Naturally occurring mutations of SARS-CoV-2 main protease confer drug resistance to nirmatrelvir

Yanmei Hu,[1,†] Eric M. Lewandowski,[2,†] Haozhou Tan,[1,†] Ryan T. Morgan,[2] Xiujun Zhang,[2] Lian M. C. Jacobs,[2] Shane G. Butler,[2] Maura V. Mongora,[2] John Choy,[3] Yu Chen,[2,*] Jun Wang[1,*]

[1]Department of Medicinal Chemistry, Ernest Mario School of Pharmacy, Rutgers, the State University of New Jersey, NJ, 08854, United States

[2]Department of Molecular Medicine, Morsani College of Medicine, University of South Florida, Tampa, FL, 33612, United States

[3]Department Biology, School of Arts and Sciences, the Catholic University of America, Washington DC, 20064, United States

[†]These authors contributed equally to this work.

*Corresponding authors:

Jun Wang, Tel: 828-445-6488, email: junwang@pharmacy.rutgers.edu

Yu Chen, Tel: 813-974-7809, email: ychen1@usf.edu

Keywords: SARS-CoV-2, COVID-19, main protease, 3CL protease, Paxlovid, nirmatrelvir, drug resistance.

ABSTRACT

The SARS-CoV-2 main protease (M^{pro}) is a cysteine protease and a validated antiviral drug target. Paxlovid is an FDA-approved oral COVID-19 antiviral that contains the M^{pro} inhibitor nirmatrelvir and the metabolic booster ritonavir. The emergence of SARS-CoV-2 variants mutations in the M^{pro} raised the alarm of potential drug resistance. In this study, we aim to discover M^{pro} drug resistant mutants from naturally observed polymorphisms. Through analyzing the SARS-CoV-2 sequences deposited in Global initiative on Sharing Avian influenza Data (GISAID) database, we identified 66 prevalent M^{pro} mutations located at the nirmatrelvir binding site. The M^{pro} mutant proteins were expressed and characterized for enzymatic activity and nirmatrelvir inhibition. While the majority of the M^{pro} mutants had reduced enzymatic activity (k_{cat}/K_m >10-fold decrease), 11 mutants including S144M/F/A/G/Y, M165T, E166Q, H172Q/F, and Q192T/S/V showed comparable enzymatic activity as the wild-type (k_{cat}/K_m <10-fold change) and resistance to nirmatrelvir (K_i > 10-fold increase). We further demonstrate that the enzymatic activity and inhibitor resistance of these single mutations can be enhanced by additional substitutions in a double mutant. X-ray crystal structures were determined for six of the single mutants with and/or without GC-376/nirmatrelvir. The structures illustrate how mutations can reduce ligand binding by impacting the conformational stability of the active site. Overall, our study identified several drug resistant hot spots that warrant close monitoring for possible clinical evidence of Paxlovid resistance.

The ongoing COVID-19 pandemic highlights the urgent need of orally bioavailable antiviral drugs. There are currently two FDA-approved SARS-CoV-2 oral drugs, molnupiravir and Paxlovid. Paxlovid is a combination of viral main protease (M^{pro} or $3CL^{pro}$) inhibitor nirmatrelvir and the metabolic booster ritonavir. SARS-CoV-2 infected patients receiving Paxlovid within three days of symptoms onset showed an 89% lower risk of hospitalization/death compared to patients who received placebo (*1*). M^{pro} is a cysteine protease that mediates the cleavage of viral polyproteins during viral replication and has a stringent substrate preference of glutamine at the P1 position, rending it a high-profile antiviral drug target. Structurally disparate M^{pro} inhibitors have been identified from screening and rational design (*2*). In addition to Paxlovid, other M^{pro} inhibitors advanced to clinical stage include PF-07304814 (phosphate form of PF-00835231), S-217622, PBI-0451, EDP-235, and 13b (*3*). As an RNA virus, SARS-CoV-2 continues to evolve either with or without selection pressure. Recent emergence of variants of concern, particularly the Omicron variant, raising the concern of possible altered susceptibility to vaccines and antiviral drugs. In this study, we report the discovery of drug resistant M^{pro} mutants from naturally occurring SARS-CoV-2 polymorphism deposited in the GISAID database. Three inhibitors GC-376, PF-00835231, and nirmatrelvir were examined for drug resistance. GC-376 is a veterinary drug candidate for the treatment of feline infectious peritonitis virus (FIPV) infection in cats (*4*). The phosphonate prodrug of PF-00835231 is developed by Pfizer as an intravenous drug for the treatment of COVID patient in hospitals and was advanced to clinical trial (*5*). Nirmatrelvir is the active component in Paxlovid (*6*). The drug sensitivity was characterized using FRET-based enzymatic assay and the thermal shift binding

assay. The X-ray crystal structures of the representative M^{pro} mutants S144A, S144L, M165Y, E166Q, H172Y, and Q192T with and/or without GC-376/nirmatrelvir provide a structural explanation for the drug resistance. Taken together, this study identified several nirmatrelvir resistant hot spots that warrants close monitoring, while highlighting the future risk of mutants with multiple substitutions at these sites.

Identification of SARS-CoV-2 M^{pro} mutants from GISAID sequence analysis

Recent sequence analysis of SARS-CoV-2 M^{pro} revealed multiple prevalent mutations including G15S (C.37 Lambda), T21I (B.1.1.318), L89F (B.1.2), K90R (B.1.351 Beta), P108S (B.1.1.284), P132H (B.1.1.529 Omicron), and L205V (P.2 Zeta) (*7-9*). All these mutants are located outside the nirmatrelvir binding site (Fig. 1A) and were found to have similar catalytic efficacy (k_{cat}/K_m) as the wild-type (WT) protein (*7, 10*). These mutants remained susceptible to nirmatrelvir with no significant IC_{50} or K_i value shifts observed (< 2-fold) (*7*). Although the results reassure that the efficacy of nirmatrelvir is not compromised by the current dominant SARS-CoV-2 variants, drug resistance to nirmatrelvir is anticipated given the experience from the clinic use of HIV and HCV protease inhibitors (*11, 12*). Several studies have been conducted to evolve or predict the nirmatrelvir resistant M^{pro} mutants (*13-16*).

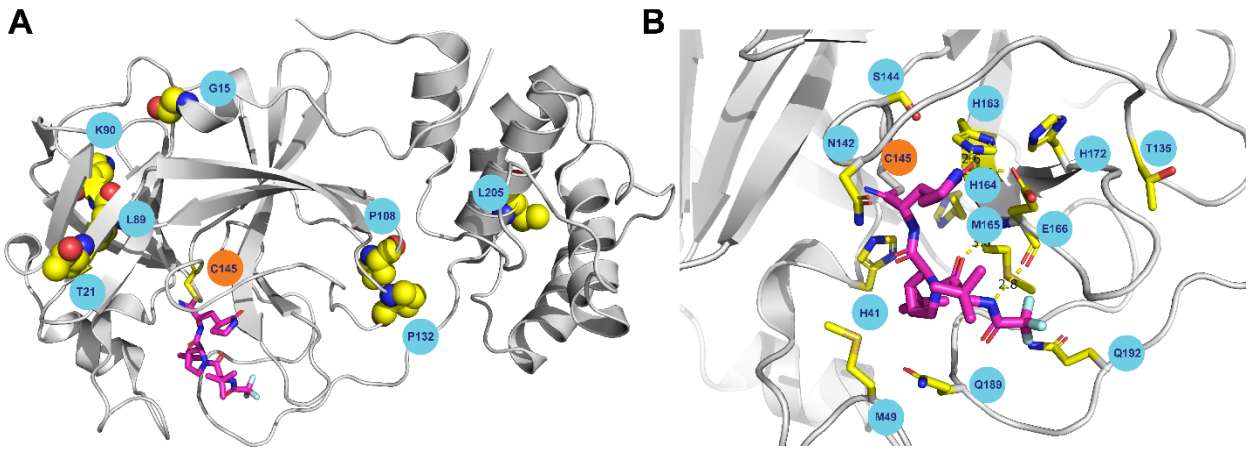

Fig. 1. SARS-CoV-2 M^{pro} mutants identified from GISAID sequence analysis. (**A**) Residues with high mutation rates that were examined in previous studies. None of the mutants showed significant drug resistance. (**B**) Residues located within 6 Å of the nirmatrelvir binding site in SARS-CoV-2 M^{pro} that were examined in this study. Figures were generated using Pymol with the X-ray crystal structure of nirmatrelvir in complex with SARS-CoV-2 M^{pro} (PDB: 7SI9). Nirmatrelvir is colored in magenta.

To identify drug resistant mutants of M^{pro}, we focus on the active site residues that are located within 6 Å of the nirmatrelvir binding site (PDB: 7SI9) (*6*). In total, 12 residues were selected including H41, M49, T135, N142, S144, H163, H164, M165, E166, H172, Q189, and Q192 (Fig. 1B). We expect that mutations at these active stie residues will have a direct impact on substrate binding and drug inhibition. To test this hypothesis, we analyzed the mutations of these 12 residues using the SARS-CoV-2 sequences deposited in GISAID (*17*). In total 5,124,505 sequences were analyzed, and the mutation frequency of each active site residue was plotted in Fig. 2A.

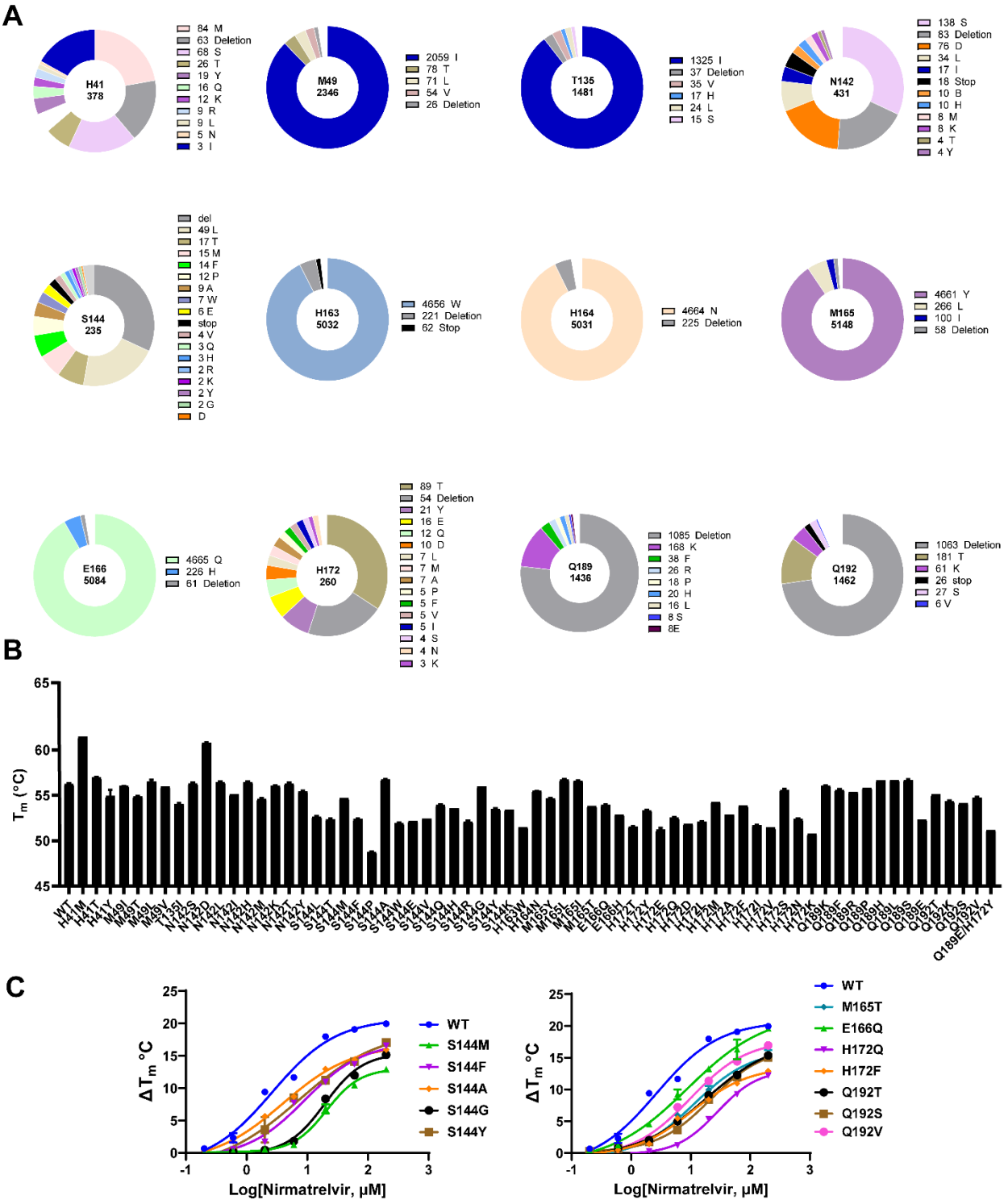

Fig. 2. SARS-CoV-2 M^{pro} mutants characterized in this study. (**A**) Mutations at 12 residues located at the nirmatrelvir binding site. Sequence data were obtained from CoVsurver of the GISAD, developed by A*STAR Bioinformatics Institute (BII), Singapore.

A total of 5,124,505 mutations of NSP5 (M^{pro}) were obtained from the database as of June 21, 2022. Occurrence of total mutation for each amino acid residue is labeled in the center of each pie chart. Occurrences of specific mutations are labeled on the right of its affiliate pie chart. (**B**) Thermal shift assay results of M^{pro} mutants. (**C**) Characterization of the nirmatrelvir resistance against M^{pro} mutants by the thermal shift assay. The results are the average of duplicates.

Based on the sequence analysis, we chose 66 mutants that cover all high frequency mutants at these 12 residues. The tag-free recombinant SARS-CoV-2 M^{pro} mutants with native N- and C-termini were expressed in *E. Coli* and purified (fig. S1). All proteins were folded correctly as shown by the thermal shift assay (Fig. 2B). The enzymatic activity (K_m, V_{max}, k_{cat}, and k_{cat}/K_m) of the mutant proteins was determined in the kinetic studies using the FRET-based enzymatic assay (*18, 19*). To profile the drug resistance, we tested the M^{pro} mutants against nirmatrelvir in the FRET assay and determined the half maximal inhibitory concentration (IC_{50}) and inhibition constant (K_i). For mutants showing resistance against nirmatrelvir, the drug sensitivity was further characterized for cross-resistance against PF-00835231 and GC-376. The comprehensive data set is shown in table S1.

The M^{pro} mutants were grouped into three categories based on the enzymatic activity k_{cat}/K_m value and drug resistance K_i value: category 1 includes residues that are critical for the enzymatic activity (H41 and H163); category 2 includes residues that are hot

spots for drug resistance (S144, M165, E166, H172, and Q192); and category 3 includes residues that can tolerate multiple mutations without significantly affecting enzymatic activity and drug inhibition (T135, M49, N142, H164, and Q189).

H41 and H163 are critical for the enzymatic activity

Among the 66 mutants, H41M, H41T, H41Y, and H163W were enzymatically inactive (table S1). H41 forms the catalytic dyad with C145, and our results showed that the top three most frequent mutants at this residue H41M (84 occurrences) H41T (26 occurrences), and H41Y (19 occurrences) were detrimental to the enzymatic activity (table S1). The X-ray crystal structure showed that the side chain imidazole of the buried H163 forms a hydrogen bond with the carbonyl from the nirmatrelvir P1 pyrrolidone (Fig. 1B) (PDB: 7SI9) (*6*). Majority of the reported M^{pro} inhibitors also contain P1 substitutions that form a hydrogen bond with H163 side chain imidazole, suggesting H163 is an essential residue for drug binding. Its extensive interactions with nearby residues such as F140 and Y161 in a tightly packed protein microenvironment also indicates a critical role in protein folding. Although H163W is a high frequency mutation with 4656 occurrences, this mutant led to inactive enzyme. Overall, mutations at H41 and H163 are unlikely to be tolerated in circulating SARS-CoV-2 variants.

S144, M165, E166, H172, and Q192 are hot spots for drug resistance

For the remaining 62 enzymatically active M^{pro} mutants, we applied two criteria to select the drug resistant mutants with clinical relevance, one is the enzymatic activity k_{cat}/K_m,

and another is the drug inhibitory constant K_i. It is generally assumed that mutations with impaired enzymatic activity will lead to attenuation of viral replication. We therefore focused on the M^{pro} mutants with k_{cat}/K_m values within 10-fold variation compared to WT. For drug sensitivity, a K_i increase by more than 10-fold is defined as a significant resistance. In total 12 M^{pro} mutants met both criteria including S144M, S144F, S144A, S144G, S144Y, M165T, E166Q, H172Q, H172F, Q192T, Q192S, and Q192V.

S144 locates at the S1 pocket and is part of the oxyanion hole consisting of two additional residues G143 and C145 (Fig. 1B). Mutations at this residue is expected to have a profound impact on enzymatic activity and drug inhibition. Among the top 15 high frequency mutations, five mutants S144M (8.0-fold lower k_{cat}/K_m value), S144F (5.8-fold), S144A (1.8-fold), S144G (2.6-fold), and S144Y (7.8-fold) had comparable enzymatic activity as the WT. Significantly, all five mutants showed drug resistance against nirmatrelvir with K_i values increase between 19.2 to 38.0-fold. Pfizer's report for healthcare providers similarly disclosed S144A as a nirmatrelvir resistant mutant with a K_i increase by 91.9-fold (*20*) compared to 20.5-fold from our study. Four mutants S144L (183.3-fold lower in k_{cat}/K_m value), S144P (523.8-fold), S144R (478.3-fold), and S144K (534.0-fold) had significantly reduced enzymatic activity compared to WT and also increased resistance to rimatrelvir. Similarly, the remaining six mutants S144T, S144W, S144E, S144V, S144Q, and S144H had compromised enzymatic activity with k_{cat}/K_m value decreased between 20.0 to 70.0-fold compared to WT. A significant drug resistance against nirmatrelvir was also observed for these six mutants.

M165 locates at the S2 pocket and forms hydrophobic interaction with the P2 dimethylcyclopropylproline (Fig. 1B). The most frequent mutant M165Y had a

significantly reduced enzymatic activity (41.7-fold decrease in k_{cat}/K_m value), while the other three mutants M165L (1.2-fold increase), M165I (1.4-fold decrease), and M165T (8.3-fold decrease) had similar enzymatic activity as the WT. No drug resistance was observed for M165L and M165I. However, a significant drug resistance against nirmatrelvir was observed for M165T (29.9-fold increase in K_i value).

E166 locates at the S1 pocket and forms three hydrogen bonds with nirmatrelvir (Fig. 1B). Therefore, E166 mutation is expected affect drug binding. E166Q is a high frequency mutation with 4665 occurrences. It has the same enzymatic activity (k_{cat}/K_m) as the WT, while being resistant to all three drugs, nirmatrelvir (11.7-fold increase in K_i value), PF-00835231 (12.9-fold), and GC-376 (2.1-fold). E166H had reduced enzymatic activity (17.5-fold lower in k_{cat}/K_m value) and showed drug resistance (368.7-fold increase in IC_{50}). In parallel to our study, two preprints reported nirmatrelvir resistant M^{pro} mutants identified from serial viral passage experiments in cell culture (*13, 14*). Jochmans et al. discovered a triple mutant L50F/E166A/L167F with a 72-fold increase in the enzymatic assay and a 51-fold increase in the antiviral assay against nirmatrelvir (*14*). However, the triple mutant only had 5.3% of the enzymatic activity of the WT, indicating an impaired fitness of replication. In another study, Zhou et al. showed that the L50F/E166V double mutant led to an 80-fold resistance in the antiviral assay against nirmatrelvir (*13*). Both E166A and E166V are natural occurring mutations with 5 and 7 occurrences. It is gratifying that our sequence analysis approach similarly identified E166 as a hot spot for nirmatrelvir resistance.

H172 locates at the S1 pocket but does not directly interact with nirmatrelvir (Fig. 1B). Among the 14 H172 mutants examined, H172Q (3.2-fold lower k_{cat}/K_m value) and H172F

(9.9-fold) had comparable enzymatic activity as the WT. Both mutants also showed a significant drug resistance against nirmatrelvir (>42-fold increase in K_i values). H172Y (13.9-fold lower k_{cat}/K_m value) and H172A (11.3-fold lower k_{cat}/K_m value) mutants had reduced enzymatic activity, while being resistant to nirmatrelvir (>113.7-fold increase in K_i). H172Y was similarly disclosed by Pfizer as a nirmatrelvir resistant mutant (233-fold increase in K_i) (*20*). The remaining H172T, H172E, H172D, H172L, H172M, H172I, H172V, H172S, H172N, and H172K had significantly reduced enzymatic activity (>21.0-fold lower k_{cat}/K_m values).

Q192 locates at the S4 pocket and forms hydrophobic interaction with the trifluoromethyl substitution from nirmatrelvir (Fig. 1B). Q192T (9.2-fold lower k_{cat}/K_m value), Q192S (8.9-fold), and Q192V (9.0-fold) had comparable enzymatic activity as the WT, and all three showed resistance against nirmatrelvir (>22.2-fold increase in K_i values). Cross resistance was also observed with PF-00835231 (>85.0-fold) and GC-376 (>7.7-fold).

The drug resistance of these 12 M^{pro} mutants against nirmatrelvir was further confirmed in the thermal shift drug titration assay. All mutants displayed a lower degree of protein stabilization than the WT M^{pro} with increasing concentrations of nirmatrelvir (Fig. 2C).

M49, T135, N142, H164, M165, and Q189 can tolerate multiple mutations without significantly affecting enzymatic activity and drug inhibition

The most frequent mutants at residue M49, M49I, M49T, M49L, and M49V remained sensitive to rimatrelvir (< 3-fold change in IC_{50} value). Interestingly, the enzymatic

activity (k_{cat}/K_m) of M49I and M49L mutants showed 1.69 and 1.74-fold increase compared to WT.

T135I is a high frequency mutation with 1325 occurrences. The T135I mutant had similar k_{cat}/K_m value as the WT and remained sensitive to all three inhibitors (<2.9-fold change in K_i values).

The top 9 high frequent mutants at residue N142 all had similar enzymatic activity as the WT (<4.1-fold change in k_{cat}/K_m value) and remained sensitive to rimatrelvir (<3.5-fold change in IC_{50} values).

H164N is a high frequency mutation with 4660 occurrences and remained sensitive to all three inhibitors (<4.1-fold change in K_i values). The catalytic activity of the H164N mutant (4.2-fold lower in k_{cat}/K_m value) is comparable to WT.

All 7 Q189 mutants retained similar enzymatic activity as the WT with the change in k_{cat}/K_m values between 1.9- and 9.2-fold. No significant drug resistance was observed for nirmatrelvir (<3.1-fold change in IC_{50} values). Interestingly, the enzymatic activity of Q189E increased by nearly 2-fold compared to WT.

Collectively, the results suggest that M49, T135, N142, H164, and Q189 residues might be able to accommodate multiple mutants without a significant compromise in enzymatic activity and drug sensitivity.

H172Y/Q189E double mutant rescued the enzymatic activity and maintained drug resistance

Given the enhanced enzymatic activity of the Q189E mutant (1.9-fold increase in k_{cat}/K_m value) and the reduced enzymatic activity of the H172Y mutant (13.9-fold decrease) compared to WT (Fig. 3A), we hypothesized that the H172Y/Q189E double mutant might rescue the reduced enzymatic activity of the drug resistant H172Y mutant. Indeed, the H172Y/Q189E double mutant increased both the enzymatic activity of H172Y from 790 $M^{-1}S^{-1}$ to 1,009 $M^{-1}S^{-1}$ (Fig. 3A) and resistance against nirmatrelvir (281.1-fold increase in K_i value in comparison to 146.3-fold increase of H172Y single mutation) (Fig. 3B). The resistance of the H172Y/Q189E against nirmatrelvir was further confirmed in the thermal shift binding assay (Fig. 3C). These results suggest that although many of the identified single mutant in this study had reduced enzymatic activity (table S1), it is plausible that the virus can generate multiple mutations to help restore the fitness of replication while maintaining or even enhanced drug resistance.

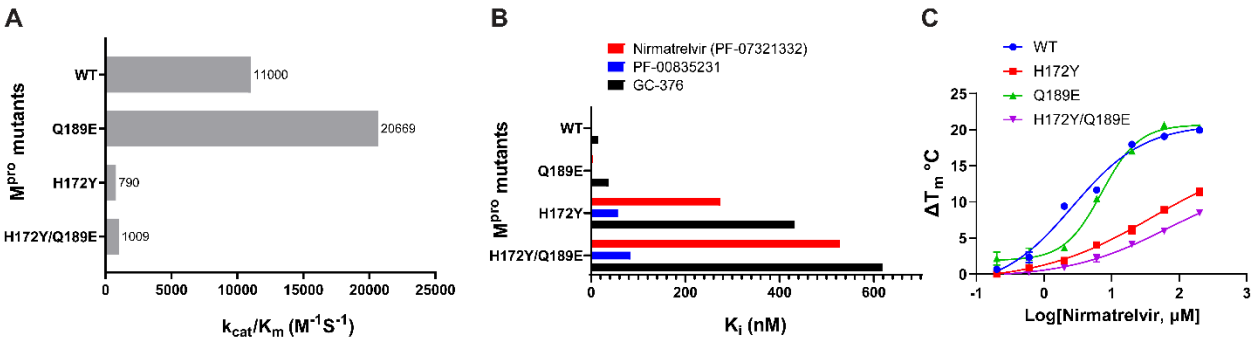

Fig. 3. Characterization of the enzymatic activity and drug inhibition of the SARS-CoV-2 M^{pro} H172Y/Q189E double mutant. (**A**) Plot of the k_{cat}/K_m values for the WT, Q189E, H172Y, and the H172Y/Q189E double mutant. (**B**) Plot of the K_i values for the WT, Q189E, H172Y, and the H172Y/Q189E double mutant. (**C**) Plot of the thermal shift

binding assay for the WT, Q189E, H172Y, and the H172Y/Q189E double mutant with nirmatrelvir.

Structural basis for resistance mutations

We have determined the X-ray crystal structures of several key representative mutants, including both unbound and GC-376 complex structures for S144A, S144L, H172Y, the unbound structure of E166Q, the GC-376 complex of Q192T, and the nirmatrelvir complex of H165Y, at 1.70-2.89 Å resolutions (Fig. 4, table S2). The unbound and GC-376 complex structures of H164N, which caused only a small decrease in enzymatic activity and inhibition, were also determined for comparison (fig. S3). GC-376 and nirmatrelvir have similar binding modes in the M^{pro} active site and place the same pyrrolidone group in the S1 pocket. Thus, the GC-376 complex structures can provide valuable information into how these mutants interact with inhibitors. It should be noted that the terminal benzene group of GC-376 exhibited two different binding modes in previously published WT structures (*18, 19*). The difference in the side chain between the GC-376 conformations in most of the mutant structures (except Q192T) and the reference WT structure likely originated from its flexibility, rather than the specific mutations.

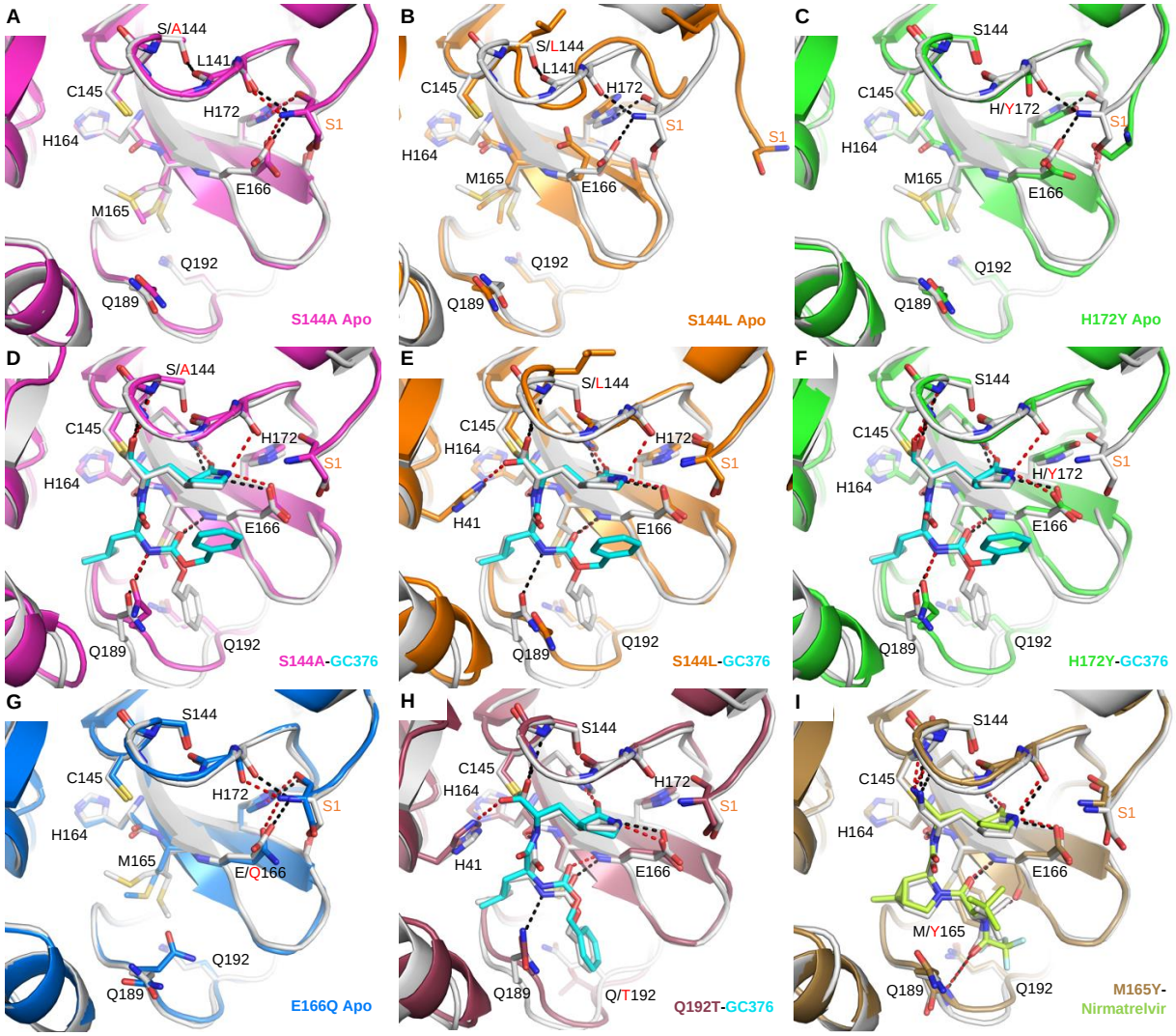

Fig. 4. X-ray crystal structures of M^{pro} mutants. Each mutant structure is aligned with the corresponding WT structure shown in white (apo, PDB 7JP1; GC-376 complex, 6WTT; nirmatrelvir complex, 7RFW). For the mutant structures, GC-376 and nirmatrelvir are shown in cyan and neon green respectively. WT HBs are shown as black dashes for selected residues at the mutation sites or between the protein and inhibitor. Mutant HBs are shown as red dashes. Mutations are indicated with red text. S1 residue from the N-terminus of the adjacent protomer is labeled in orange. The side chain of L141 is not shown. (**A**) Apo M^{pro} S144A (magenta, PDB 8D4L). (**B**) Apo M^{pro} S144L (orange, PDB

8DFE). The view for panel B is shifted slightly to show for the movement of the adjacent N-terminus. (**C**) Apo Mpro H172Y (green, PDB 8D4J). (**D**) M^{pro} S144A GC-376 complex (magenta, PDB 8D4M). (**E**) M^{pro} S144L GC-376 complex (orange, PDB 8DD9). (**F**) M^{pro} H172Y GC-376 complex (green, PDB 8D4K). S1 residue is disordered. (**G**) Apo M^{pro} E166Q (blue, PDB 8D4N). (**H**) M^{pro} Q192T GC376 complex (mauve, PDB 8DGB). (**I**) M^{pro} M165Y Nirmatrelvir complex (brown, PDB 8DCZ).

All the WT residues at the five mutation sites are involved in intra-molecular interactions and are at least partially buried. However, with the only exception of S144L, the changes caused by the six mutations are mostly local and relatively small. In addition, all mutant complex structures including S144L showed the inhibitor, GC-376 or nirmatrelvir, and the protein assumed conformations very similar to the WT, suggesting that the mutants adopt the productive configuration in the presence of the substrate or inhibitor, explaining their ability to catalyze the reaction.

The decrease in the mutants' catalytic activity and inhibition by GC-376/nirmatrelvir appears to stem from two causes, a largely enthalpic effect through direct disruption of ligand-binding interactions and an entropic effect through increasing conformational instability of the active site. S144L, M165Y, H172Y and Q192T represented some of the biggest changes in both the size of the residue side chain and the enzyme catalytic efficiency, with a decrease of ~127x, 39x, 13x and 10x respectively in k_{cat} values. Those mutations resulted in notable changes in ligand interactions in the structures. In the unbound state (Fig. 4B) and, to a lesser extent, the complex structure (Fig. 4E), the S144L mutation led to a drastically different conformation in the 140-146 loop as well as

neighboring regions to avoid steric clashes caused by the bulky leucine side chain. Importantly, this loop constitutes the oxyanion hole stabilizing the transition state of the enzymatic reaction, formed by the main chain amide groups of G143, S144, and C145. Interestingly, in the GC-376 complex structure of S144L (Fig. 4E) as well as Q192T (Fig. 4H), the thioacetal hydroxide is placed outside the oxyanion hole and hydrogen bonds with H41, unlike most previously determined structures (*18*). This suggests that the interactions between the oxyanion hole and the inhibitor thioacetal hydroxide as well as, by extension, the substrate transition state, are likely weakened in the S144L and Q192T mutants, resulting in an alternate conformation in the inhibitor crystal structure. Similarly, the M165Y mutation also seemed to lead to diminished interactions between the oxyanion hole and the ligand, as exemplified by the lengthened HB between the imine nitrogen and G143 amide NH (from 3.0 Å in the WT to 3.6 Å in the M165Y mutant) (Fig. 4I). These distortions are likely caused by the bulky Y165 residue through series of ripple effects relayed by both the protein and inhibitor in a tightly packed active site. Meanwhile, the H172Y mutation altered the conformation of the N-terminus of the adjacent protomer in the biological dimer, which is adjacent to the active site and was previously shown to be important to M^{pro} catalysis. In addition to these effects on the reaction center, the M165Y and Q192T mutation also directly impacted ligand interactions in the S4 site. The M165Y mutation pushes the terminal trifluoromethyl group of nirmatrelvir out of its original position in the WT complex, disrupting its interaction with Leu167 and the HB with Thr190O (Fig. 4I). The Q192T mutation also increased the plasticity of the surrounding residues (Fig. 4H), allowing them to better

accommodate the terminal benzene ring of GC-376, which assumed a different conformation compared with the other mutant complex structures.

In contrast to the above four mutations, the remaining two mutations, S144A and E166Q, represented the smallest changes in the side chain size, and led to nearly no alternations of the unbound structure compared with the WT (Figs. 4A, D, G). Nevertheless, these mutations have impaired the inhibitory activity of nirmatrelvir and GC-376. The mutations of these two residues, all located in the active site, abolish or reduce the intramolecular interactions involving their side chains in the WT, and subsequently increase the conformational instability in the protein active site. Even though the enthalpic interactions between the substrate/inhibitor and the protein may be similar to the WT in the lowest energy conformation, the entropic cost will be higher for the mutants, thus making the free energy of ligand binding less favorable. Because the inhibitor relies on better shape complementarity with a smaller portion of the active site and also contains more rigid features than the substrate, we hypothesize that this entropic cost may impact inhibitor binding more than the larger and more flexible substrate. For example, compared with the glutamine side chain at the P1 position, the pyrrolidone ring of nirmatrelvir and GC-376 forms additional interactions with the peptide bond between Leu141 and Asn142 through the two extra carbon atoms. The S144A mutation eliminates the HB with the carbonyl group of this peptide bond, likely increasing its flexibility and increasing the energetic cost of inhibitor binding. Such entropic effects are not limited to S144A and E166Q mutations causing minimal changes in the unbound protein structure, but also those aforementioned mutations

such as S144L and H165Y that can both directly influence the protein-ligand contacts and increase the active site flexibility.

SUMMARY

Collectively, our results have several implications. First, all 66 M^{pro} mutants characterized in this study are naturally occurring SARS-CoV-2 M^{pro} polymorphisms that could potentially affect the efficacy of Paxlovid, and continuous prescription of Paxlovid might likely increase the frequency of these pre-existing drug resistance mutants. Second, S144, M165, E166, H172, and Q192 appear to be hot spots for nirmatrelvir resistance and need to be closely monitored among circulating viruses. Mutations at these residues are most likely to maintain the enzymatic activity while cause a significant drug resistance. It remains to be further characterized whether these mutants impair the fitness of viral replication and transmission. Third, as exemplified the H172Y/Q189T mutant, double or triple mutants might emerge to compensate for the loss of enzymatic activity from the single mutant while maintaining or enhancing drug resistance. Therefore, the M^{pro} mutants with reduced enzymatic activity from this study should also be monitored.

REFERENCES AND NOTES

1. J. Hammond *et al.*, Oral Nirmatrelvir for High-Risk, Nonhospitalized Adults with Covid-19. *N. Engl. J. Med.* **386**, 1397-1408 (2022).
2. M. Xiong *et al.*, What coronavirus 3C-like protease tells us: From structure, substrate selectivity, to inhibitor design. *Med. Res. Rev.* **41**, 1965-1998 (2021).

3. A tale of two antiviral targets — and the COVID-19 drugs that bind them. https://www.nature.com/articles/d41573-021-00202-8.

4. Y. Kim *et al.*, Reversal of the Progression of Fatal Coronavirus Infection in Cats by a Broad-Spectrum Coronavirus Protease Inhibitor. *PLoS Pathog.* **12**, e1005531 (2016).

5. B. Boras *et al.*, Preclinical characterization of an intravenous coronavirus 3CL protease inhibitor for the potential treatment of COVID19. *Nat. Commun.* **12**, 6055 (2021).

6. D. R. Owen *et al.*, An oral SARS-CoV-2 M(pro) inhibitor clinical candidate for the treatment of COVID-19. *Science* **374**, 1586-1593 (2021).

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

8. M. D. Sacco *et al.*, 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).

9. K. Abe *et al.*, Pro108Ser mutation of SARS-CoV-2 3CLpro reduces the enzyme activity and ameliorates the clinical severity of COVID-19. *Sci. Rep.* **12**, 1299 (2022).

10. S. E. Greasley *et al.*, Structural basis for Nirmatrelvir in vitro efficacy against the Omicron variant of SARS-CoV-2. *J. Biol. Chem.* 101972 (2022).

11. A. N. Matthew *et al.*, Drug Design Strategies to Avoid Resistance in Direct-Acting Antivirals and Beyond. *Chem. Rev.* **121**, 3238-3270 (2021).

12. S. Mason, J. P. Devincenzo, S. Toovey, J. Z. Wu, R. J. Whitley, Comparison of antiviral resistance across acute and chronic viral infections. *Antiviral Res.* **158**, 103-112 (2018).

13. Y. Zhou *et al.*, Nirmatrelvir Resistant SARS-CoV-2 Variants with High Fitness in Vitro. *bioRxiv*, 2022.2006.2006.494921 (2022).

14. D. Jochmans *et al.*, The substitutions L50F, E166A and L167F in SARS-CoV-2 3CLpro are selected by a protease inhibitor in vitro and confer resistance to nirmatrelvir. *bioRxiv*, 2022.2006.2007.495116 (2022).

15. 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*. doi: 10.1021/acs.jmedchem.2c00404 (2022).

16. A. M. Shaqra *et al.*, Defining the substrate envelope of SARS-CoV-2 main protease to predict and avoid drug resistance. *Nat. Commun.* **13**, 3556 (2022).

17. www.gisaid.org/epiflu-applications/covsurver-mutations-app.

18. C. Ma *et al.*, Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease. *Cell Res.* **30**, 678-692 (2020).

19. M. D. Sacco *et al.*, Structure and inhibition of the SARS-CoV-2 main protease reveal strategy for developing dual inhibitors against M(pro) and cathepsin L. *Sci. Adv.* **6**, eabe0751 (2020).

20. FDA, Fact sheet for healthcare providers: Emergency use authorization for Paxlovid (2021); https://www.fda.gov/media/155050/download.

ACKNOWLEDGEMENTS

Funding: This work was supported by the National Institute of Allergy and Infectious Diseases of Health (NIH-NIAID) grants AI147325, AI157046, and AI158775 to J. W. We thank the scientists and staff at SER-CAT, especially Norma Duke, for their assistance with X-ray diffraction data collection. SER-CAT is supported by its member institutions, and equipment grants (S10_RR25528, S10_RR028976 and S10_OD027000) from the National Institutes of Health.

Author contributions: Conceptualization: J.W., Y.C., Y.H., H.T., E.M.L.; Protein expression: Y.H., H.T., X.Z.; Enzymatic assay and thermal shift assay: Y.H., H.T.; Protein crystallization: E.M.L., R.T.M., M.V.M, Y.C; Structure determination: E.M.L., L.M.C.J, S.G.B., Y.C.; Writing: J.W., Y.C.

Competing interests: The authors declare no competing interests.

Data and materials availability: The X-ray crystal structures have been deposited into the Protein Data Bank with accession codes 8D4J (H172Y apo), 8D4K (H172Y-GC376), 8D4L (S144A), 8D4M (S144A-GC376), 8D4N (E166Q apo), 8DFN (H164N apo), 8DD1 (H164N-GC376), 8DFE (S144L apo), 8DD9 (S144L-GC376), 8DGB (Q192T-GC376), and 8DCZ (M165Y-nirmatrelvir).

SUPPLEMENTARY MATERIALS

Materials and methods

Figs

Tables

References