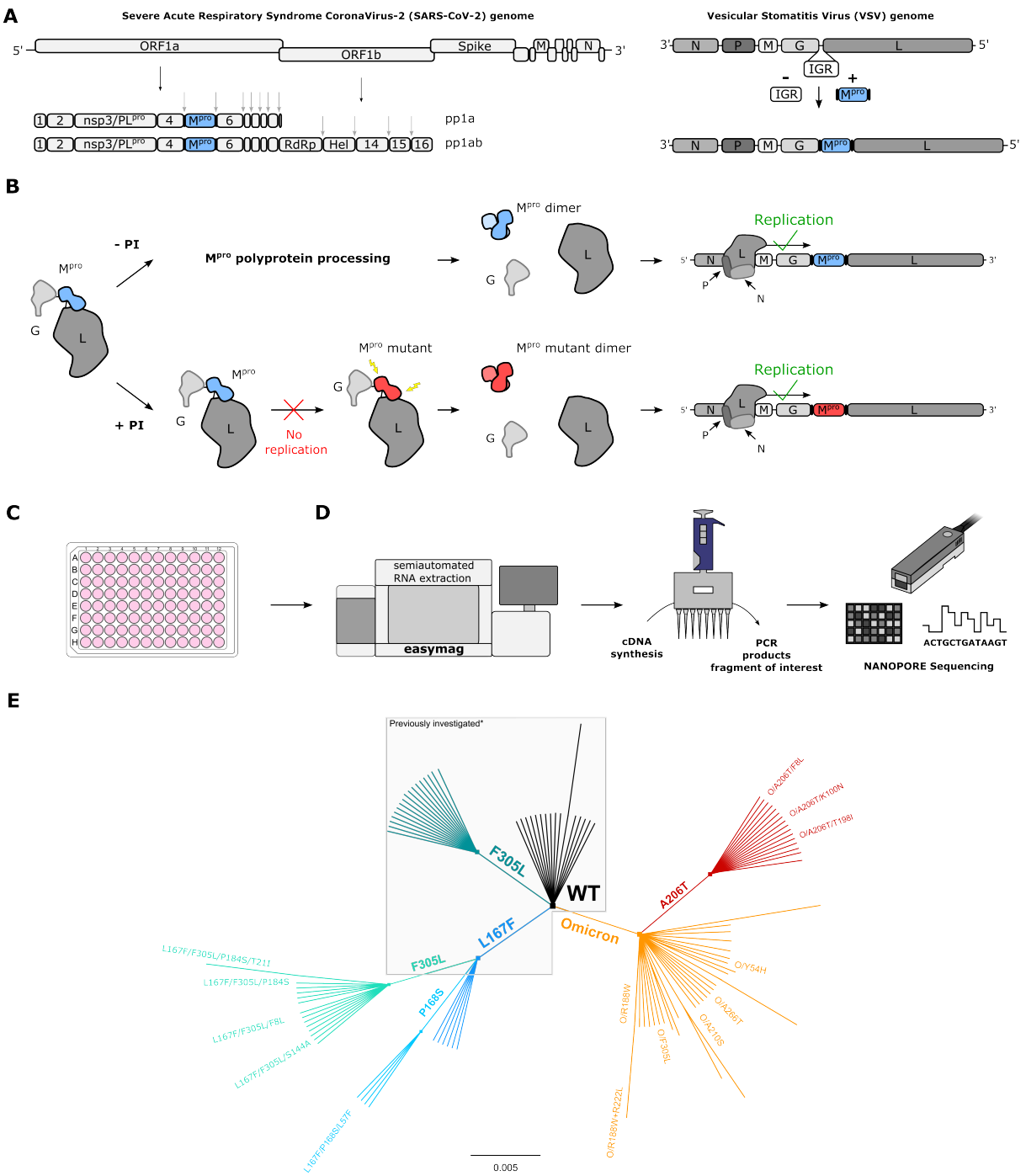

Fig. 1. VSV-G-M^{pro}-L construct: molecular mechanism, sequencing workflow, and mutant lineage phylogeny. (**A**) Schematic representation of SARS-CoV-2 genome, polyproteins pp1a and pp1ab and the VSV-based mutation selection tool. The InterGenic Region (IGR) between proteins G and L of WT VSV was replaced with the SARS-CoV-2 M^{pro} Wuhan-1 sequence and its cognate autocleavage sites. M^{pro} genome positions in SARS-CoV-2 and in VSV-M^{pro} are highlighted in light blue. (**B**) The virus is fully dependent on M^{pro} for replication. Upon translation of G-M^{pro}-L, two outcomes are possible: 1. without an inhibitor, M^{pro} is free to process the polyprotein, and the transcription and replication complexes can assemble; 2. with an inhibitor, M^{pro} is inhibited, the polyprotein is not processed, and the virus is thus not able to replicate, unless it acquires a mutation rendering the M^{pro} less susceptible to the inhibitor. Then, the virus can replicate despite the inhibitor. (**C**) BHK21 cells in a 96-well plate are infected with VSV-M^{pro}. Here, the two outcomes described in panel b can occur. (**D**) Nanopore sequencing workflow. (**E**) Unrooted phylogenetic tree showing the relationship between original/parental viruses and mutants. M^{pro} variants belonging to the same parental virus are coloured accordingly: WT (black), F305L (sea green), L167F (blue), L167F/F305L (light sea green), L167F/P168S (light blue), Omicron (orange), Omicron/A206T (red). For clarity, only the names of the mutants investigated in this work are displayed. *Previously generated/investigated set of mutants in our first study (*34*).

To achieve multiple-mutated viruses, the most interesting mutants were selected according to specific parameters for further selection experiments, where we increased the concentration of the inhibitor, imposing stronger selection pressure at each passage. An overview on the M^{pro}-genome location of the generated mutations is displayed in **Fig. S2**. Several variants were chosen for additional passaging in the presence of the PI according to the following criteria: 1. Proximity of the amino acid substitution to the inhibitor (within 5 Å of the catalytic site or near the catalytic site, within 5 – 10 Å); 2. Prevalence in the GISAID database of specific mutations (*37–39*) (**Fig. 2A,B**), and for further analyses, substitutions whose count in the GISAID database was above 500 entries, were considered (**Fig. 2A**); 3. Frequency of a specific mutation occurring in different samples independently (**Fig. 2C**).

Some mutations combined more than one criterion. For example, some substitutions located in the catalytic site were also frequent in the GISAID repository. However, residues located in this area were generally more conserved (**Fig. S3**) (*40*, *41*), whereas residues outside the catalytic site were more frequent in our selection experiments and more frequent in GISAID. We found ten substitutions with more than 500 entries: T21I, P184S, K100N, T198I, F8L, A234T, A194S, A210S, A266T, and P168S (**Fig. 2A**). Despite not reaching 500 entries in GISAID, we also included the F305L mutation for further study. We recently described F305L as an autocleavage site optimization mutant with similarities to T304I selected in authentic SARS-CoV-2 resistance studies. All mutations occurred during selection experiments, except for T21I and R222L. R222L arose during plaque purification of VSV-Omicron-R188W-M^{pro}. T21I appeared during the generation of VSV-L167F/F305L/P184S from its plasmid (also referred to as "rescue" (*42*)). This virus variant had to be produced from its plasmid, as it could not be plaque purified.

Based on the abovementioned parameters (proximity, GISAID frequency and/or independent occurrences), two VSV-L167F-M^{pro} variants, L167F/F305L and L167F/P168S, were chosen for further selection experiments. From VSV-O-M^{pro}, the variants O/A206T, O/R188W, O/Y54H, O/A210S, and O/A266T were chosen for further study. Among these variants, only O/A206T was selected for additional selection experiments. O/A206T did not fulfil the first two filtering criteria, but only the third (multiple occurrences), as it occurred 17 times independently in VSV-O-M^{pro} selection experiments. To confirm that this variant was not present in our VSV-Omicron-M^{pro} stocks already, Nanopore sequencing was performed on those stocks. We confirmed the absence of any VSV-O-M^{pro} variant subpopulation to the extent of sensitivity that Nanopore sequencing provides. Subsequently, the next round of selection experiments with the variants L167F/F305L and L167F/P168S was carried out, in which the triple-mutant variants L167F/F305L/F8L, L167F/F305L/S144A, L167F/F305L/P184S, and L167F/P168S/L57F were selected. From VSV-O/A206T-M^{pro}, the mutants O/A206T/F8L, O/A206T/K100N and O/A206T/T198I were selected. To assess whether mutants generated from the selection experiments and selected for further analyses were not detrimental for viral propagation, the Ultrafast Sample placement on Existing tRee (UShER) tool was applied. Phylogenetic analyses using sequences harbouring T21I (**Fig. 2C**), T198I, P184S, K100N, and F8L (**Fig. S4A,B**) were generated. We chose these mutations from our mutation pool because they were the most frequently represented in the GISAID database. The prevalence of these mutations before and after the Omicron surge (21st December 2021, up to 18th January 2023) (**Fig. S4C**) has also been examined. From this analysis we could observe that M^{pro}

variants bearing these substitutions appear to not be harmful to the replication and spread of the virus.

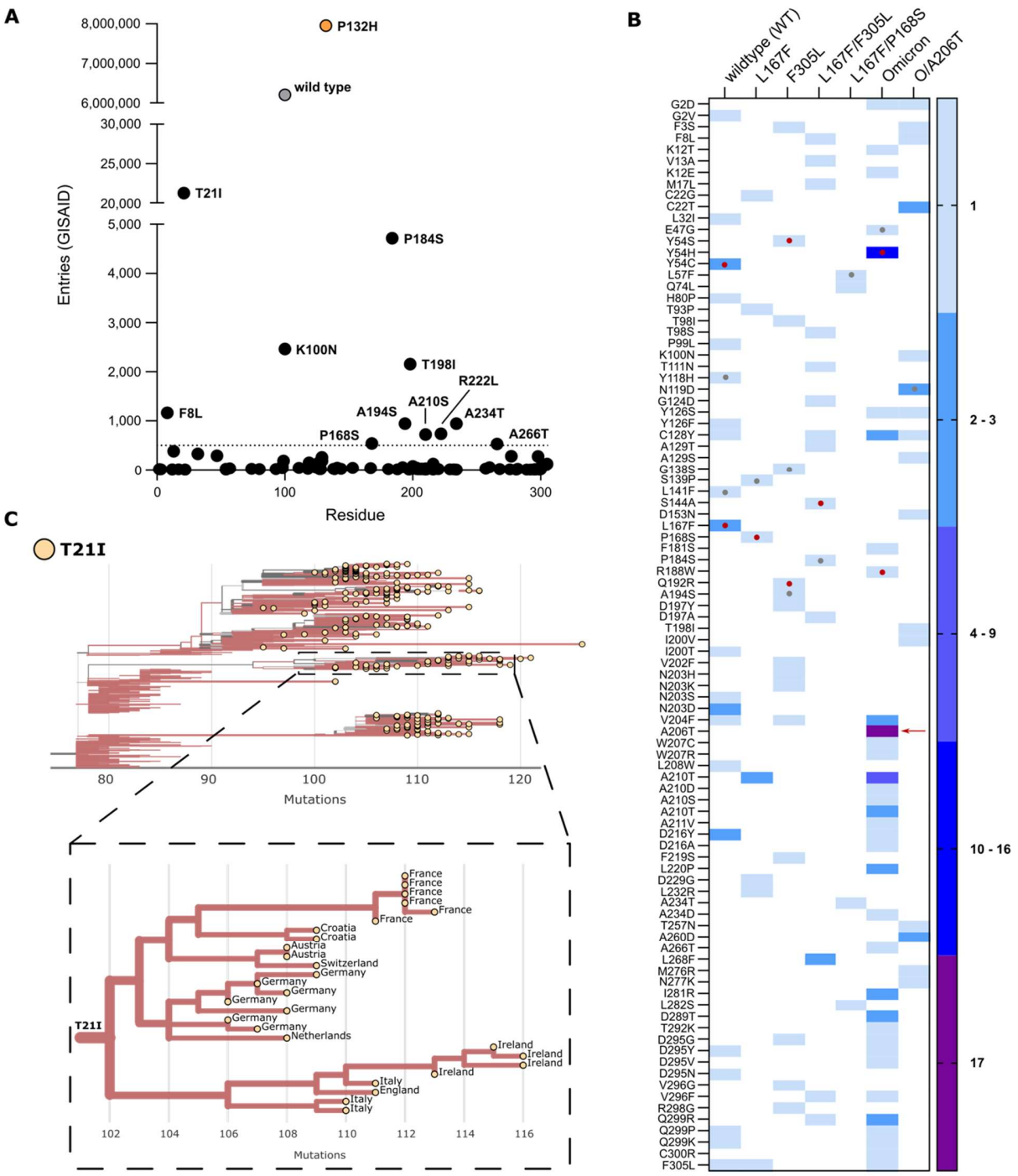

Fig. 2. Frequency of M^{pro} mutations in the GISAID database. (**A**) The total number of nsp5/M^{pro} substitution entries for each mutation found during the selection experiments. WT and Omicron-M^{pro} sequences are displayed in grey and orange, respectively. The dotted line at 500 represent the

cut-off value. (**B**) Heat map representing the number of a specific substitution's independent occurrences from different samples of VSV-M^{pro} selection experiments. The red arrow indicates the substitution A206T. Red dots indicate residues within the catalytic site and grey dots indicate residues near the catalytic site. (**C**) Phylogenetic subtree of nsp5/M^{pro}-T21I mutant generated with the Ultrafast Sample placement on Existing tRee (UShER) tool and magnified view of a subtree area, showing transmission of this variant possibly from a single founder event (9th of April 2023). Only sequences deposited after the Omicron emergence were used.

Mutations selected in VSV-G-M^{pro}-L confer resistance to nirmatrelvir & ensitrelvir

To quantify the resistance of the selected variants, the most interesting mutations were introduced into two previously described live-cell-based protease activity assays, namely 3CL/M^{pro}-On and 3CL/M^{pro}-Off (*43*) (**Fig. S5B,C**). These two protease activity measurement tools rely on replication-incompetent VSV-dsRed variants missing either the phosphoprotein (P) or the polymerase (L), which are replaced with the red fluorescent protein dsRed. Briefly, cells were transfected either with a plasmid encoding the P protein modified with an INTRAmolecular-M^{pro}-tag (P_{Nterm}:M^{pro}:P_{Cterm}) or an INTERmolecular fusion protein made of the green fluorescent protein, M^{pro} and L (GFP-M^{pro}-L). The cells were then infected with VSV-ΔP or VSV-ΔL, respectively. The P_{Nterm}:M^{pro}:P_{Cterm} intramolecular tag in combination with VSV-ΔP-dsRed constitutes a gain-of-signal assay also called "M^{pro}-On", whereas the artificial polyprotein (or fusion protein) GFP-M^{pro}-L in combination with VSV-ΔL-dsRed constitutes a loss-of-signal assay also called "M^{pro}-Off".

First, the M^{pro}-On system was used (**Fig. S5C**) to quantify the resistance phenotype of four triple mutants that arose from the WT protease: L167F/F305L/F8L, L167F/P168S/L57F, L167F/F305L/P184S, and L167F/F305L/S144A, against nirmatrelvir and ensitrelvir (**Fig. 3A**). We also included the related single and double mutants to investigate the contribution of each mutation to the resistance phenotype. Single substitutions conferred low resistance to protease inhibitors, but the resistance increased by combining them. However, there were limitations in quantifying the resistance of some WT-related combinations using the gain-of-signal assay. This was largely observed in triple mutants, where the gain in signal was strongly supressed by the high resistance of these protease variants and IC_{50} values could not be quantified. Since inhibitor concentrations of 100 μM and higher were cytotoxic, thereby interfering with signal generation, we did not increase the inhibitor concentration further. Instead, to attain better resistance quantification, we cloned the same set of WT M^{pro} mutations in the loss-of-signal assay. This cellular system is more sensitive, as the IC_{50} value of the WT M^{pro}-Off is considerably lower than

in M^{pro}-On. Furthermore, the fluorescent signal was measured at different time points (48, 72 and 84 hours), because some variants were replicating slower, affecting final signal quantification (**Fig. 3C**). As expected, while single mutants had a mild effect on M^{pro} susceptibility to nirmatrelvir and ensitrelvir, triple mutants showed increased resistance (**Fig. S6B,D,F**). Initially, the most prominent differences between PIs were observed for L167F/F305L/F8L with 18.4-fold and 40.5-fold, and L167F/P168S/L57F, with 25.8-fold changes and 34.6-fold changes for nirmatrelvir and ensitrelvir (**Fig. 3A** and **Fig. S7**), respectively. When T21I was added to L167F/F305L/P184S, the resistance phenotype substantially increased, with 49.0- and 35.5-fold changes in the IC_{50} values of nirmatrelvir and ensitrelvir, respectively (**Fig. 3A**).

Mutations that arose during selection experiments with the Omicron protease (P132H) were also characterized in the dose-response assays. In **Fig. S8**, non-linear regression analyses of dose-response experiments are shown. IC_{50} values were extrapolated from each curve and calculated the IC_{50} fold-changes for all the variants compared to the parental protease, Omicron-M^{pro} (**Fig. 3B**). In the M^{pro}-On assay, neither single nor combined mutants showed strong resistance compared to the parental protease. We therefore cloned the same set of mutations for loss-of-signal experiments to confirm the phenotype. Some Omicron-M^{pro} variants were measured at different time points to adjust for their slower kinetics. The resistance levels in the M^{pro}-Off system were far greater than in the M^{pro}-On assay, as noticeable from the IC_{50} fold-changes. M^{pro}-O/A206T-Off, M^{pro}-O/A206T/K100N-Off, M^{pro}-O/A206T/T198I-Off, M^{pro}-O/A206T/F8L-Off and M^{pro}-O/Y54H/F305L-Off showed large differences in the susceptibility to the two inhibitors. Nirmatrelvir showed IC_{50} fold-changes of 27.3, 26.0, 75, 100.2 and 11.6, respectively. Moreover, we observed differences between responses against nirmatrelvir and ensitrelvir, in particular for M^{pro}-O/Y54H/F305L-Off, which was more resistant against ensitrelvir with an IC_{50} fold-change of 20.1.

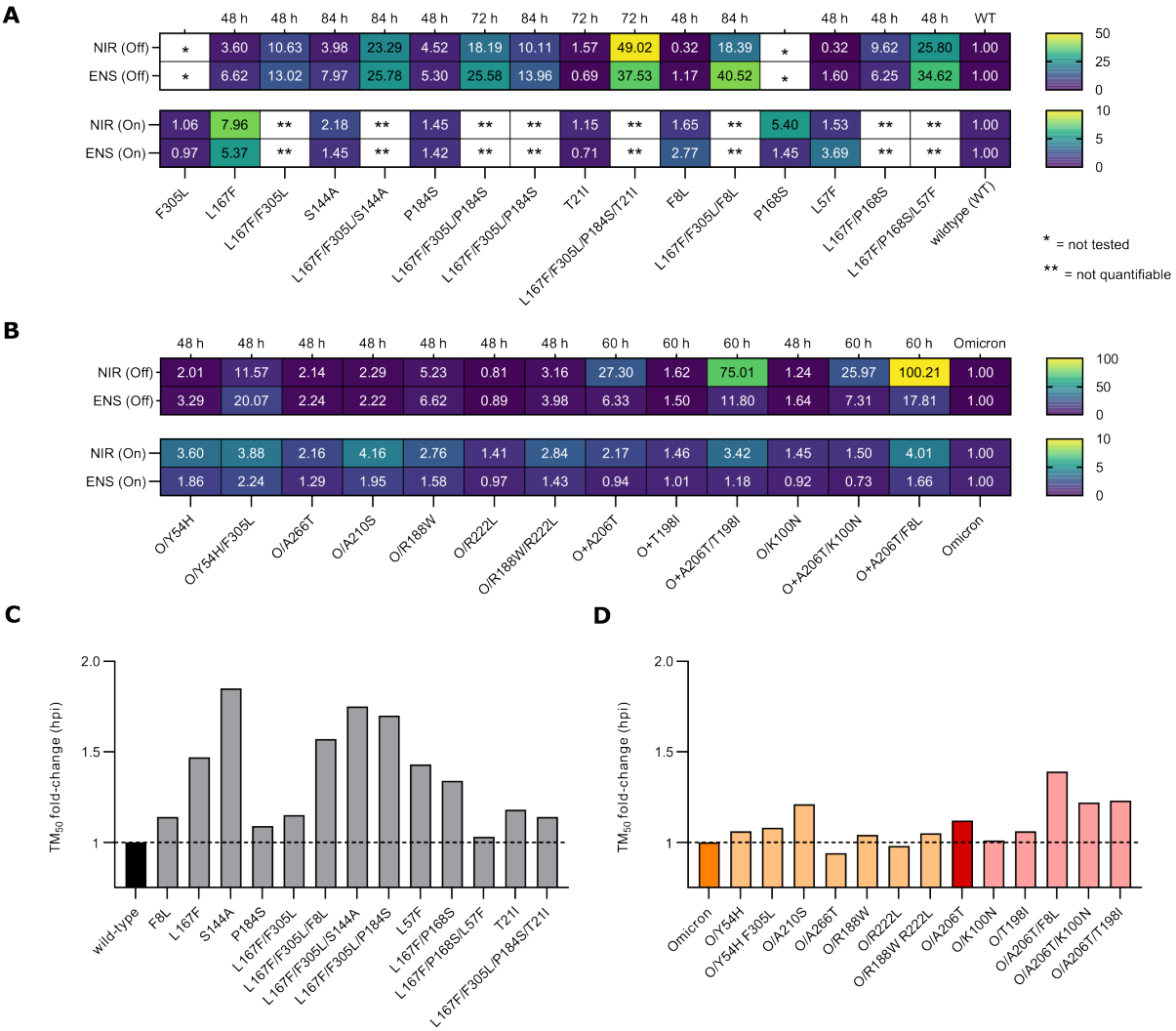

Fig. 3. Resistance data and viral replication kinetics of WT, Omicron and mutant main proteases. Rectangular fields with one and two stars represent constructs/variants that have not been tested or for which IC_{50} values were not quantifiable, respectively. Hours after infection / hours post infection (hpi). (**A**) Heat maps of IC_{50} fold-changes of M^{pro}-On (bottom) and M^{pro}–Off (top) dose responses for nirmatrelvir and ensitrelvir. (**B**) Heat maps of IC_{50} fold-changes for Omicron-M^{pro}-On (bottom) and Omicron-M^{pro}–Off (top) dose responses for nirmatrelvir and ensitrelvir. (**C**) Bar plot of TM_{50} fold-changes vs the WT protease. The WT M^{pro} is coloured in black. (**D**) Bar plot of TM_{50} fold-changes vs Omicron (orange) and Omicron/A206T (bordeaux) proteases. Variants are coloured in light orange and light bordeaux according to the parental protease from which they originated.

Protease inhibitor resistant mutations can alter VSV replication kinetics, which can be restored upon acquisition of compensatory mutations

As described above, some variants, e.g. S144A and triple mutants except L167F/P168S/L57F, did not show strong resistant phenotypes at the time point we typically measure dsRed signal (48 h). To investigate this unexpected result, viral replication was assessed over time using the M^{pro}-Off system. In M^{pro}-Off transfected cells, the signal increases over time in the absence of an inhibitor (**Fig. S9A**). Fluorescence output was measured every 12 h, beginning 12 h after infection (= hours post infection, hpi). Fluorescent signals were then plotted against hpi and non-linear regression analyses were performed, which returned TM_{50} (time maximum fifty), a value that represents the time required to reach the half-maximum value of the curve when the signal plateaus. The WT protease had a TM_{50} between 35-40 hpi (**Fig. S9E**), and the plateau started at 48 hpi, which matched the time point that we generally consider for dose-response experiments in M^{pro}-Off assays. As expected, the WT protease led to the fastest increase in fluorescence signal, and mutants showed either similar or higher TM_{50} values (**Fig. 3C, S9F,G**). This phenomenon indicates that certain resistance mutations can reduce protease activity, which is expected to attenuate virus replication if the mutations would occur in wild-type SARS-CoV-2. In contrast, by adding T21I to M^{pro}-L167F-P184S-F305L-Off resulting in the quadruple-mutant M^{pro}-T21I-L167F-P184S-F305L-Off, a decreased TM_{50} of 14 hours compared to its parental virus was observed (**Fig. S9H**).

We also investigated viral replication kinetics of the selected Omicron-M^{pro} mutants. Overall, mutations that have occurred in the Omicron-M^{pro} context, had a lower impact on viral replication rate than in the WT M^{pro} (**Fig. 3D**). The slowest variant was O/A206T/F8L, with a 1.4 fold-change (~16 hours) increase in the time required to reach the plateau.

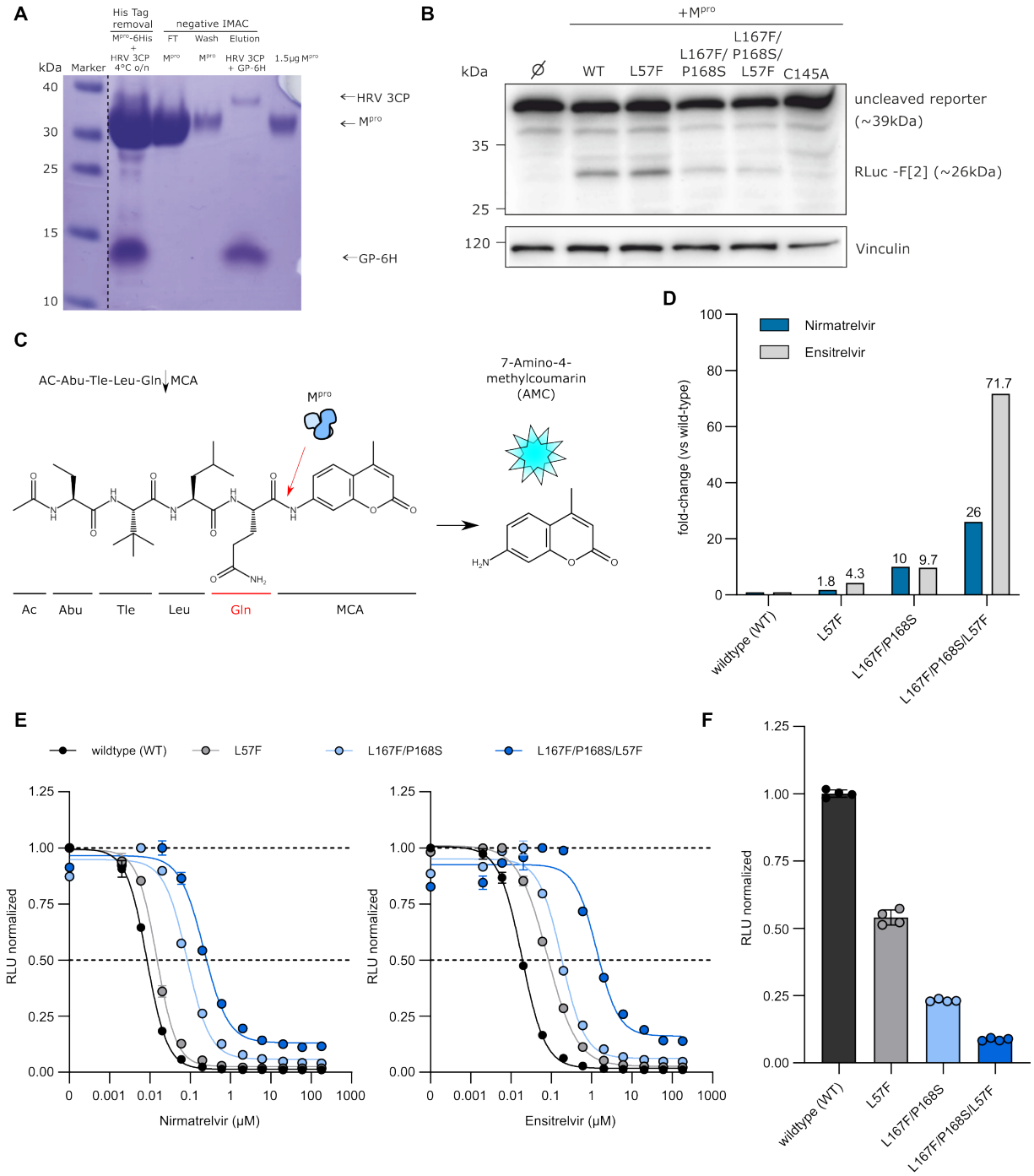

Fig. 4. Biochemical assay dose responses of nirmatrelvir and ensitrelvir against selected mutant recombinant proteases. (**A**) Representative Coomassie-staining for the M^{pro}-L57F variant after FPLC purification and His-tag removal with HRV-3CP followed by a negative Immobilized Metal Affinity Chromatography (IMAC). The dotted line indicated that the marker lane photograph was cut and placed near the production of the recombinant M^{pro} lanes. Lane 1: Tag removal reaction of M^{pro} with HRV-3CP. Combined band of M^{pro} and HRV 3CP at 35 kDa and free 6xHis-tag (GP-6H) at around 10 kDa. Lane 2: Flow-through (FT) of negative IMAC, yields the M^{pro}-L57F without His tag. Lane 3: Wash step, indicating a slight loss of protein. Lane 4: The elution step leads to the

release of the bound HRV-3CP protease (His tagged) and cleaved His tag of M^{pro}. Lane 5: 1.5 µg M^{pro}-L57F without His Tag after desalting. (**B**) Western blot analysis of the cleavage of the "long" cutting reporter after the addition of the different M^{pro} variants. The RLuc-F[2] band at approximately 26 kDa indicates successful reporter cleavage. One representative western blot of n = 3 independent experiments is shown. (**C**) Schematic representation of the fluorescence cleavage-based M^{pro} activity assay. (**D**) Fold-changes of IC_{50} values of nirmatrelvir (blue) and ensitrelvir (light grey) against WT, L57F, L167F/P168S and L167F/P168S/L57F proteases. (**E**) Dose response experiment with nirmatrelvir (left) and ensitrelvir (right) with WT, L57F, L167F/P168S and L167F/P168S/L57F (± SEM; n = 2). (**F**) WT and mutant proteases relative activity at 4 hours (± SD; n = 4). The WT signal in the absence of inhibitor was normalized to 1 and mutants' relative activities were normalized accordingly using WT as reference scale.

Purification and resistance characterization of M^{pro} variants through dose response experiments via a fluorescent-based biochemical assay

To further support our findings, recombinant proteases of a few selected mutants were produced as described in **Fig. S10A** (WT, L57F, L167F/P168S and L167F/P168S/L57F) for a subsequent fluorescent-based protease activity inhibition assay. The activity of the purified proteases was tested by western blot analyses using the recombinant proteases and a cellular expressed cutting reporter bearing the nsp4/nsp5 autocleavage sequence in the SARS-CoV-2 polyprotein. The M^{pro} variants were tested also with a fluorescence cleavage-based (**Fig. 4C-E**) assay and the resistance phenotype of the L167F/P168S/L57F variant observed in our cell-based assays was confirmed. To display dose-response fitting curves from the fluorescence cleavage-based assay, we normalized to the highest average mean of each dataset independently, thereby obtaining curves with the same plot scale. As shown in **Fig. 4F**, mutant proteases had progressively decreased cleavage activity.

Computational analyses reveal destabilization of M^{pro}-nirmatrelvir complex formation and provide insights on protease stability and dimerization affinity upon mutation

To provide a molecular interpretation of the mechanism of resistant M^{pro} variants against nirmatrelvir compared to the WT, we performed molecular modelling with Bioluminate (*44–47*). This software uses already existing structural data to calculate the impact of a mutation on the stability of a specific protein conformation. The software returned models of mutant structures and delta stability values (Δ_Stability in kcal/mol), indicating whether a mutation stabilizes (negative values) or destabilizes (positive values) the investigated conformation. Many of the investigated mutations returned positive values, indicating destabilization of the nirmatrelvir-binding conformation in the WT and Omicron-M^{pro} structures, (PDB entries 8DZ2 (*48*) and 7TLL (*49*), plus related variants (**Fig. S13**)). Most notably, Δ_Stability values of +36.8, +36.5, +46.9 and +14.2 kcal/mol were calculated for the M^{pro} L167F/P168S/L57F, L167F/F305L/P184S/T21I, Omicron/A206T and Omicron/A206T/K100N mutations respectively, which were among the mutations that conferred the strongest resistance phenotypes (**Fig. 3A**, **Fig. 3B**).

Some mutated residue side chains are within the ligand binding site, therefore close to the inhibitor, but do not directly interact with it, namely L167, P168, Y54, L141 and several others (**Fig. 5A**). Furthermore, the majority of selected and observed resistance mutations were located outside of the drug-binding site. We therefore assumed that both near-catalytic site and distant mutations exert an allosteric effect. I.e., they change protein dynamics in favor of conformations that are less compatible with ligand binding, especially when in combination with other mutations.

For example, Omicron/A206T that conferred strong resistance to drug treatment, was a frequently occurring mutation in our experiments and is distant from the catalytic site, prompting us to investigate the structure in more detail. Parental Omicron has a hydrophobic side chain alanine residue buried inside the protein at position 206, which is surrounded by the hydrophobic side chains of L253, P293, and V296 (**Fig. 5A**). The larger, polar side chain of the threonine mutant would be incompatible with the hydrophobic local environment, sterically clashing with the residues nearby, thereby requiring conformational changes to accommodate the side chain of threonine.

We also looked at L167F/F305L/F8L in more detail since it has both an N-terminal and C-terminal cleavage site mutation as well as the catalytic site L167F. In the different M^{pro} assays, mutations F305L and F8L alone did not confer resistance. In addition, the F305L mutant showed a gain-of-function phenotype (**Fig. 3A**) (*34*). Consistent with these findings, we observed a considerable

decrease in stability for L167F (+61.9 kcal/mol), a similar stability for F8L (+1.09 kcal/mol) and an increase in protein stability for F305L (-5.7 kcal/mol), compared to the WT.

To investigate the effect of F305L in more detail, dimerization affinity prediction calculations were performed using Osprey, version 3.3 (*50*). This program uses rotamers of residue side chains to provide an estimation of binding affinity between two interaction partners, with higher Log_{10} K* scores indicating stronger and lower Log_{10} K* scores weaker binding. An increase in the binding affinity of +9.3 in the F305L mutant was calculated compared to the WT, suggesting an increase of M^{pro} dimerization affinity and formation of the mature dimer. F8L, conversely, did not influence dimerization affinity (**Fig. 5C**). F8 and F305 interact with each other via a pi-pi T-stack (**Fig. 5D and S11D**). In addition, F305 might be affected by electronic repulsion between the aromatic ring of F305 with the backbone oxygen atoms of I152 on protomer A and S123 on protomer B. Upon mutation to L305, the shorter leucine side chain is located further away from the backbone, potentially reducing this electronic repulsion between both protomers A and B (**Fig. 5D**).

Moreover, the contemporary presence of the mutations F8L and F305L likely causes an increase in C-terminus flexibility that could compete with the ligand for the binding site by interfering with its entry or acting as a "wiper" by occasionally leaping towards the active site (**Fig. 5E-G**).

The predicted loss of stability was less severe for the M^{pro} L167F/F305L/F8L (+2.2 kcal/mol) mutant. The Omicron A206T/F8L (+6.1 kcal/mol) also experienced a mild destabilization. For the mutant O/A206T/T198I we predicted an increased stability (-18.1 kcal/mol), indicating that the effects of this mutation might be mediated through a different mode of action.

Besides their impact on drug sensitivity, many of the mutations that decreased protease stability also increased the TM_{50} (**Fig. 3C**), potentially indicating impaired protease activity. For this reason, the stability of all the generated mutants was calculated, in addition to affinity changes of residues within the dimerization interface, for the apo dimer (without nirmatrelvir) (**Fig. S13A,C**).

Overall, we predicted a decrease of protein stability for most mutations similar to the nirmatrelvir results. This is not surprising, given that nirmatrelvir binds to the active M^{pro} conformation and the apo and nirmatrelvir-bound structures are therefore very similar (RMSD of 0.89 Å). The dimerization affinity was less affected than protease stability, with most mutants having a predicted Log_{10} K* score similar to the WT (**Fig. S13B**). Notable exceptions include C128Y, G2D, and G124D which were predicted to disrupt the dimer interaction instead (**Fig. S13B**). This finding was consistent with in vitro experiments, where we failed to plaque-purify VSV-C128Y-M^{pro}.

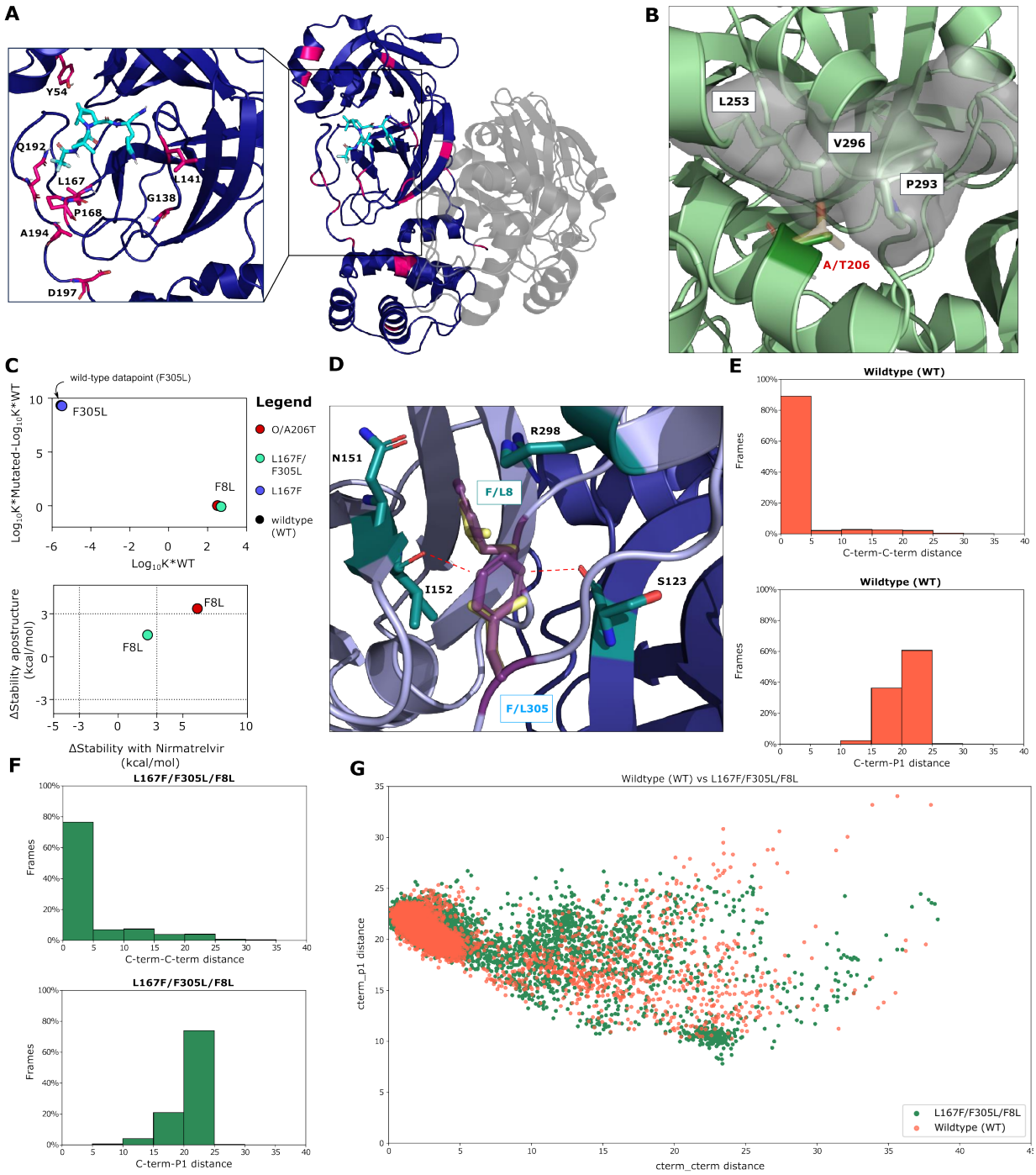

Fig. 5. Structural analysis of mutants. (**A**) Overview of investigated mutants (pink) mapped onto the M^{pro} homodimer (dark violet and grey) bound to nirmatrelvir (light blue sticks, PDB 8DZ2) (*48*). Side chains of residues within the drug binding site are shown as sticks to highlight their distance to nirmatrelvir. (**B**) The WT alanine side chain of residue 206 (dark green sticks) is tightly packed against the hydrophobic side chains of residues L253, P293, and V296 (green sticks, PDB entry 7ALI) (*51*). The threonine side chain (yellow sticks) is polar and would clash with these residues as indicated by the light grey surface. (**C**) Dimerization affinity plot of F8L and F305L mutants (top) and Δ_Stability values (bottom) of apo and nirmatrelvir bound structures for the F8L

mutants within different variant backgrounds. The additional dotted lines at x = ± 3 and y = ± 3 in the stability plot are cut-off values that we considered as a meaningful difference in protease stability (a positive value indicates a decrease in protease stability and a negative value indicates an increase in protein stability). (**D**) The F8 and F305 aromatic side chains (dark violet) form a pi-pi T-stack and F8 in addition interacts with N151, I152, and R298 (green sticks, PDB entry 7ALI (*51*)). F8 is located within the homodimer interface, with chains A and B colored light violet and blue, respectively. The conjugated p-orbitals of the F305 side chain are located nearby of the I152 and S123 backbone oxygen atoms, potentially leading to electronic repulsion. In the F305L mutant, the distance of the leucine side chain (yellow sticks) to these oxygen atoms is considerably increased. (**E**) Bar plots showing the percentage of frames against C-term/C-term distance or C-term/P1 distance for WT (**F**) Bar plots showing the percentage of frames against C-term/C-term distance or C-term/P1 distance for L167F/F305L/F8L mutant. (**G**) Scatter/cluster plot of WT (orange) and the L167F/F305L/F8L (dark green) mutant C-term-C-term and C-term/P1 distances.

To analyze specific multiple-catalytic site mutants in more detail, we performed Thermal Titration Molecular Dynamics (TTMD) (*52*) simulations. This approach allows the characterization in high detail of the binding mode and residence time of a ligand within the catalytic site of a given protein-ligand complex. It thereby enables the comparison between: 1. different ligands bound on the same protein (*52*); 2. different ligand conformations within the same binding site; 3. different mutants of the same protein in complex with the same ligand, as in the present work. The simulation consists of a series of classic molecular dynamics simulations performed at progressively increasing temperatures. Whether the ligand remains bound or not, is monitored through a scoring function based on protein-ligand interaction fingerprints (IFP_{CS}) (*53*). In the end, the simulation returns the so-called MS coefficient, which is a measure for how long the protein-ligand complex is stably formed. This measure facilitates a qualitative comparison of the different protein-ligand complex stabilities, for example, nirmatrelvir bound to different mutant M^{pro} variants. A schematic representation of TTMD is shown in **Fig. 6A**.

Five catalytic site mutants (**Fig. 3A**) that arose from the WT-L167F-M^{pro} were selected for further investigation and compared to WT M^{pro}-nirmatrelvir. Their structures, plus so-called rainbow blots, root-mean-square-deviations and interaction energy blots are depicted for the most resistant variants L167F/P168S/L57F and L167F/F305L/P184S/T21I (**Fig. 6B,C**).

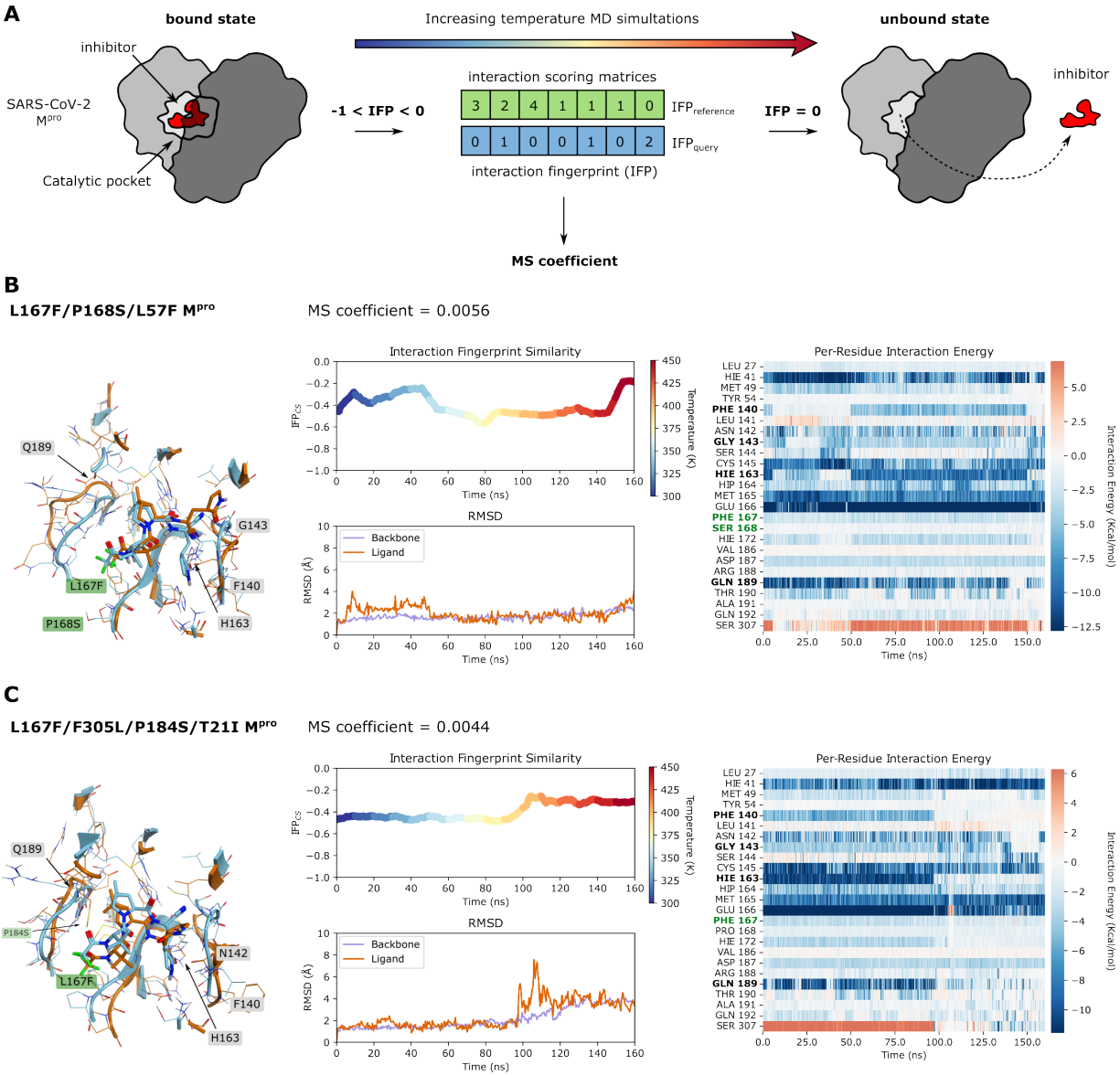

Fig. 6. Thermal Titration Molecular Dynamics (TTMD) experiments of L167F/P168S/L57F and L167F/F305L/P184S/T21I. (**A**) Schematic representation of the TTMD simulation workflow. Interaction FingerPrints (IFP) of reference (bound state) and query (protease-inhibitor complex at each temperature condition over time) are continuously sampled and compared during the simulation. The closer from negative to zero the IFP_{query} is, the more the protease-inhibitor complex differs from the bound state. (**B**) TTMD simulation data of the nirmatrelvir-L167F/P168S/L57F-M^{pro} complex. Left: overlay of M^{pro} structure at the beginning (turquoise) and at the end of the simulation (orange); middle top: a rainbow plot including the IFP_{CS} (left y-axis) is plotted against time in nanoseconds (x-axis). Additionally, the temperature in Kelvin is indicated by colours from blue to red on the right y-axis; middle bottom: the root-mean-square-deviation (left y-axis) is plotted against time in nanoseconds (x-axis); right: a heat map of interaction energies between ligand and surrounding residues. (**C**) TTMD simulation data of the nirmatrelvir-L167F/F305L/P184S/T21I-M^{pro} complex. Left: overlay of M^{pro} structure at the beginning (turquoise) and at the end of the simulation (orange); middle top: a rainbow plot including the IFP_{CS} (left y-axis) is plotted against time in nanoseconds (x-axis). Additionally, the temperature in Kelvin

is indicated by colours from blue to red on the right y-axis; middle bottom: the root-mean-square-deviation (left y-axis) is plotted against time in nanoseconds (x-axis); right: a heat map of interaction energies between ligand and surrounding residues. Residues that are mentioned in the results/discussion are highlighted in black, and mutated residues are highlighted in dark green.

Each TTMD simulation was repeated five times and the MS coefficients for each run are shown in **Table 1**. A plausible cutoff for the MS coefficient to discriminate between stable and unstable protein-ligand complexes is 0.004, according to recent literature (*52*, *54*). We observed that all the mutant MS values were above 0.004, meaning that the stability of the non-covalent nirmatrelvir-M^{pro} complex was lower for all of them compared to the WT protease. To illustrate the stabilities of the different complexes, we ranked the mutants according to MS coefficients, showing that the nirmatrelvir-WT complex is the most stable, followed by L167F/F305L/F8L > L167F/F305L/S144A > L167F/F305L/P184S/T21I > L167F/F305L/P184S and > L167F/P168S/L57F.

As can be observed in **Fig. 6B**, **Fig. S11B,C** and in **Movies S1-S3**, the active site constantly reshapes throughout the simulation, particularly in the area affected by the catalytic site mutations. L167F/P168S/L57F leads to a loss of interaction with the backbone of crucial catalytic site residues (164-166) by altering the floor of the binding site, causing a drag effect on the ligand (**Fig. 6B** and **Movie S1**). The resistance mechanism of the L167F/F305L/S144A mutant seemed to be increased instability of the oxyanion loop (residues 138 to 145), which finally leads to the loss of the hydrogen bond with H163 subsequently resulting in the loss of interaction with residue Q189 (*55*) (**Fig. S11B** and **Movie S2**). In L167F/F305L/P184S, the P184S mutation enhances the plasticity of the loop spanning from residue S184 to Q192, thereby reducing the ligand-pocket shape complementarity, which is essential for ligand positioning. At the same time, it did not alter the S1 pocket and the hydrogen bonds network within the central portion of the cavity (**Fig. S11C** and **Movie S3**). In mutant L167F/F305L/P184S/T21I the TTMD indicated a stabilizing effect of the T21I mutation (**Fig. 6C**, **Movie S5**). L167F/F305L/F8L (**Fig. S11D**) has only one mutation within the catalytic site (L167F) and two other ones located at the dimeric interface, F8L and F305L. As mentioned above, the C-terminus is freer to move in the available space, potentially competing with the ligand (**Movie S6**).

Further modeling results on the potential mechanisms leading to the resistance phenotype can be found in **Fig. 6B,C**; **Fig. S11A-D**, **Fig. S12,** and in **Movies S1-S6**.

	Stability rank (decreasing)	Average	MD1	MD2	MD3	MD4	MD5
Wildtype (WT)	1	0.0037	0.0029	0.0028	0.0048	0.0037*	0.0047
L167F/P168S/L57F	6	0.0056	0.0055*	0.0069	0.0111	0.0038	0.0045
L167F/F305L/S144A	3	0.0044	0.0029	0.0044*	0.0052	0.0042	0.0047
L167F/F305L/P184S	5	0.0051	0.0049	0.0071	0.0050*	0.0054	0.0026
L167F/F305L/F8L	2	0.0042	0.0046	0.0041*	0.0038	0.0039	0.0083
L167F/F305L/P184S/T21I	4	0.0045	0.0040	0.0044*	0.0052	0.0049	0.0038

Table 1. MS coefficients of WT and mutants. MS coefficients were determined for each TTMD replicate, as well as the average MS value was calculated across three different replicates after discarding the highest and lowest values. The representative replicate (i.e., the closest to the average MS value), is highlighted with a star (*).

Discussion

In this study, we used a previously developed, safe, VSV-based tool to select for protease inhibitor resistance mutations. We demonstrated the capability of our selection tool to achieve SARS-CoV-2 M^{pro} variants with multiple substitutions and potential Omicron-related mutants relevant to the current pandemic situation. The Omicron-M^{pro} sequence has a signature substitution at amino acid position 132, exchanging proline for histidine (P132H). This mutation has been shown to not confer resistance to approved protease inhibitors, such as nirmatrelvir and ensitrelvir (*29–32*), in agreement with our findings, but altered thermal stability (*33*). To reduce biosafety risks, previous studies based on authentic SARS-CoV-2 to select resistance mutations have used early variants of SARS-CoV-2 (*17–19*). Protease inhibitor resistant mutants derived from such early variants would unlikely be able to compete with Omicron variants if released accidentally. This caveat does not apply to mutation selection performed with our VSV-M^{pro} system. Consequently, we selected protease inhibitor resistant Omicron M^{pro} variants.

First, we generated a large number of mutants, which we subsequently filtered on a three-parameter basis for further characterization: proximity to the inhibitor, prevalence in the GISAID database, and frequency of a specific mutation in our selection experiments. Substitutions that appeared close to the inhibitor are likely to yield a meaningful, structural explanations that can then inform structural derivatization of inhibitors. Highly frequent substitutions in GISAID are relevant because they occur in viruses that can efficiently propagate in humans and therefore should not be detrimental to the enzymatic activity. Lastly, mutations that occur frequently in selection experiments could be favoured routes of evolution for the M^{pro} enzyme (*56*) and thus the mechanism might be of interest.

Wuhan-1 M^{pro} resistance mutations. In the VSV-L167F-M^{pro} variant, additional substitutions close to the inhibitor were selected, namely S144A, P184S and L57F. In some of those mutants, we observed differences in the susceptibility to the two protease inhibitors nirmatrelvir and ensitrelvir. Interestingly, the mutant L167F/F305L/F8L demonstrated a considerably greater degree of resistance to ensitrelvir than to nirmatrelvir. Both L167F/F305L/P184S and L167F/F305L/L57F also showed a similar trend. L167F/P168S/S144A had a comparable resistance profile against both protease inhibitors. Mutations of S144 can simultaneously cause an increase in drug resistance and a decrease in the catalytic activity of the protease. Consistently, in our kinetic experiments and in the literature (*17*) this mutation impaired replication.

As mentioned above, the second filtering parameter we applied was the absolute frequency of a specific nsp5/M^{pro} substitution in the GISAID database. Residues in catalytic sites tend to be more conserved (*40*, *41*), and we expected them to not be frequently represented. Indeed, most high-frequent mutants were observed in positions further away from the active pocket. P184S is an exception as it is frequently represented (4713 entries, as of 01/06/2023) and is near the catalytic pocket. During the rescue (generation of a virus from a plasmid) of L167F/F305L/P184S, one sample had acquired an additional mutation, namely T21I. The resistance profile of L167F/F305L/P184S changed from more resistant against ensitrelvir to stronger resistance against nirmatrelvir upon introduction of the T21I mutation. However, T21I alone neither conferred resistance against nirmatrelvir nor ensitrelvir in both our gain- and loss-of-signal assays. Consistently, it has been reported in the literature as a compensatory mutation for the fitness loss caused by other resistance mutations, such as E166V (*17*). T21I could support or compensate the increased plasticity of L167F/F305L/P184S we found in TTMD simulations by an allosteric, re-stabilizing effect (*26*, *57*).

Taken together, it seems that L167F, P168S, S144A and P184S mutations deform the binding pocket and loosen the electrostatic interactions between nirmatrelvir and the protease. This observation indicates that not only mutations of residues directly interacting with the inhibitor can decrease susceptibility, but mutations of residues within the catalytic pocket can trigger conformational changes in the enzymatic cleft that impact in the inhibitor-protease complex formation as well.

The triple mutant L167F/F305L/F8L is among our most resistant M^{pro} variants and showed the most marked difference in susceptibility between the two inhibitors nirmatrelvir and ensitrelvir. The residues F305 and F8 are in close contact with each other in the folded protein and this is expected to be a stabilizing interaction for the dimer interface. However, other residues are viable at this position, for example leucine (L) has been shown to be the preferred residue instead of phenylalanine (F) (*40*). In fact, we calculated an increase in the dimerization affinity from F305 to L305, suggesting an improved formation of the fully functional dimeric protein upon mutation. We hypothesize that the simultaneous mutation of both F305 and F8 residues into leucines results in a modified interaction between the C-term of the first protomer and the N-term of the second, increasing the conformational degree of freedom of the C-term, thereby leading to a “wiper” mechanism that could compete with the inhibitor for the catalytic site (**Movie S6**).

Omicron-M^{pro} resistance mutations. In the Omicron-M^{pro}, we selected resistance mutations within and outside the catalytic site. The catalytic site mutant O/Y54H, which occurred 10 times independently, and O/R188W showed only mild resistance alone as well as in combination with R222L. O/Y54H/F305L showed an increased, selective resistance against ensitrelvir, with a 20-fold-change of the IC_{50} in the loss-of-signal assay (M^{pro}-Off). The mutants K100N, T198I, A210S, and A266T were frequent in GISAID, therefore we picked them for resistance characterization. Double mutants O/A206T/F8L and O/A206T/T198I showed strong selective resistance against nirmatrelvir (up to 100-fold-change), while being less effective for ensitrelvir. We speculate that in O/A206T/F8L, the contribution of F8L to resistance is the same as in mutant L167F/F305L/F8L ("wiper" effect). O/A206T and O/A206T/K100N also showed considerable resistance (~26-fold), and the computed destabilization correlated with resistance.

Most Omicron and WT substitutions, however, occurred in the amino acid sequence that constitutes domain III (residues 200-306) of M^{pro}. We chose to characterize the A206T mutation due to the unusually high frequency in the selected VSV-O-M^{pro} pool of mutants. A206T only fits our third filtering parameter: frequent selection in our experiments, as it is far from the catalytic site and scarcely represented in the GISAID database. Likewise, some mutations may confer an advantage for viral replication only in the presence of a protease inhibitor, but would probably be unfavourable in an untreated individual. In a study on SARS-CoV M^{pro} polyprotein maturation (*58*), the authors proposed that despite the presence of deleterious mutations that hinder mature dimerization, M^{pro} retains its ability to cleave its N-terminal cleavage site owing to a weak, immature dimerization between two polyprotein monomers (1 and 2) catalysed by domain III (1) – domain III (2) interaction. We speculate that mutations of residues between 200 and 306, most frequently found in VSV-O-M^{pro}, positively contribute to this interaction. As with A206T, we hypothesize that destabilization of the mature M^{pro} to which nirmatrelvir binds, is the resistance mechanism.

Viral replication kinetics. In kinetic studies, we observed that the gain in signal was slower for some mutants, which might be an indirect read-out of M^{pro} activity. The L167F/F305L/S144A mutant was among the slowest variants and the time to reach the plateau was 1.75 times longer than for the WT M^{pro}. However, mutations that were described as compensatory in the literature, such as T21I (*17*), could partially reverse the kinetic attenuation. T21I appeared after generating VSV-L167F-F305L-P184S-M^{pro} from its plasmid (also called "rescue"), and R222L after plaque

purification of VSV-O-R188W-M^{pro}. T21I and R222L are likely compensating for substitutions that occur in the catalytic site. For example, T21I seems to compensate for the loss of replicative fitness of the previously described, highly resistant M^{pro}-E166V variant (*17*). Furthermore, T21I is a very frequently represented substitution in the GISAID database (21248 entries, 1st of June 2023). From mutations generated in the VSV-O-M^{pro} variant, viral replication kinetics exhibited a different behaviour in comparison to VSV-L167F-M^{pro} variants. Overall, mutations related to the Omicron-M^{pro} did not decrease the viral replication rate to the same extent as in the L167F-M^{pro}. The slowest variant was O/A206T/F8L, with a 1.4 fold-change (increase of ~16 hours) to reach the plateau. As expected, most mutations that have been computationally predicted to decrease protease stability indeed had a negative impact on viral replication, with higher TM_{50} values related to M^{pro} variants bearing such mutations. Furthermore, they were less represented in isolates in the GISAID repository (1 or 2 digit entries), whereas those that had similar TM_{50} values compared to the WT such i.e. F8L, P184S or K100N and T198I alone, were more frequent (3 or 4 digit entries). These correlations support the validity of computational predictions.

Caveats of the study. Chimeric VSV-Spike, where the VSV glycoprotein G was replaced by the SARS-CoV-2 spike, has been used previously to predict antibody escape mutations (*59–61*). Before that, VSV was shown many times to be promiscuous in the context of its glycoprotein, which could easily be replaced for example with those of Ebola and Marburg viruses (EBOV, MARV) (*62*) or lymphocytic choriomeningitis virus (LCMV) (*63*). To replace one of the VSV intergenic regions with a protease and thereby use VSV as a protease mutation tool, however, was a previously untapped area of research. Therefore, our resulting mutations need careful evaluation. We observed many mutations in chimeric VSV-M^{pro}, some of which aligned exactly with those identified in SARS-CoV-2 gain-of-function experiments reported in the literature (*17*, *19*), whereas others were completely different. One explanation may be the artificial system used in our experiments, which uses cis-cleavage, the only requirement for M^{pro} processing in VSV-M^{pro} replication. In contrast, in experiments with authentic SARS-CoV-2, both cis- and trans-cleavage must be preserved in the development of resistance mutations. As described above, during auto-cleavage there is an intermediate dimerization state, different from the mature protease. This intermediate state may require specific interactions that differ from those involved in the mature dimer. Another reason is the difference between the natural polyprotein of coronaviruses and the artificial G-M^{pro}-L polyprotein expressed by chimeric VSV-M^{pro}. The observed divergence

between SARS-CoV-2 and VSV may also be attributed to the differences between two viral polymerases: SARS-CoV-2 RNA-dependent RNA polymerase (RdRp) and VSV polymerase L. It has been reported in the literature that the SARS-CoV-2 polymerase can proofread (*64*), whereas VSV lacks such mechanisms (*36*), leading to a much higher error rate of 1/10,000 nucleotides, as described previously (*35*, *65*). However, this could also be seen as an advantage as resistance mutations may develop faster in the error prone VSV-M^{pro} replication system.

Overall, we conclude that most PI resistance mutations destabilize the mature conformation of M^{pro} and consequently the inhibitor-M^{pro} complex formation, which finally impairs inhibitor binding. These resistance mutations can be selective either for nirmatrelvir or for ensitrelvir, or impact both. Moreover, the extended use of protease inhibitors will increase the risk of selecting SARS-CoV-2 protease-inhibitor resistant variants. Therefore, to combat such variants, there is still a need for new PIs that ideally target alternative conformations or regions of the protease (*55*, *66*) and for rationale design of new inhibitors less affected by mutations (*67*).

Materials and Methods

Study Design

The overall rationale of the study was to use a previously developed mutation selection tool based on VSV to select and characterize a comprehensive collection of SARS-CoV-2 main protease inhibitor escape mutants. The study was performed on cell lines, in biochemical settings and in-silico, and no animal husbandry or human participants were involved. Human cell lines with replicating BSL-1 and -2 viruses were treated with protease inhibitors to observe resistance phenotypes in appropriate facilities. Viral titres were determined using $TCID_{50}$. Measurement readouts were fluorescence-based, detected by ELISpot/FluoroSpot and multi-well readers. Autofluorescent fibers were excluded automatically and manually from spot counting in the ELISpot readout. Experiments were neither blinded nor randomly distributed to experimenters. We chose sample sizes empirically based on experience from former studies. At least two and up to four biologically independent replicates were performed per condition. Biologically independent meant distinct wells with the same condition, not multiple measurements of the same wells (technical replicates). Resistance phenotypes were reproduced at least twice, usually more often and in different combinations. Representative measurements were chosen to compile graphs and figures.

Cloning strategies

The chimeric VSV variant with M^{pro} instead of the intergenic region between G and L was cloned as previously described (*34*). VSV-G-M^{pro}-L carrying the Omicron signature mutation, P132H (VSV-O-M^{pro}) was cloned as follows: N-terminal fragment comprising part of G and M^{pro} was amplified by PCR using primers 33n-before-KpnI-for (Q195) and Omicron rev VSV-G-M^{pro}-L as a template, and the C-terminal fragment comprising the remaining part of M^{pro} and part of L was amplified by PCR using primers Omicron for and 33n-after-HpaI-rev (Q176). N- and C-terminal fragments were ligated together by Gibson assembly in a VSV-backbone vector digested with KpnI and HpaI. VSV-G-M^{pro}-L L167F (VSV-L167F-M^{pro}) was plaque purified as it was generated in a previous work.

M^{pro}-Off and -On point mutants were generated by point directed mutagenesis on parental plasmids (GenBank accession codes: M^{pro}-Off: ON262565; $M^{pro\ pro}$-On: ON262564) with mutation primers and the Herculase II Fusion DNA Polymerase. Herculase is a polymerase that can overcome

difficult (GC-rich) sequences and amplify large plasmids. However, primers could not be chosen for optimal design because the mutation site was fixed. For this reason, this simple point directed mutagenesis did not work for each construct. For M^{pro}-On and M^{pro}-Off plasmids, where point-directed mutagenesis did not work, we used mutagenic Gibson assembly as previously described (*34*). Cloning primers used in this study are shown in **Table S1**.

Cell lines

BHK-21 cells (American Type Culture Collection, ATCC) were cultured in Glasgow Minimum Essential Medium (GMEM) (Lonza) supplemented with 10 % fetal calf serum (FCS), 5 % tryptose phosphate broth, and 100 units/ml penicillin and 0.1 mg/ml streptomycin (P/S) (Gibco). 293T cells (293tsA1609neo, ATCC), and 293-VSV (293 expressing N, P-GFP and L of VSV) (*68*) were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10 % FCS, P/S, 2 % glutamine, 1x sodium pyruvate and 1x non-essential amino acids (Gibco). Vero E6 (ATCC CRL-1586) was cultured in DMEM supplemented with 5 % FCS (VWR) and 1 % penicillin−streptomycin−glutamine (PSG) solution (Corning).

Virus recovery ('rescue')

VSV-G-M^{pro}-L WT and a few mutants were rescued in 293T cells by $CaPO_4$ transfection of whole-genome VSV plasmids together with T7-polymerase, N-, P-, M-, G- and L expression plasmids as helper plasmids (*42*). Briefly, genome and helper plasmids were transfected into 293T in the presence of 10 μM chloroquine to avoid lysosomal DNA degradation. After 6 to 16 hours, chloroquine was removed, and cells were cultured until cytopathic effects occurred. M and G proteins were used as helper plasmids; although these proteins are optional in the recovery of VSV, they were chosen here as a precaution to support the rescue of a potentially attenuated virus variant. After the rescue, viruses were passaged on 293-VSV cells and plaque purified twice on BHK-21 cells. ΔP and ΔL VSV variants expressing dsRed were produced on replication supporting 293-VSV cells. VSV-G-M^{pro}-L was fully replication competent and produced on BHK-21 cells.

Mutation selection assay

10^4 BHK-21 cells/well were seeded in a 96-well plate one day before infection with the chimeric VSV-M^{pro} viruses. VSV-L167F-M^{pro} and VSV-O-M^{pro} infected cells (MOI = 0.01) were treated

with nirmatrelvir with concentration ranging from 10 to 100 µM. The resistant variants VSV-L167F/P168S-M^{pro} and VSV-L167F/F305L-M^{pro} at MOI = 0.01 were passaged 5 times, and with increasing concentrations of inhibitor (from 50 to 100 µM). VSV-O-M^{pro} and VSV-O/A206T-M^{pro} at MOI ranging from 0.001 and 0.01 were passaged only once with increasing concentrations of inhibitor (from 10 to 100 µM). Each virus variant occupied from 24 wells to 96 of the 96-well plate. Supernatants from wells that displayed cytopathic effect after 48-72 hours were collected for downstream viral RNA isolation, cDNA synthesis, PCR amplification and Sanger or Nanopore sequencing as described in this section.

Viral RNA isolation, cDNA synthesis and M^{pro} sequencing

VSV-G-M^{pro}-L RNA was isolated either by using the E.Z.N.A. Viral RNA Kit (Omega Bio-Tek Inc.), the NucleoSpin RNA Virus (Macherey-Nagel GmbH) or by semiautomated RNA extraction using the Easymag (BioMerieux^{TM}). BHK-21 cells were infected with the respective VSV-G-M^{pro}-L variant in 96-well plates as described above. Virus-containing supernatants were collected from individual 96-wells and the RNA was purified from the supernatants according to manufacturers' instructions (Easymag). Then, cDNA was synthesized from isolated viral RNA by RevertAid RT Reverse Transcription Kit (Thermo Fisher Scientific). G_{Cterm}-M^{pro}-L_{Nterm} fragment sequence was amplified by PCR with primers (primer_for: CTCAGGTGTTCGAACATCCTCAC and primer_rev: GATGTTGGGATGGGATTGGC) and either sent for Sanger sequencing (MicroSynth AG) or sequenced using Nanopore (as described below). Obtained sequences were mapped to the L167F-M^{pro} or Omicron-M^{pro} reference sequences in Geneious Prime 2023.0.1 build 2022-11-28 and examined for mutations.

Nanopore sequencing workflow

After the first two selection experiments where we sequenced our samples with Sanger sequencing (MicroSynth AG), we used Nanopore sequencing. First, PCR amplification products of the G_{Cterm}-M^{pro}-L_{Nterm} fragments, as described above, were used. The sequencing libraries were prepared using the Rapid Barcoding Kit SQK-RBK110.96 (Oxford Nanopore Technologies, ONT), and up to 96 samples were multiplexed and sequenced on a MinION Mk1B sequencer with R9.4.1 flowcells (ONT). Raw data in the form of electrical signals are translated into nucleotide sequences (base-calling) and saved in fastq files. Basecalling using the super high accuracy model,

demultiplexing, as well as barcode and adapter sequence trimming, was performed in Guppy (version 6.1.5, ONT).

Raw reads were filtered to a PHRED quality score of ≥ Q15 and length between 200 bp and 1800 bp using SeqKit (version 2.4.0) (*69*), aligned to the reference sequence using minimap2 (version 2.22) (*70*), followed by sorting and indexing using SAMtools (version 1.13) (*71*). SAMtools depth was used to check sufficient depth. To find SNVs, we used LoFreq for variant calling with --min-cov option set to 300 (v2.1.3.1) (*72*).

The resulting VCF files were imported to Geneious Prime 2023.0.1 and called variants manually checked for plausibility.

Unrooted phylogenetic tree generation and visualization

Mutations had to be manually inserted into the main protease sequence for each variant. The phylogenetic tree was generated using the online tool http://www.phylogeny.fr/ with standard parameters. The output as .newick file was imported to Geneious Prime 2023.0.1, visualized and exported as PDF. Graphic rearrangements and overhauls were made using Inkscape version 1.1.

Collection of GISAID frequencies of generated mutations

To obtain updated nsp5 frequencies, metadata from the EpiCoV database (correct as of 1st June 2023) was used; specifically, the `all_mutations` field from a JSON dump obtained with GISAID credentials. Using Python 3.9 and the pyarrow and pandas' libraries, sequences were classified into WT (not assigned Omicron lineage, and not carrying the P132H mutation) and Omicron (either assigned Omicron lineage, or carrying the P132H mutation). Then, for both these groups of sequences, the counts of 95 nsp5 mutations of interest were aggregated.

SARS-CoV-2 variant phylogenetic analyses using Ultrafast Sample placement on Existing tRees (UShER)

Phylogenetic trees were generated for the resistance variants of interest using patient derived sequences deposited in the GISAID database. For each variant of interest, viral genomes harboring the corresponding mutation in Nsp5 were retrieved using the GISAID EpiCoV web server after filtering to consider only viruses descending from the Omicron lineage and removing genomes with low sequence coverage. These sequences were subsequently uploaded to the UShER (*73*) web server to generate phylogenetic trees using all the sequences available in the GISAID database,

visualized using the Auspice.us web application and edited to highlight the variants of interest using Adobe Illustrator. The frequency of each resistance mutation within Omicron and Pre-Omicron lineages was similarly determined using the GISAID EpiCoV web server by dividing the number of occurrences of the corresponding Nsp5 mutation within Omicron or all other SARS-CoV-2 lineages by the total number of viral genomes deposited belonging to each respective lineage. Histograms comparing the frequency of the variants within the Omicron and Pre-Omicron lineages were generated using GraphPad Prism 9.

Screening assay with ELISpot/FluoroSpot read-out

3×10^5 cells were seeded per well in 6-well plates and transfected one day after seeding with $3CL^{pro}$ plasmids using TransIT-^{PRO} (Mirus Bio LLC) and incubated 8-9 hours for M^{pro}-On and 10 hours for M^{pro}-Off assays. Then, cells were seeded into a 96-well plate with 2×10^4 cells per well in 50 µl complete growth medium. Directly after seeding, compounds and virus (MOI 0.1) were added in 50 µl complete growth medium to wells. After 48 to 72 hours, supernatants were removed, and fluorescent spots counted in a Fluoro/ImmunoSpot counter (CTL Europe GmbH). For experiments where FluoroBrite™ medium was used, supernatants were not removed and the signal was read out every 12 hours. The manufacturer-provided software CTL switchboard 2.7.2. was used to scan 90% (70% when FluoroBrite™ medium was used) of each well area concentrically to exclude reflection from the well edges, and counts were normalized to the full area. Automatic fiber exclusion was applied while scanning. The excitation wavelength for dsRed was 570 nm, and the D_F_R triple band filter was used to collect fluorescence. To increase comparability between M^{pro}-On and -Off signals, we normalized dsRed events with the following strategies. In M^{pro}-On, the highest compound concentrations would not reach the same value due to the different responses of each mutant. Therefore, we normalized to the highest mean of the experiment. In M^{pro}-Off, we normalized the signal to each individual's highest mean of the construct. For both assays, fluorescent spot count (y-axis) data were plotted against PI concentration (x-axis) and IC_{50} values were extrapolated as described later in this section.

Replication kinetic experiments

For replication kinetic experiments with WT M^{pro}, Omicron and mutants, we adapted the 3CL/M^{pro}-Off assay to enable multiple read-outs over time every 12 hours. We exchanged the standard DMEM supplemented with 10 % FCS, P/S, 2 % glutamine, 1x sodium pyruvate and 1x non-

essential amino acids (Gibco), with FluoroBrite™ DMEM equally supplemented. FluoroBrite™ DMEM has the same composition as DMEM, except it lacks phenol red which interferes with spot recognition. We experimented as described in the previous paragraph. After seeding 2 x 10^4 transfected cells in 50 µl of FluoroBrite™ complete medium into 96-well plates, only virus (MOI 0.1) was added (50 µl). The signal was read out at 12 hours intervals. As the virus starts replicating, dsRed is expressed by the infected cells, and the expression of dsRed is correlated with the amount of viral progeny and therefore with the number of spot count. We generally stopped signal acquisition after 84 up to 108 hpi, because the cells die due to virus-induced cytopathy, and the spot count decreases. The manufacturer-provided software CTL switchboard 2.7.2. was used as described above. Replication kinetic curves were normalized individually to the highest mean of the construct. We normalized fluorescent spot count data by the highest mean of each dataset. Then, we plotted the fluorescent dsRed signal (spot count) against time (hpi). We then performed non-linear regression analyses using the built-in function "Sigmoidal 4PL, X is concentration" in GraphPad Prism 9.5, where concentration was replaced with "hours" and extrapolated the TM_{50} (time-maximum fifty) value, which represents the time required to achieve half of the maximum signal/plateau.

Protein Expression and purification

M^{pro} with a C-terminal Hexahistdine tag (M^{pro}-6H) was expressed in *E.coli* followed by a tag removal with the human rhinovirus 14 3C protease (HRV3C protease).
WT M^{pro}-6H construct was codon optimized for the expression in E. coli and ordered from Thermo Scientific (Waltham, MA, USA). Point mutations to generate the mutant variants were introduced using the Q5® Site-Directed Mutagenesis Kit purchased from New England Biolabs (NEB, Ipswich, MA, USA). Primers were acquired from Sigma Aldrich (St. Louis, MO, USA).
M^{pro}-6H variants were transformed into electrocompetent *E. coli* BL21(DE3) cells. An overnight culture was inoculated 1:50 in 1L TB medium and cells were grown at 37°C, 220 rpm to an OD_{600nm} of 1. After induction with 0.4 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) temperature was lowered to 25°C for 4 hours. Cells were isolated by centrifugation at 3200 g, 20 minutes at 4°C, and cell pellets frozen at -20°C.

For purification, frozen pellets were resuspended in 50 mM Tris/50mM NaCl, pH 7.5 buffer and disrupted with a French press (Thermo Scientific). Lysed cells were centrifuged at 20000 g, 4°C,

20 min and the clarified supernatant was adjusted to 50 mM Tris/300 mM NaCl/20 mM Imidazole. IMAC purification was carried out with a ÄKTApurifier (GE Healthcare Life Sciences; Little Chalfon, United Kingdom) and a HisTrap FF Crude 1 mL (GE Healthcare) column. Unbound proteins washed out using 4 column volumes (CV) 50 mM Tris/300 mM NaCl/20 mM Imidazole, pH 7.5 and 1 CV 50 mM Tris/300 mM NaCl/40 mM Imidazole, pH 7.5. The protein-containing fractions were pooled, and the buffer was changed to 50 mM Tris/50mM NaCl, pH 7.5 using a HiTrap Desalting 5 mL (GE Healthcare) column. Protein concentration was determined with a NanoDrop ND-1000 Spectrometer (Thermo Scientific).

Hexahistidine Tag removal by negative IMAC

The highly soluble Hexahistidine tagged NT*-HRV3Cprotease was expressed as previously described (*74*) and purified the same way as M^{pro}. M^{pro}-6H was incubated together with NT*-HRV3protease at 4°C overnight followed by negative IMAC purification with a manually packed Ni-NTA column. The flowthrough yields the mature M^{pro} with native N- and C-termini without His tag, because the His tag as well as the NT*-HRV3C protease are bound to the Ni-NTA resin.

Cross validation with biochemical M^{pro} inhibition assay

The biochemical assay used to confirm mutations was based on the $3CL^{pro}$ activity assay from BPS Biosciences, catalog number #78042-2. The M^{pro} in the kit was replaced by an in-house produced M^{pro} and mutants thereof (L57F, L167F7P168S, L167F/P168S/L57F), as described in the M^{pro} purification paragraph. Solutions of WT and mutant M^{pro} variants – 84.5 ng/reaction in 30 µl buffer – (20 mM Tris/HCl pH = 8, 150 mM NaCl, bovine serum albumin 0.1 mg/ml, 1 mM dithiothreitol – DTT) were prepared according to the kit's manual, to reach a final concentration of 50 nM in 50 µl. 10 µl of five-fold excess to tested nirmatrelvir concentrations (180 µM) were added to the 30 µl of M^{pro} solution and incubated for 30 minutes. Then, 10 µl of a fluorescent substrate (Ac-Abu-Tle-Leu-Gln-MCA) were added to get a final concentration of 40 µM. This generates a 1:5 dilution of the excess of protease inhibitors – and therefore final concentrations – and the reactions were incubated for 4 hours. Fluorescence was measured by excitation at 365 nm and read-out at 415-445 nm emission with a Glomax Explorer fluorometer (Promega).

Expression constructs

M^{pro} cutting reporter: Oligonucleotides for the "short" (GCAGTGCTCCAAAGCGGATTTCGC) and the "long" (ATCACGAGTGCAGTGCTCCAAAGCGGATTTCGCAAAATGGCC) cleavage sequence of M^{pro} (nsp4/nsp5 autocleavage site), flanked by N-terminal AgeI and a C-terminal HpaI restriction enzyme sites were ordered at Eurofins. The annealed oligos were ligated into a vector presenting N-terminally the fragment 1 (F[1]) and C-terminally the fragment 2 (F[2]) of the *Renilla* luciferase harboring the corresponding restriction enzyme sites (pcDNA3.1 backbone vector) as previously described (*75*, *76*). Between each fragment and the cleavage sequence we inserted interjacent 10-aa linkers.

Preparation of lysates for western blotting

The cells expressing the M^{pro} cutting reporter and the purified proteins were prepared as described before. 50 µl of the protein suspension (5 µg of protein) were pre-incubated with inhibitor (10µM) or DMSO for 15 minutes at room temperature (R.T.). After that, 25 µl of cell suspension were added and the mixture was incubated for 3 hours at R.T. The cleavage reaction was stopped by the addition of 5x SDS-Loading buffer. Before the SDS-Page the samples were cooked at 95°C for 10 minutes. Western blot images were obtained using the ChemiDoc MP from Bio-Rad.

$TCID_{50}$ assay and dose responses

For initial dose response experiments, 5 x 10^4 BHK-21 cells per well were seeded in 48-well plates one day before infection. Cells were infected in duplicates with a MOI of 0.05 of VSV-M^{pro} WT or VSV-O-M^{pro} and indicated concentrations of nirmatrelvir were added to the wells. After 48 hours, supernatants were collected and titrated with $TCID_{50}$. For quantification, the 50 % tissue culture infective dose ($TCID_{50}$) assay was performed as described previously (*77*). In short, 100 µl of serial dilutions of virus were added in octuplicates to 10^3 BHK-21 cells seeded in a 96-well plate. Six days after infection, the $TCID_{50}$ were read out and titers were calculated according to the Kaerber method (*78*).

IC_{50} calculations

In this study, different assay systems were used to generate resistance data, namely VSV-based cellular assays with FluoroSpot read out, a luminescence-based cellular assay that employs recombinant M^{pro}, as well as a peptide-MCA cleavage-based biochemical assay. IC_{50} calculations and statistical analysis for all assays were performed with GraphPad Prism 10. To calculate IC_{50}s, we used GraphPads pre-set sigmoidal models: 4PL, X is concentration.

$$Y = Bottom + \frac{Top - Bottom}{1 + (\frac{IC_{50}}{X})^{HillSlope}}$$

However, under some circumstances, slightly bell-shaped curves arise due to an initial signal in- or decrease (On and Off, respectively), followed by a signal peak and then a de- or increase. In the On assay, the decrease at the end occurs because of inhibitor toxicity at high concentrations. In the Off assay, the initial increase can be attributed to the fast death of cells in the absence of any inhibitor, whereas cells die slower at small doses, allowing for more cell division and thereby more "substrate" (cells and virus) generating dsRed. The bell-shaped curves can cause overfitting with the sigmoidal models, leading to extremely steep de- or increasing curves. Although in general, this does not lead to important changes in IC_{50}s, we still chose to compensate for such unrealistically steep slopes by constraining the HillSlope to 3 or lower (positive slope) for the gain-of-signal assay and minus 3 or higher (negative slope) for the loss-of signal assay.

Hardware overview (for subsequent TTMD simulations)

Most of the molecular modeling operations, such as the structure preparation, the setup for molecular dynamics (MD) simulations, and trajectory analyses, were performed on a Linux workstation, equipped with a 20 cores Intel Core i9-9820X 3.3 GHz processor, running the Ubuntu 20.04 operating system. Molecular dynamics simulations were carried out on an in-house Linux GPU cluster composed of 20 NVIDIA devices ranging from GTX1080Ti to RTX3090.

Structure preparation

A computational study was conducted to rationalize the effect of SARS-CoV-2 M^{pro} mutations on Nirmatrelvir resistance. In detail, a recently developed enhanced sampling molecular dynamics

method named Thermal Titration Molecular Dynamics (TTMD) (*52*, *54*) was exploited to compare the stability of the non-covalent complex between Nirmatrelvir and both the WT and mutated proteases, based on the assumption that the formation of the reversible, non-covalent complex, is the rate-determining step that leads to the formation of the irreversible (or very slowly reversible), covalent complex (*79*). The top five mutants concerning their experimentally determined drug resistance were investigated.

The experimentally determined three-dimensional coordinates of the WT nirmatrelvir-M^{pro} complex were retrieved from the Protein Data Bank (PDB) (*80*) and prepared for further calculations exploiting various tools provided within the Molecular Operating Environment (MOE) 2022.02 suite (*81*). Particularly, the complex deposited with accession code 7RFW 6 was used for describing the WT protease, while mutants were generated through the "protein builder" module of MOE.

The catalytically competent dimeric conformation of the protease was reconstructed by applying symmetric crystallographic transformations. Inconsistencies in the experimental structure were checked and fixed through the "structure preparation" tool. Missing hydrogen atoms were added according to the predominant tautomeric and protomeric state of each residue at pH = 7.4 through the "Protonate3D" module. Finally, each water molecule was removed before storing the prepared complex structure for subsequent calculations. In the end, a 1:1 protein-ligand non-covalent complex between nirmatrelvir and various M^{pro} proteases was considered in the simulated systems.

System setup for MD simulations and Equilibration Protocol

Each complex obtained in the previous step was prepared for MD simulation making use of various tools provided within Visual Molecular Dynamics 1.9.2 (VMD) (*82*) and the AmberTools 22 (*83*, *84*) suite. All parameters were attributed according to the ff14SB (*85*) force field, except for the ligand ones which were instead assigned through the general Amber force field (GAFF) (*86*). Partial charges for the ligand were calculated through the AM1-BCC method (*87*). Each one of the protein-ligand systems was solvated using the TIP3P (*88*) water model into a rectangular base box, ensuring a minimum 15 Å distance between the box border and the nearest protein or ligand atom. To neutralize the net charge of the system and to reach a salt concentration of 0.154 M, the proper number of sodium and chlorine ions was added. Finally, to remove clashes and unfavorable

contacts, 500 steps of energy minimization through the conjugate gradient algorithm were performed.

Before the production phase, a two-step equilibration process was performed. First, 0.1 ns of simulation in the canonical ensemble were carried out, imposing harmonic positional restraints on both the protein and ligand atoms. Then, 0.5 ns of simulation in the isothermal-isobaric ensemble were performed, with harmonic positional restraints applied only on the protein backbone and ligand atoms. In each case, a 5 kcal mol−1 Å−2 force constant was imposed on each restrained atom for the whole duration of the simulation, and the temperature was maintained at a constant value of 300 K through a Langevin thermostat (*89*). In the second equilibration stage, the pressure was kept constant at 1 atm through a Monte Carlo barostat (*90*).

All MD simulations were conducted exploiting the ACEMD16 3.5 engine, which is based upon the open-source Python library for biomolecular simulations OpenMM (*91*). A 2 fs integration timestep was used, and the M-SHAKE algorithm was used to constrain the length of bonds involving hydrogen atoms. The particle-mesh Ewald method was used to compute electrostatic interactions, using cubic spline interpolation and a 1 Å grid spacing. A 9.0 Å cutoff was applied to the computation of Lennard-Jones interactions.

Thermal Titration Molecular Dynamics (TTMD) simulations

As thoroughly described in the work of Pavan et al. (*52*), Thermal Titration Molecular Dynamics (TTMD) is an enhanced sampling MD protocol originally developed for the qualitative estimation of protein-ligand unbinding kinetics.

The TTMD workflow relies on a series of short classic molecular dynamics simulations, defined as “TTMD-steps”, performed at progressively increasing temperatures. The temperature increase is used to augment the kinetic energy of the system, thus shortening the simulation time required to observe protein-ligand unbinding events compared to classic MD simulations. To monitor the progress of the simulation, the conservation of the native binding mode is evaluated through a protein-ligand interaction fingerprint (PLIF) based scoring function (*53*). The protocol described hereafter is implemented as a Python 3.10 code relying on the NumPy, MDAnalysis (*92*, *93*), Open Drug Discovery Toolkit (ODDT) (*94*) and Scikit-learn packages. The code is released under a

permissive MIT license and available free of charge at github.com/molecularmodelingsection/TTMD.

In detail, the user must define a "temperature ramp", i.e., the number, the temperature, and the length of each "TTMD-step". Consistently with previous works on the target (*52*, *54*), in this case, the starting and end temperatures were set at 300K and 450K respectively, with a temperature increase between each "TTMD-step" of 10K and the duration of each simulation window set at 10 ns. The extension of the temperature ramp was determined based on the conservation of the native fold of the protein throughout the simulation, monitored through the backbone RMSD.

Structure preparation (for subsequent analyses with Bioluminate and Osprey)

The PDB entries 7ALI (*51*), 8DZ2 (*48*), and 7TLL (*49*) were used to investigate the impact of mutations on apo WT, nirmatrelvir bound WT and nirmatrelvir bound Omicron dimers. All structures were prepared with the default settings of the Protein Preparation Wizard in Maestro version 2022-3 (*95*) except that selenomethionines were converted to methionines and no water molecules were deleted. The variant mutations Omicron, Omicron+A206T, L167F, L167F+P168S and L167F+F305L were introduced manually into the entry 7ALI (*51*) L167F, L167F+P168S and L167F+F305L were introduced manually into 8DZ2 (*48*) and Omicron+A206T was manually introduced in 7TLL (*49*). All mutations were introduced in Maestro version 2022-3 (*95*, *96*). The effects of additional mutations found in these most prevalent variants were investigated with the methods described below. Please note, only the structure 7ALI (*51*) includes residue 305 in both chains, 8DZ2 (*48*) only includes residue 305 in chain B, and 7TLL does not include residue 305 at all. Structure visualization and figure creation were performed in PyMol (*97*) version 2.5.0. Protein alignment was conducted in Maestro release 2022-3.

Stability prediction

The stability of all the mutations was calculated using the default settings of the residue_scanning_backend.py module in Bioluminate (*44–47*) release 2022-3. Mutations were either introduced individually or in combination in the two protomers.

Affinity prediction

For the affinity predictions in Osprey version 3.3 (*50*, *98–100*) water molecules were not considered. Osprey calculates so called Log10 K* scores (*99*) which provide an estimation of binding affinity. The required yaml and frcmod input files were created as described in detail in the Guerin et al. STAR Protocol (*101*). For the dimer affinity calculations, chain B was considered as the ligand. The stability threshold was disabled, epsilon was defined as 0.03 and WT and mutant side chain conformations were calculated as continuously flexible. Osprey is an open source software and is available for free at https://github.com/donaldlab/OSPREY3.

References

1. I. Mohammed, A. Nauman, P. Paul, S. Ganesan, K. H. Chen, S. M. S. Jalil, S. H. Jaouni, H. Kawas, W. A. Khan, A. L. Vattoth, Y. A. Al-Hashimi, A. Fares, R. Zeghlache, D. Zakaria, The efficacy and effectiveness of the COVID-19 vaccines in reducing infection, severity, hospitalization, and mortality: a systematic review. *Hum. Vaccines Immunother.* **18** (2022), doi:10.1080/21645515.2022.2027160.

2. N. G. Davies, S. Abbott, R. C. Barnard, C. I. Jarvis, A. J. Kucharski, J. D. Munday, C. A. B. Pearson, T. W. Russell, D. C. Tully, A. D. Washburne, T. Wenseleers, A. Gimma, W. Waites, K. L. M. Wong, K. van Zandvoort, J. D. Silverman, K. Diaz-Ordaz, R. Keogh, R. M. Eggo, S. Funk, M. Jit, K. E. Atkins, W. J. Edmunds, Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science (80-.).* **372** (2021), doi:10.1126/science.abg3055.

3. T. Kustin, N. Harel, U. Finkel, S. Perchik, S. Harari, M. Tahor, I. Caspi, R. Levy, M. Leshchinsky, S. Ken Dror, G. Bergerzon, H. Gadban, F. Gadban, E. Eliassian, O. Shimron, L. Saleh, H. Ben-Zvi, E. Keren Taraday, D. Amichay, A. Ben-Dor, D. Sagas, M. Strauss, Y. Shemer Avni, A. Huppert, E. Kepten, R. D. Balicer, D. Netzer, S. Ben-Shachar, A. Stern, Evidence for increased breakthrough rates of SARS-CoV-2 variants of concern in BNT162b2-mRNA-vaccinated individuals. *Nat. Med.* **27**, 1379–1384 (2021).

4. A. Rössler, L. Riepler, D. Bante, D. von Laer, J. Kimpel, Correspondence SARS-CoV-2 Omicron Variant Neutralization in Serum from Vaccinated and Convalescent Persons, 1–3 (2021).

5. A. M. Carabelli, T. P. Peacock, L. G. Thorne, W. T. Harvey, J. Hughes, T. I. de Silva, S. J. Peacock, W. S. Barclay, T. I. de Silva, G. J. Towers, D. L. Robertson, SARS-CoV-2 variant biology: immune escape, transmission and fitness. *Nat. Rev. Microbiol.* **21**, 162–177 (2023).

6. C. K. H. Wong, I. C. H. Au, K. T. K. Lau, E. H. Y. Lau, B. J. Cowling, G. M. Leung, Real-world effectiveness of molnupiravir and nirmatrelvir plus ritonavir against mortality, hospitalisation, and in-hospital outcomes among community-dwelling, ambulatory patients with confirmed SARS-CoV-2 infection during the omicron wave in Hong Kong: a. *Lancet.* **400**, 1213–1222 (2022).

7. Y. Unoh, S. Uehara, K. Nakahara, H. Nobori, Y. Yamatsu, S. Yamamoto, Y. Maruyama, Y. Taoda, K. Kasamatsu, T. Suto, K. Kouki, A. Nakahashi, S. Kawashima, T. Sanaki, S. Toba, K. Uemura, T. Mizutare, S. Ando, M. Sasaki, Y. Orba, H. Sawa, A. Sato, T. Sato, T. Kato, Y. Tachibana, Discovery of S-217622, a Noncovalent Oral SARS-CoV-2 3CL Protease Inhibitor Clinical Candidate for Treating COVID-19. *J. Med. Chem.* (2022), doi:10.1021/acs.jmedchem.2c00117.

8. X. Chen, X. Huang, B. Zhou, J. Xu, B. Liu, H. Jiang, W. Zhang, C. Yang, S. Wu, J. Huang, H. Li, C. Long, X. Zhao, H. Xu, Y. Sheng, Y. Guo, C. Niu, L. Xue, Y. Xu, J. Liu, T. Zhang, J. Spencer, W. Deng, S. Chen, Z. Yang, N. Zhong, U. States, T. Technology, W. Free, T. Zone, X. Chen, X. Huang, SARS-CoV-2 has demonstrated extraordinary ability to evade antibody immunity by, 1–60 (2023).

9. W. Shang, W. Dai, C. Yao, L. Xu, X. Tao, H. Su, J. Li, X. Xie, Y. Xu, M. Hu, D. Xie, H. Jiang, L. Zhang, H. Liu, In vitro and in vivo evaluation of the main protease inhibitor

FB2001 against SARS-CoV-2. *Antiviral Res.* **208**, 105450 (2022).

10. J. R. Spengler, S. R. Welch, J. Deval, B. G. Gentry, A. Brancale, K. Carter, J. Moffat, C. Meier, K. L. Seley-Radtke, L. M. Schang, Meeting report: 35th International Conference on Antiviral Research in Seattle, Washington, USA – March 21–25, 2022. *Antiviral Res.* **211**, 105521 (2023).

11. C. E. Officer, Pardes Biosciences Announces Top-line Results from Phase 2 Trial Evaluating Pomotrelvir for the Treatment of COVID-19, 1–3 (2023).

12. C. N. M. P. A. (NMPA), NMPA conditionally approved the marketing of the combined packaging of Simnotrelvir Tablets/Ritonavir Tablets and Renmindevir Hydrobromide Tablets for COVID-19 treatment (2023); https://www.nmpa.gov.cn/zhuanti/yqyjzxd/yqyjxd/20230129143814163.html.

13. Y. Li, T. I. Zang, L. Xu, D. Leonard, K. Hoang, T. Greizer, S. Zhang, C. Foster, M. Huang, J. Kibel, J. Klaene, Y. S. Or, L.-J. Jiang, *Open Forum Infect. Dis.*, in press, doi:10.1093/ofid/ofac492.962.

14. A. K. Padhi, T. Tripathi, Hotspot residues and resistance mutations in the nirmatrelvir-binding site of SARS-CoV-2 main protease: Design, identification, and correlation with globally circulating viral genomes. *Biochem. Biophys. Res. Commun.* **629**, 54–60 (2022).

15. Y. Hu, E. M. Lewandowski, H. Tan, X. Zhang, R. T. Morgan, X. Zhang, L. M. C. Jacobs, S. G. Butler, M. V Gongora, J. Choy, X. Deng, Y. Chen, J. Wang, Naturally Occurring Mutations of SARS-CoV-2 Main Protease Confer Drug Resistance to Nirmatrelvir. *ACS Cent. Sci.* (2023), doi:10.1021/acscentsci.3c00538.

16. J. Clayton, V. M. de Oliveira, M. F. Ibrahim, X. Sun, P. Mahinthichaichan, M. Shen, R. Hilgenfeld, J. Shen, Integrative Approach to Dissect the Drug Resistance Mechanism of the H172Y Mutation of SARS-CoV-2 Main Protease. *J. Chem. Inf. Model.* (2023), doi:10.1021/acs.jcim.3c00344.

17. S. Iketani, H. Mohri, B. Culbertson, S. J. Hong, Y. Duan, M. I. Luck, M. K. Annavajhala, Y. Guo, Z. Sheng, A.-C. C. Uhlemann, S. P. Goff, Y. Sabo, H. Yang, A. Chavez, D. D. Ho, Multiple pathways for SARS-CoV-2 resistance to nirmatrelvir. *Nature*. **613**, 558–564 (2022).

18. Y. Zhou, K. A. Gammeltoft, L. A. Ryberg, L. V. Pham, H. D. Tjørnelund, A. Binderup, C. R. D. Hernandez, C. Fernandez-Antunez, A. Offersgaard, U. Fahnøe, G. H. J. Peters, S. Ramirez, J. Bukh, J. M. Gottwein, Nirmatrelvir-resistant SARS-CoV-2 variants with high fitness in an infectious cell culture system. *Sci. Adv.* **8** (2022), doi:10.1126/sciadv.add7197.

19. D. Jochmans, C. Liu, K. Donckers, A. Stoycheva, S. Boland, S. K. Stevens, C. De Vita, B. Vanmechelen, P. Maes, B. Trüeb, N. Ebert, V. Thiel, S. De Jonghe, L. Vangeel, D. Bardiot, A. Jekle, L. M. Blatt, L. Beigelman, J. A. Symons, P. Raboisson, P. Chaltin, A. Marchand, J. Neyts, J. Deval, K. Vandyck, "The Substitutions L50F, E166A, and L167F in SARS-CoV-2 3CLpro Are Selected by a Protease Inhibitor In Vitro and Confer Resistance To Nirmatrelvir" (2023), , doi:10.1128/mbio.02815-22.

20. S. A. Moghadasi, E. Heilmann, A. M. Khalil, C. Nnabuife, F. L. Kearns, C. Ye, S. N. Moraes, F. Costacurta, M. A. Esler, H. Aihara, D. von Laer, L. Martinez-Sobrido, T. Palzkill, R. E. Amaro, R. S. Harris, "Transmissible SARS-CoV-2 variants with resistance to clinical

protease inhibitors" (2023), , doi:10.1126/sciadv.ade8778.

21. L. Parigger, A. Krassnigg, T. Schopper, A. Singh, K. Tappler, K. Köchl, M. Hetmann, K. Gruber, G. Steinkellner, C. C. Gruber, Recent changes in the mutational dynamics of the SARS-CoV-2 main protease substantiate the danger of emerging resistance to antiviral drugs. *Front. Med.* **9**, 1–12 (2022).

22. A. M. Shaqra, S. N. Zvornicanin, Q. Y. J. Huang, G. J. Lockbaum, M. Knapp, L. Tandeske, D. T. Bakan, J. Flynn, D. N. A. Bolon, S. Moquin, D. Dovala, N. Kurt Yilmaz, C. A. Schiffer, Defining the substrate envelope of SARS-CoV-2 main protease to predict and avoid drug resistance. *Nat. Commun.* **13**, 1–11 (2022).

23. K. S. Yang, S. Z. Leeuwon, S. Xu, W. R. Liu, Evolutionary and Structural Insights about Potential SARS-CoV-2 Evasion of Nirmatrelvir. *J. Med. Chem.* **65**, 8686–8698 (2022).

24. J. M. Flynn, Q. Y. J. Huang, S. N. Zvornicanin, G. G. Schneider-Nachum, A. M. Shaqra, N. K. Yilmaz, S. A. Moquin, D. Dovala, C. A. Schiffer, D. N. A. Bolon, Systematic Analyses of the Resistance Potential of Drugs Targeting SARS-CoV-2 Main Protease. *ACS Infect. Dis.* (2023), doi:10.1021/acsinfecdis.3c00125.

25. S. Iketani, S. J. Hong, J. Sheng, F. Bahari, B. Culbertson, F. F. Atanaki, A. K. Aditham, A. F. Kratz, M. I. Luck, R. Tian, S. P. Goff, H. Montazeri, Y. Sabo, D. D. Ho, A. Chavez, Functional map of SARS-CoV-2 3CL protease reveals tolerant and immutable sites. *Cell Host Microbe*. **30**, 1354-1362.e6 (2022).

26. J. M. Flynn, N. Samant, G. Schneider-Nachum, D. T. Barkan, N. K. Yilmaz, C. A. Schiffer, S. A. Moquin, D. Dovala, D. N. A. Bolon, Comprehensive fitness landscape of SARS-CoV-2 Mpro reveals insights into viral resistance mechanisms. *Elife*. **11**, 1–27 (2022).

27. J. Ou, E. M. Lewandowski, Y. Hu, A. A. Lipinski, A. Aljasser, M. Colon-Ascanio, R. T. Morgan, L. M. C. Jacobs, X. Zhang, M. J. Bikowitz, P. R. Langlais, H. Tan, J. Wang, Y. Chen, J. S. Choy, A yeast-based system to study SARS-CoV-2 Mpro structure and to identify nirmatrelvir resistant mutations. *PLOS Pathog.* **19**, e1011592 (2023).

28. S. Lan, G. Neilsen, R. L. Slack, W. A. Cantara, A. E. Castaner, Z. C. Lorson, N. Lulkin, H. Zhang, J. Lee, M. E. Cilento, P. R. Tedbury, S. G. Sarafianos, *bioRxiv*, in press (available at https://www.biorxiv.org/content/10.1101/2022.12.31.522389v1%0Ahttps://www.biorxiv.org/content/10.1101/2022.12.31.522389v1.abstract).

29. S. Ullrich, K. B. Ekanayake, G. Otting, C. Nitsche, Main protease mutants of SARS-CoV-2 variants remain susceptible to nirmatrelvir. *Bioorganic Med. Chem. Lett.* **62**, 128629 (2022).

30. L. Vangeel, W. Chiu, S. De Jonghe, P. Maes, B. Slechten, J. Raymenants, E. André, P. Leyssen, J. Neyts, D. Jochmans, Remdesivir, Molnupiravir and Nirmatrelvir remain active against SARS-CoV-2 Omicron and other variants of concern. *Antiviral Res.* **198**, 10–12 (2022).

31. P. Li, Y. Wang, M. Lavrijsen, M. M. Lamers, A. C. de Vries, R. J. Rottier, M. J. Bruno, M. P. Peppelenbosch, B. L. Haagmans, Q. Pan, SARS-CoV-2 Omicron variant is highly sensitive to molnupiravir, nirmatrelvir, and the combination. *Cell Res.* **32**, 322–324 (2022).

32. S. Kawashima, Y. Matsui, T. Adachi, Y. Morikawa, K. Inoue, S. Takebayashi, H. Nobori, M. Rokushima, Y. Tachibana, T. Kato, Ensitrelvir is effective against SARS-CoV-2 3CL protease mutants circulating globally. *Biochem. Biophys. Res. Commun.* **645**, 132–136

(2023).

33. M. D. Sacco, Y. Hu, M. V. Gongora, F. Meilleur, M. T. Kemp, X. Zhang, J. Wang, Y. Chen, The P132H mutation in the main protease of Omicron SARS-CoV-2 decreases thermal stability without compromising catalysis or small-molecule drug inhibition. *Cell Res.* **32**, 498–500 (2022).

34. E. Heilmann, F. Costacurta, S. A. Moghadasi, C. Ye, M. Pavan, D. Bassani, A. Volland, C. Ascher, A. K. H. Weiss, D. Bante, R. S. Harris, S. Moro, B. Rupp, L. Martinez-Sobrido, D. von Laer, SARS-CoV-2 3CLpro mutations selected in a VSV-based system confer resistance to nirmatrelvir, ensitrelvir, and GC376. *Sci. Transl. Med.* **15**, 1–19 (2023).

35. J. W. Drake, J. J. Holland, Mutation rates among RNA viruses. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 13910–13913 (1999).

36. D. A. Steinhauer, E. Domingo, J. J. Holland, Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase. *Gene*. **122**, 281–288 (1992).

37. S. Khare, C. Gurry, L. Freitas, M. B. Schultz, G. Bach, A. Diallo, N. Akite, J. Ho, R. T. C. Lee, W. Yeo, G. C. C. Team, S. Maurer-Stroh, GISAID's Role in Pandemic Response. *China CDC Wkly.* **3**, 1049–1051 (2021).

38. Y. Shu, J. McCauley, GISAID: Global initiative on sharing all influenza data – from vision to reality. *Eurosurveillance*. **22** (2017), , doi:10.2807/1560-7917.ES.2017.22.13.30494.

39. S. Elbe, G. Buckland-Merrett, Data, disease and diplomacy: GISAID's innovative contribution to global health. *Glob. Challenges*. **1**, 33–46 (2017).

40. S. Ullrich, C. Nitsche, The SARS-CoV-2 main protease as drug target. *Bioorganic Med. Chem. Lett.* **30**, 127377 (2020).

41. K. Anand, J. Ziebuhr, P. Wadhwani, J. R. Mesters, R. Hilgenfeld, (3CL pro) Structure : Basis for Design of Anti-SARS Drugs. *Science (80-.).* **1763**, 1763–1768 (2014).

42. S. E. Witko, C. S. Kotash, R. M. Nowak, J. E. Johnson, L. A. C. Boutilier, K. J. Melville, S. G. Heron, D. K. Clarke, A. S. Abramovitz, R. M. Hendry, M. S. Sidhu, S. A. Udem, C. L. Parks, An efficient helper-virus-free method for rescue of recombinant paramyxoviruses and rhadoviruses from a cell line suitable for vaccine development. *J. Virol. Methods*. **135**, 91–101 (2006).

43. E. Heilmann, F. Costacurta, S. Geley, S. A. Mogadashi, A. Volland, B. Rupp, R. S. Harris, D. von Laer, A VSV-based assay quantifies coronavirus Mpro/3CLpro/Nsp5 main protease activity and chemical inhibition. *Commun. Biol.* **5**, 1–12 (2022).

44. K. Zhu, T. Day, D. Warshaviak, C. Murrett, R. Friesner, D. Pearlman, Antibody structure determination using a combination of homology modeling, energy-based refinement, and loop prediction. *Proteins Struct. Funct. Bioinforma.* **82**, 1646–1655 (2014).

45. N. K. Salam, M. Adzhigirey, W. Sherman, D. A. Pearlman, D. Thirumalai, Structure-based approach to the prediction of disulfide bonds in proteins. *Protein Eng. Des. Sel.* **27**, 365–374 (2014).

46. H. Beard, A. Cholleti, D. Pearlman, W. Sherman, K. A. Loving, Applying physics-based scoring to calculate free energies of binding for single amino acid mutations in protein-protein complexes. *PLoS One*. **8**, 1–11 (2013).

47. Schrödinger Release 2023-2: BioLuminate, Schrödinger, LLC, New York, NY, 2023.

48. G. D. Noske, E. de Souza Silva, M. O. de Godoy, I. Dolci, R. S. Fernandes, R. V. C. Guido, P. Sjö, G. Oliva, A. S. Godoy, Structural basis of nirmatrelvir and ensitrelvir activity against naturally occurring polymorphisms of the SARS-CoV-2 main protease. *J. Biol. Chem.* **299**, 103004 (2023).

49. S. E. Greasley, S. Noell, O. Plotnikova, R. A. Ferre, W. Liu, B. Bolanos, K. Fennell, J. Nicki, T. Craig, Y. Zhu, A. E. Stewart, C. M. Steppan, Structural basis for the in vitro efficacy of nirmatrelvir against SARS-CoV-2 variants. *J. Biol. Chem.* **298**, 101972 (2022).

50. M. A. Hallen, J. W. Martin, A. Ojewole, J. D. Jou, A. U. Lowegard, M. S. Frenkel, P. Gainza, H. M. Nisonoff, A. Mukund, S. Wang, G. T. Holt, D. Zhou, E. Dowd, B. R. Donald, OSPREY 3.0: Open-source protein redesign for you, with powerful new features. *J. Comput. Chem.* **39**, 2494–2507 (2018).

51. E. Costanzi, M. Kuzikov, F. Esposito, S. Albani, N. Demitri, B. Giabbai, M. Camasta, E. Tramontano, G. Rossetti, A. Zaliani, P. Storici, Structural and biochemical analysis of the dual inhibition of MG-132 against SARS-CoV-2 main protease (Mpro/3CLpro) and human cathepsin-L. *Int. J. Mol. Sci.* **22** (2021), doi:10.3390/ijms222111779.

52. M. Pavan, S. Menin, D. Bassani, M. Sturlese, S. Moro, Qualitative Estimation of Protein-Ligand Complex Stability through Thermal Titration Molecular Dynamics Simulations. *J. Chem. Inf. Model.* **62**, 5715–5728 (2022).

53. M. Pavan, S. Menin, D. Bassani, M. Sturlese, S. Moro, Implementing a Scoring Function Based on Interaction Fingerprint for Autogrow4: Protein Kinase CK1δ as a Case Study. *Front. Mol. Biosci.* **9**, 1–17 (2022).

54. S. Menin, M. Pavan, V. Salmaso, M. Sturlese, S. Moro, Thermal Titration Molecular Dynamics (TTMD): Not Your Usual Post-Docking Refinement. *Int. J. Mol. Sci.* **24** (2023), doi:10.3390/ijms24043596.

55. E. Fornasier, M. L. Macchia, G. Giachin, A. Sosic, M. Pavan, M. Sturlese, C. Salata, S. Moro, B. Gatto, M. Bellanda, R. Battistutta, A new inactive conformation of SARS-CoV-2 main protease. *Acta Crystallogr. Sect. D Struct. Biol.* **78**, 363–378 (2022).

56. J. D. Bloom, R. A. Neher, *bioRxiv*, in press (available at https://www.biorxiv.org/content/10.1101/2023.01.30.526314v1%0Ahttps://www.biorxiv.org/content/10.1101/2023.01.30.526314v1.abstract).

57. S. Albani, E. Costanzi, G. L. Hoang, M. Frings, N. Ansari, N. Demitri, T. Toan, Unexpected single ligand occupancy and negative cooperativity in the SARS-CoV-2 Main protease, doi:10.26434/chemrxiv-2023-n1x1n D O I: 10.26434/chemrxiv-2023-n1x1n [opens in a new tab].

58. S. Chen, F. Jonas, C. Shen, R. Higenfeld, Liberation of SARS-CoV main protease from the viral polyprotein: N-terminal autocleavage does not depend on the mature dimerization mode. *Protein Cell*. **1**, 59–74 (2010).

59. F. Schmidt, Y. Weisblum, M. Rutkowska, D. Poston, J. Da Silva, F. Zhang, E. Bednarski, A. Cho, D. J. Schaefer-Babajew, C. Gaebler, M. Caskey, M. C. Nussenzweig, T. Hatziioannou, P. D. Bieniasz, High genetic barrier to SARS-CoV-2 polyclonal neutralizing antibody escape. *Nature*. **600**, 512–516 (2021).

60. Y. Yahalom-Ronen, N. Erez, M. Fisher, H. Tamir, B. Politi, H. Achdout, S. Melamed, I. Glinert, S. Weiss, I. Cohen-Gihon, O. Israeli, M. Izak, M. Mandelboim, Y. Caraco, N. Madar-Balakirski, A. Mechaly, E. Shinar, R. Zichel, D. Cohen, A. Beth-Din, A. Zvi, H. Marcus, T. Israely, N. Paran, Neutralization of SARS-CoV-2 Variants by rVSV-ΔG-Spike-Elicited Human Sera. *Vaccines*. **10** (2022), , doi:10.3390/vaccines10020291.

61. Z. Liu, L. A. VanBlargan, L. M. Bloyet, P. W. Rothlauf, R. E. Chen, S. Stumpf, H. Zhao, J. M. Errico, E. S. Theel, M. J. Liebeskind, B. Alford, W. J. Buchser, A. H. Ellebedy, D. H. Fremont, M. S. Diamond, S. P. J. Whelan, Identification of SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization. *Cell Host Microbe*. **29**, 477-488.e4 (2021).

62. T. W. Geisbert, H. Feldmann, Recombinant vesicular stomatitis virus-based vaccines against Ebola and marburg virus infections. *J. Infect. Dis.* **204**, 1075–1081 (2011).

63. A. Muik, I. Kneiske, M. Werbizki, D. Wilflingseder, T. Giroglou, O. Ebert, A. Kraft, U. Dietrich, G. Zimmer, S. Momma, D. von Laer, Pseudotyping Vesicular Stomatitis Virus with Lymphocytic Choriomeningitis Virus Glycoproteins Enhances Infectivity for Glioma Cells and Minimizes Neurotropism. *J. Virol.* **85**, 5679–5684 (2011).

64. M. Romano, A. Ruggiero, F. Squeglia, G. Maga, R. Berisio, A Structural View of SARS-CoV-2 RNA Replication Machinery: RNA Synthesis, Proofreading and Final Capping. *Cells*. **9** (2020), , doi:10.3390/cells9051267.

65. D. A. Steinhauer, J. J. Holland, Direct method for quantitation of extreme polymerase error frequencies at selected single base sites in viral RNA. *J. Virol.* **57**, 219–228 (1986).

66. Z. Sun, L. Wang, X. Li, C. Fan, J. Xu, Z. Shi, H. Qiao, Z. Lan, X. Zhang, L. Li, X. Zhou, Y. Geng, An extended conformation of SARS-CoV-2 main protease reveals allosteric targets. *Proc. Natl. Acad. Sci. U. S. A.* **119**, 1–9 (2022).

67. A. Dodaro, M. Pavan, S. Moro, Targeting the I7L Protease: A Rational Design for Anti-Monkeypox Drugs? *Int. J. Mol. Sci.* **24** (2023), doi:10.3390/ijms24087119.

68. D. Panda, P. X. Dinh, L. K. Beura, A. K. Pattnaik, Induction of Interferon and Interferon Signaling Pathways by Replication of Defective Interfering Particle RNA in Cells Constitutively Expressing Vesicular Stomatitis Virus Replication Proteins. *J. Virol.* **84**, 4826–4831 (2010).

69. W. Shen, S. Le, Y. Li, F. Hu, SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One*. **11**, 1–10 (2016).

70. H. Li, Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics*. **34**, 3094–3100 (2018).

71. P. Danecek, J. K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M. O. Pollard, A. Whitwham, T. Keane, S. A. McCarthy, R. M. Davies, Twelve years of SAMtools and BCFtools. *Gigascience*. **10**, 1–4 (2021).

72. A. Wilm, P. P. K. Aw, D. Bertrand, G. H. T. Yeo, S. H. Ong, C. H. Wong, C. C. Khor, R. Petric, M. L. Hibberd, N. Nagarajan, LoFreq: A sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. *Nucleic Acids Res.* **40**, 11189–11201 (2012).

73. Y. Turakhia, B. Thornlow, A. S. Hinrichs, N. De Maio, L. Gozashti, R. Lanfear, D. Haussler, R. Corbett-Detig, Ultrafast Sample placement on Existing tRees (UShER) enables real-time phylogenetics for the SARS-CoV-2 pandemic. *Nat. Genet.* **53**, 809–816 (2021).

74. E. H. Abdelkader, G. Otting, NT*-HRV3CP: An optimized construct of human rhinovirus 14 3C protease for high-yield expression and fast affinity-tag cleavage. *J. Biotechnol.* **325**, 145–151 (2021).

75. R. Röck, J. E. Mayrhofer, O. Torres-Quesada, F. Enzler, A. Raffeiner, P. Raffeiner, A. Feichtner, R. G. Huber, S. Koide, S. S. Taylor, J. Troppmair, E. Stefan, BRAF inhibitors promote intermediate BRAF(V600E) conformations and binary interactions with activated RAS. *Sci. Adv.* **5**, 14–16 (2019).

76. J. E. Mayrhofer, F. Enzler, A. Feichtner, R. Röck, J. Fleischmann, A. Raffeiner, P. Tschaikner, E. Ogris, R. G. Huber, M. Hartl, R. Schneider, J. Troppmair, O. Torres-Quesada, E. Stefan, Mutation-oriented profiling of autoinhibitory kinase conformations predicts RAF inhibitor efficacies. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 31105–31113 (2020).

77. C. Dold, C. Rodriguez Urbiola, G. Wollmann, L. Egerer, A. Muik, L. Bellmann, H. Fiegl, C. Marth, J. Kimpel, D. Von Laer, Application of interferon modulators to overcome partial resistance of human ovarian cancers to VSV-GP oncolytic viral therapy. *Mol. Ther. - Oncolytics*. **3**, 16021 (2016).

78. G. Kärber, Beitrag zur kollektiven Behandlung pharmakologischer Reihenversuche. *Naunyn. Schmiedebergs. Arch. Exp. Pathol. Pharmakol.* **162**, 480–483 (1931).

79. J. Singh, R. C. Petter, T. A. Baillie, A. Whitty, The resurgence of covalent drugs. *Nat. Rev. Drug Discov.* **10**, 307–317 (2011).

80. H. M. Berman, The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).

81. Molecular Operating Environment (MOE), Molecular Operating Environment (MOE), 2019.01; Chemical Computing Group ULC, 1010 Sherbooke St. West, Suit. *Sci. Comput. Instrum.*, 32 (2009).

82. W. Humphrey, A. Dalke, K. Schulten, Sartorius products. *J. Mol. Graph.* **14**, 33–38 (1996).

83. D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, R. J. Woods, The Amber biomolecular simulation programs. *J. Comput. Chem.* **26**, 1668–1688 (2005).

84. D. A. Case, H. M. Aktulga, K. Belfon, I. Y. Ben-Shalom, J. T. Berryman, S. R. Brozell, D. S. Cerutti, T. E. I. Cheatham, G. A. Cisneros, V. W. . Cruzeiro, T. A. Darden, R. E. Duke, G. Giambasu, M. K. Gilson, H. Gohlke, A. W. Goetz, R. Harris, S. Izadi, S. A. Izmailov, K. Kasavajhala, M. C. Kaymak, E. King, A. Kovalenko, T. Kurtzman, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, M. Machado, V. Man, M. Manathunga, K. M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, K. A. O'Hearn, A. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rhanamoun, D. R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, A. Shajan, J. Shen, C. L. Simmerling, N. R. Skrynnikov, J. Smith, J. Swails, R. C. Walker, H. Wang, J. Wang, H. Wei, R. M. Wolf, X. Wu, Y. Xiong, Y. Xue, D. M. York, S. Zhao, P. A. Kollman, Amber 2022, University of California, San Francisco. (2022).

85. J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from

ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).

86. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, Development and testing of a general Amber force field. *J. Comput. Chem.* **25**, 1157–1174 (2004).

87. A. Jakalian, D. B. Jack, C. I. Bayly, Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* **23**, 1623–1641 (2002).

88. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).

89. R. L. Davidchack, R. Handel, M. V. Tretyakov, Langevin thermostat for rigid body dynamics. *J. Chem. Phys.* **130** (2009), doi:10.1063/1.3149788.

90. R. Faller, J. J. De Pablo, Constant pressure hybrid molecular dynamics-Monte Carlo simulations. *J. Chem. Phys.* **116**, 55–59 (2002).

91. P. Eastman, J. Swails, J. D. Chodera, R. T. McGibbon, Y. Zhao, K. A. Beauchamp, L. P. Wang, A. C. Simmonett, M. P. Harrigan, C. D. Stern, R. P. Wiewiora, B. R. Brooks, V. S. Pande, OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Comput. Biol.* **13**, 1–17 (2017).

92. N. Michaud-Agrawal, E. J. Denning, T. B. Woolf, O. Beckstein, MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **32**, 2319–2327 (2011).

93. R. Gowers, M. Linke, J. Barnoud, T. Reddy, M. Melo, S. Seyler, J. Domański, D. Dotson, S. Buchoux, I. Kenney, O. Beckstein, MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. *Proc. 15th Python Sci. Conf.*, 98–105 (2016).

94. M. Wójcikowski, P. Zielenkiewicz, P. Siedlecki, Open Drug Discovery Toolkit (ODDT): A new open-source player in the drug discovery field. *J. Cheminform.* **7**, 1–6 (2015).

95. G. Madhavi Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, W. Sherman, Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided. Mol. Des.* **27**, 221–234 (2013).

96. Schrödinger Release 2022-3: Protein Preparation Wizard; Epik, Schrödinger, LLC, New York, NY, 2022; Impact, Schrödinger, LLC, New York, NY; Prime, Schrödinger, LLC, New York, NY, 2022.

97. The PyMOL Molecular Graphics System, Version 2.5.0 Schrödinger, LLC.

98. I. Georgiev, R. H. Lilien, B. R. Donald, The minimized dead-end elimination criterion and its application to protein redesign in a hybrid scoring and search algorithm for computing partition functions over molecular ensembles. *J. Comput. Chem.* **29**, 1527–1542 (2008).

99. R. H. Lilien, B. W. Stevens, A. C. Anderson, B. R. Donald, A Novel Ensemble-Based Scoring and Search Algorithm for Protein Redesign and Its Application to Modify the Substrate Specificity of the Gramicidin Synthetase A Phenylalanine Adenylation Enzyme. *J. Comput. Biol.* **12**, 740–761 (2005).

100. K. E. Roberts, P. R. Cushing, P. Boisguerin, D. R. Madden, B. R. Donald, Computational design of a PDZ domain peptide inhibitor that rescues CFTR activity. *PLoS Comput. Biol.* **8** (2012), doi:10.1371/journal.pcbi.1002477.

101. N. Guerin, A. Feichtner, E. Stefan, T. Kaserer, B. R. Donald, Resistor: An algorithm for predicting resistance mutations via Pareto optimization over multistate protein design and mutational signatures. *Cell Syst.* **13**, 830-843.e3 (2022).

Acknowledgements

We thank Florian Krammer, Bernhard Rupp and Sebastiaan Werten for useful discussions. The MMS lab is grateful to Chemical Computing Group, OpenEye, and Acellera for the scientific and technical partnership. MMS lab thankfully acknowledges the support of NVIDIA Corporation with the donation of the Titan V GPU used for this research.

Funding

This work was funded by:

Austrian Science Fund (FWF) grant P35148

Austrian Science Fund (FWF) grant P34376.

National Institute of Allergy and Infectious Disease grant U19-AI171954.

Author contributions

E.H. conceived the initial concept. F.C., E.H., A.D., H.S., B.S., J.F. and D.v.L. designed the experiments. E.H. and F.C. conceived the cloning strategies. E.H., F.C. generated recombinant viruses. E.H., F.C., S.R. and L.K. cloned plasmids for resistance experiments. E.H. and F.C. performed selection and resistance experiments with VSV-based methods and biochemical assays. A. Sauerwein, A.H. and T.R. performed viral RNA isolation and Nanopore library preparation. D. Bante elaborated Nanopore sequencing data. B.S. provided recombinant M^{pro} with point mutations expressed in E. coli. J.F. performed western blot analyses. S.A.M. provided phylogenetic analyses. J.H. provided substitution frequencies from the GISAID database. F.C. and E.H. wrote the manuscript and designed figures. A.D., M.P., D. Bassani, S. Menin and S. Moro performed structural modelling. H.S. and T.K. performed stability and dimerization affinity calculations. S. Moro and T.K. supervised in silico computational analyses. R.S. supervised B.S. recombinant M^{pro} production and E.S. supervised J.K. cross-validation assay. All authors read and approved the final manuscript.

Competing interests

D.v.L. is founder of ViraTherapeutics GmbH. D.v.L serves as a scientific advisor to Boehringer Ingelheim and Pharma KG. E.H. and D.v.L have received an Austrian Science Fund (FWF) grant in the special call "SARS urgent funding". E.H. is a registered consultant at Guidepoint. D. Bante holds stocks of Pfizer Inc. and Oxford Nanopore Technologies plc. S.A.M and R.S.H. are inventors on the patent application "Live cell assay for protease inhibition", application number WO/2022/094463. All other authors declare they have no competing interests.

Data and materials availability

All data associated with this study are in the paper or supplementary materials. Further data can be obtained by contacting francesco.costacurta@i-med.ac.at, emmanuel.heilmann@i-med.ac.at and virologie@i-med.ac.at. Materials can be obtained by contacting listed Email addresses and after completion of a materials transfer agreement.