

## RESEARCH ARTICLE

# A yeast-based system to study SARS-CoV-2 M<sup>PRO</sup> structure and to identify nirmatrelvir resistant mutations

Jin Ou<sup>1</sup>, Eric M. Lewandowski<sup>2</sup>, Yanmei Hu<sup>3</sup>, Austin A. Lipinski<sup>4</sup>, Ali Aljasser<sup>1</sup>, Mariliz Colon-Ascanio<sup>1</sup>, Ryan T. Morgan<sup>2</sup>, Lian M. C. Jacobs<sup>2,5</sup>, Xiujun Zhang<sup>2</sup>, Melissa J. Bikowitz<sup>2</sup>, Paul R. Langlais<sup>4</sup>, Haozhou Tan<sup>3</sup>, Jun Wang<sup>3\*</sup>, Yu Chen<sup>2\*</sup>, John S. Choy<sup>1\*</sup>

**1** Department of Biology, School of Arts and Sciences, The Catholic University of America, Washington, Washington DC, United States of America, **2** Department of Molecular Medicine, Morsani College of Medicine, University of South Florida, Tampa, Florida, United States of America, **3** Department of Medicinal Chemistry, Ernest Mario School of Pharmacy, Rutgers, the State University of New Jersey, Piscataway, New Jersey, United States of America, **4** Department of Medicine, College of Medicine, University of Arizona, Tucson, Arizona, United States of America, **5** Drug Discovery Department, Moffitt Cancer Center, Tampa, Florida, United States of America

\* [junwang@pharmacy.rutgers.edu](mailto:junwang@pharmacy.rutgers.edu) (JW); [ychen1@usf.edu](mailto:ychen1@usf.edu) (YC); [choy@cua.edu](mailto:choy@cua.edu) (JSC)



## OPEN ACCESS

**Citation:** Ou J, Lewandowski EM, Hu Y, Lipinski AA, Aljasser A, Colon-Ascanio M, et al. (2023) A yeast-based system to study SARS-CoV-2 M<sup>PRO</sup> structure and to identify nirmatrelvir resistant mutations. *PLoS Pathog* 19(8): e1011592. <https://doi.org/10.1371/journal.ppat.1011592>

**Editor:** Wenshe Liu, Texas A&M University, UNITED STATES

**Received:** November 21, 2022

**Accepted:** July 31, 2023

**Published:** August 31, 2023

**Copyright:** © 2023 Ou et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Proteomics data has been deposited to ProteomeXchange. Project accession: PXD036325 and Project DOI:[10.6019/PXD036325](https://doi.org/10.6019/PXD036325). The X-ray crystal structures have been deposited into the Protein Data Bank with accession codes 8DDI (E166N Apo) and 8DDM (E166R GC-376). All other relevant data are within the manuscript or in [Supporting Information](#) files.

**Funding:** This work was supported by the National Institutes of Health (NIH) grants 1R15GM129766-01 to JSC and AI158775 to JW, and a National

## Abstract

The SARS-CoV-2 main protease (M<sup>PRO</sup>) is a major therapeutic target. The M<sup>PRO</sup> inhibitor, nirmatrelvir, is the antiviral component of Paxlovid, an orally available treatment for COVID-19. As M<sup>PRO</sup> inhibitor use increases, drug resistant mutations will likely emerge. We have established a non-pathogenic system, in which yeast growth serves as an approximation for M<sup>PRO</sup> activity, enabling rapid identification of mutants with altered enzymatic activity and drug sensitivity. The E166 residue is known to be a potential hot spot for drug resistance and yeast assays identified substitutions which conferred strong nirmatrelvir resistance and others that compromised activity. On the other hand, N142A and the P132H mutation, carried by the Omicron variant, caused little to no change in drug response and activity. Standard enzymatic assays confirmed the yeast results. In turn, we solved the structures of M<sup>PRO</sup> E166R, and M<sup>PRO</sup> E166N, providing insights into how arginine may drive drug resistance while asparagine leads to reduced activity. The work presented here will help characterize novel resistant variants of M<sup>PRO</sup> that may arise as M<sup>PRO</sup> antivirals become more widely used.

## Author summary

The SARS-CoV-2 virus has proven to be highly adept at evading the newly developed vaccines. Antiviral drugs provide an important alternative that can reduce disease severity. Nirmatrelvir, an orally available drug approved by the FDA to treat COVID-19, inhibits the SARS-CoV-2 main or 3C-like protease (M<sup>PRO</sup> or 3CL<sup>PRO</sup>) and reduces the severity of COVID-19 infections. However, continued use of nirmatrelvir will likely drive the emergence of drug resistant M<sup>PRO</sup> mutations. Our study reports a rapid, inexpensive, and non-pathogenic system using baker's yeast that can determine if potential M<sup>PRO</sup> mutations may

Science Foundation CCF PIPP Grant 2200138 to JSC. A graduate scholarship from King Saud University, Saudi Arabia, supported AJ. Additional support was provided by funds from the Adeline E. Fracassi Trust (Gregory A. Pope) and the Litowitz Family Fund to JSC. SEC-CAT is supported by its member institutions, and equipment grants (S10\_RR25528, S10\_RR028976 and S10\_OD027000) from the National Institutes of Health. Use of the Advanced Photon Source was supported by the U. S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng-38. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors declare that there are no competing interests.

confer drug resistance. Furthermore, we solve the structures of two M<sup>Pro</sup> mutants first characterized in yeast that provide insights into how E166 can change to drive resistance versus leading to inactivity. Our results can aid in advancing our understanding of resistance mechanisms that will be important as nirmatrelvir and other protease inhibitors become more widely used.

## Introduction

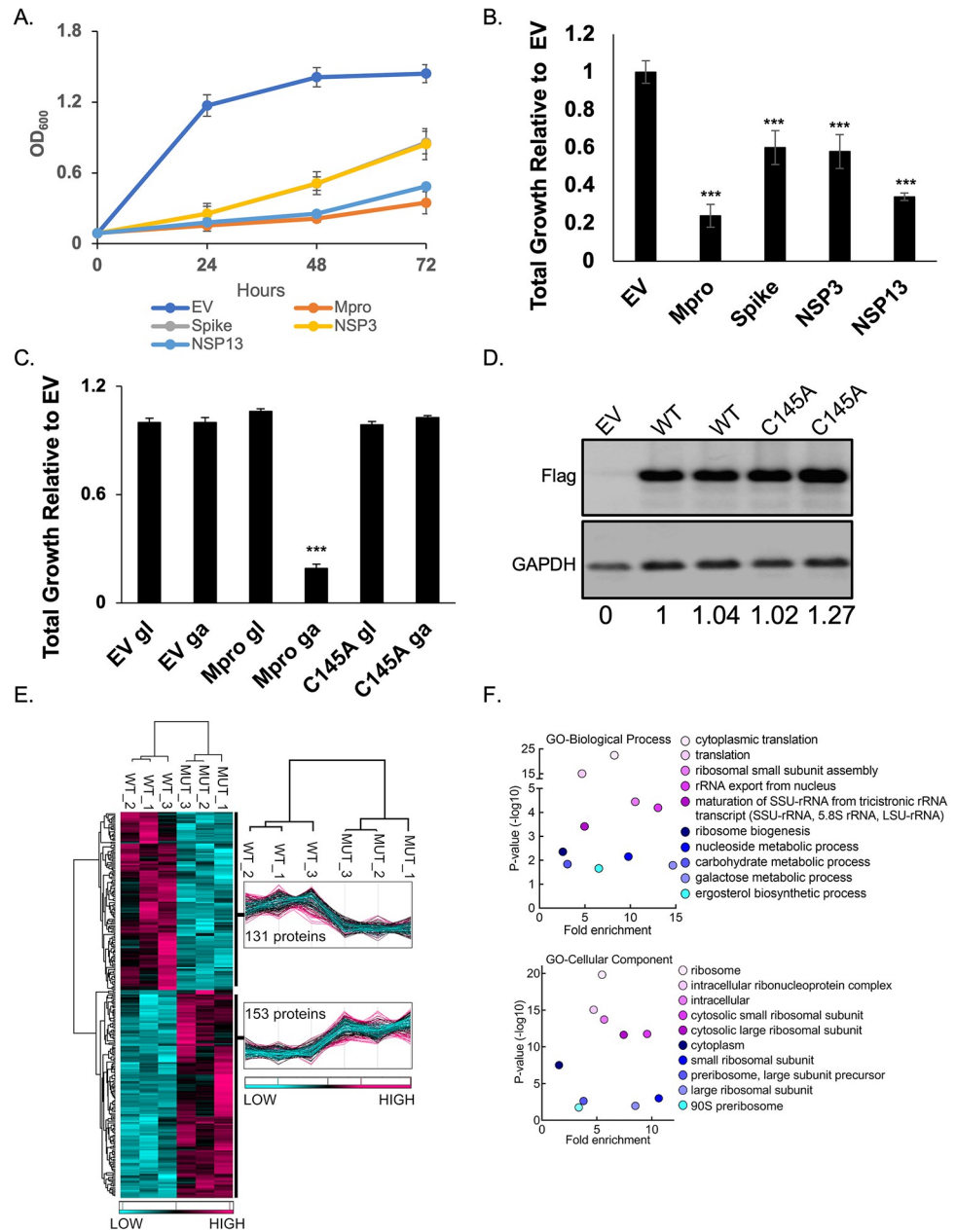
The evolution of new SARS-CoV-2 variants that evade vaccines, cause breakthrough COVID-19 infections in vaccinated individuals, and the limited vaccine availability in many parts of the world, highlight the need for complementary approaches [1]. Antiviral drugs provide an important alternative and can contribute to minimizing disease severity and death. The SARS-CoV-2 main or 3C-like protease (M<sup>Pro</sup> or 3CL<sup>Pro</sup>) is essential for viral replication and is a promising drug target [2,3]. There have been intense efforts to repurpose or to develop new drugs that directly target M<sup>Pro</sup> [4,5]. In December 2021, emergency authorization use of Paxlovid to treat COVID-19 was granted by the US Food and Drug Administration [6]. Paxlovid is a combination of the M<sup>Pro</sup> inhibitor, nirmatrelvir, and the cytochrome CYP3A inhibitor, ritonavir, which slows metabolism of nirmatrelvir [7,8]. There are several other M<sup>Pro</sup> inhibitors in clinical trials, including PF-07304814, the phosphate form of PF-00835231 [9,10], and ensitrelvir [11]. As M<sup>Pro</sup> inhibitors become more widely used the emergence of resistant mutations will increase as greater selection pressure is present in the population.

Knowledge of resistant mutants can inform on drug design modifications to identify new drugs that target resistant variants. However, standard approaches to characterize resistant mutants using live virus [12], recombinant proteins, and *in vitro* assays can be highly limiting due to infrastructure requirements, cost, and time [13]. Here we report a yeast system that is non-pathogenic, rapid, inexpensive, and reports on M<sup>Pro</sup> activity and drug resistance simply by measuring yeast growth. Using this assay, we found that compared to wild-type, the E166R mutation conferred strong nirmatrelvir resistance ( $K_i > 1000$ -fold). As the E166 site appears to be a hot spot for drug resistance from *in vitro* viral evolution experiments [14,15], we solved the structures of two substitution mutants M<sup>Pro</sup> E166N and M<sup>Pro</sup> E166R, revealing how E166 mutations may compromise activity versus drug resistance, respectively. Furthermore, we tested two known resistant mutants identified by *in vitro* SARS-CoV-2 evolution experiments [14,16,17] and find that they are resistant in the yeast assay. Our results demonstrate the yeast system can be a reliable tool to determine the activity and drug responses of M<sup>Pro</sup> mutants. Results from the yeast assays can help rapidly prioritize mutants for further analysis using more resource intensive systems. In doing, so we can efficiently test M<sup>Pro</sup> mutants as they arise in the population and aid in mitigating COVID-19 infections.

## Results

### SARS-CoV-2 M<sup>Pro</sup>, PL<sup>Pro</sup>, spike, and helicase proteins are toxic in *S. cerevisiae*

Six SARS-CoV-2 (Wuhan-Hu-1) NSPs and the structural genes were expressed from a high copy plasmid [18] to determine if any would result in growth effects (Figs 1A and S1A). We observed no marked growth phenotypes as determined by spot tests when M, E, N, NSP7, NSP8, or NSP12 were expressed (S1B Fig). In contrast, spot tests revealed nearly a complete absence of growth when cells expressed NSP3 (PL<sup>Pro</sup>), NSP5 (M<sup>Pro</sup> or 3CL<sup>Pro</sup>), NSP13



**Fig 1. M<sup>P<sup>ro</sup></sup> confers a significant reduction in growth in yeast caused by decreases in a variety of cellular proteins.**

A) The indicated SARS-CoV-2 genes regulated by a galactose inducible promoter were expressed in yeast and conferred growth defects compared to empty vector (EV). B) Bar graph shows the total growth of cultures after 72 hours normalized to EV. C) Galactose-induced (ga) expression of the catalytically inactive M<sup>P<sup>ro</sup></sup> C145A mutant (C145A ga) does not confer a growth reduction compared to WT (M<sup>P<sup>ro</sup></sup> ga). When grown in glucose (gl) all three strains grew equally well. D) Protein levels of the M<sup>P<sup>ro</sup></sup> C145A mutant and wild-type M<sup>P<sup>ro</sup></sup> (WT) are comparable. Shown are two biological replicates for each form of M<sup>P<sup>ro</sup></sup>. Ratios of Flag:GAPDH signals relative to M<sup>P<sup>ro</sup></sup> WT is shown at the bottom of each lane. E) Total protein lysates made from yeast expressing the wild-type M<sup>P<sup>ro</sup></sup> (WT) or M<sup>P<sup>ro</sup></sup> C145A mutant (MUT) were subjected to mass spectrometric analyses revealing 153 proteins were higher in abundance in the mutant relative to the wild-type. F) Gene Ontology (GO) analyses indicates an enrichment of proteins with functions in translation that are significantly reduced in the presence of M<sup>P<sup>ro</sup></sup> versus M<sup>P<sup>ro</sup></sup> C145A. Plots in A, B, C show averages from three biological replicates and error bars are standard deviations. (\*\*\*) indicates differences (p<0.001) between EV and tested genes.

<https://doi.org/10.1371/journal.ppat.1011592.g001>

(Helicase), and spike (S1B Fig). Analysis of growth profiles of cells expressing PL<sup>Pro</sup>, M<sup>Pro</sup>, Helicase, and spike showed all four genes caused a reduction in growth. Previous reports have shown that expression of the SARS-CoV-1 PL<sup>pro</sup> reduces yeast growth [19]. M<sup>Pro</sup> and the Helicase were the most toxic conferring a ~70 to 80% reduction in total growth by 72 hours compared to cells carrying empty vector (Fig 1A and 1B). As M<sup>Pro</sup> is highly conserved between classes of coronavirus and a key drug target we focused our efforts on using the yeast system to study M<sup>Pro</sup> structure and function.

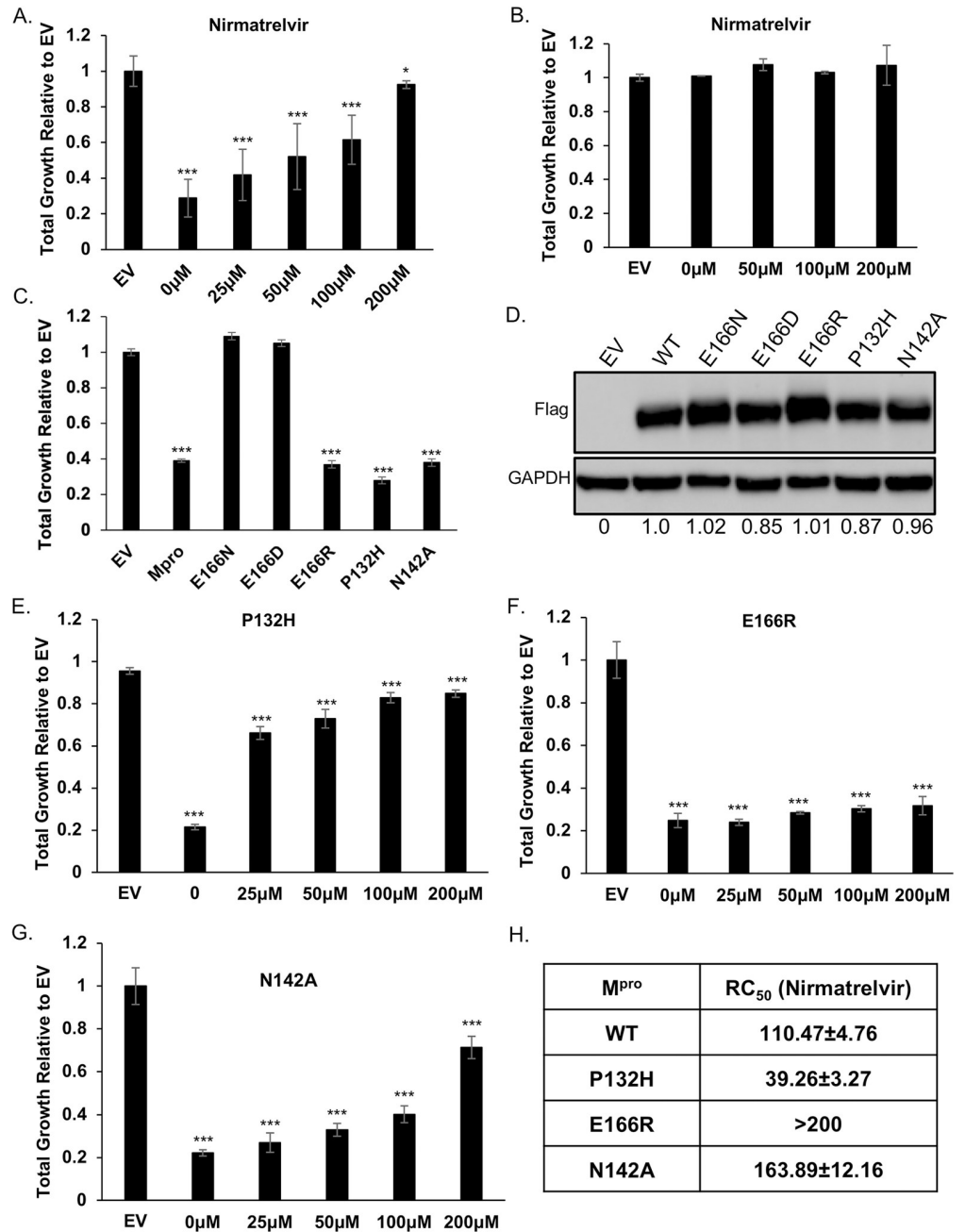
### Growth defect conferred by M<sup>Pro</sup> expression depends on its catalytic activity and associated with decreased abundance in essential and non-essential yeast proteins

To determine if the growth reduction depended on M<sup>Pro</sup> proteolytic activity we constructed a catalytic mutant of M<sup>Pro</sup> by replacing the key cysteine at position 145 to an alanine, which prevents the initial protonation step needed for peptide bond hydrolysis [20,21]. Liquid growth assays showed that yeast expressing the M<sup>Pro</sup> C145A mutant grew as well as the yeast control carrying empty vector (Fig 1C). In contrast, all three strains grew similarly in glucose, which represses expression of M<sup>Pro</sup> (Fig 1C). Western analysis showed that yeast expressed similar levels of wild-type and M<sup>Pro</sup> C145A mutant (Fig 1D). These results demonstrate that the growth reduction observed in yeast expressing M<sup>Pro</sup> is dependent on its proteolytic activity.

Next, we measured the relative abundance of proteins in yeast expressing M<sup>Pro</sup> compared to yeast expressing the M<sup>Pro</sup> C145A catalytic mutant to determine the mechanism(s) that lead to loss of cell viability. Whole cell lysates were made from three independent cultures of cells expressing wild-type M<sup>Pro</sup> or the catalytic M<sup>Pro</sup> C145A mutant (Figs 1E and S2). The biological replicates were highly reproducible, and we observed peptides from 153 proteins (S1 Table) were significantly reduced in yeast expressing M<sup>Pro</sup> compared to the M<sup>Pro</sup> C145A mutant (Fig 1E.) Gene ontology analysis revealed an enrichment for genes with functions in translation (Fig 1F). In particular, multiple ribosomal proteins and translational regulators were reduced. There were a number of proteins that were significantly enriched in the M<sup>Pro</sup> catalytic mutant with functions in a variety of activities beyond translation (S1 Table) and several are known to be essential (S1 Table). Furthermore, approximately 16% (25/153) of reduced proteins contained at least one potential canonical M<sup>Pro</sup> site (LQ/S, G, A) (S2 Table) suggesting that some may be direct substrates of M<sup>Pro</sup> [22,23]. Another 36% (55/153) of reduced proteins carried non-canonical sites in which LQ was followed by any amino acid besides S, G, or A. These results show that expression of M<sup>Pro</sup> leads to decreases in a variety of proteins and eventual loss of translation that is likely the cause of the growth defects.

### Nirmatrelvir restores growth to yeast expressing M<sup>Pro</sup> from multiple coronaviruses

Considering that the growth reduction conferred by M<sup>Pro</sup> activity is dependent on its proteolytic activity we tested if treating yeast with nirmatrelvir, would suppress the growth reduction. We tested nirmatrelvir at several concentrations and observed no cytotoxic effects (Fig 2B). Treating cells with increasing doses of nirmatrelvir led to a corresponding increase in growth (Fig 2A). At 100 $\mu$ M and 200 $\mu$ M of nirmatrelvir, growth was restored to similar levels as cells carrying empty vector (Fig 2A). As a qualitative measure to compare the effects of nirmatrelvir, we estimated the concentration of drug required to restore 50% of growth (RC<sub>50</sub>) relative to that of untreated M<sup>Pro</sup> expressing cells. Based on this criterion we calculated RC<sub>50</sub> for nirmatrelvir to be 110.47  $\pm$  4.76 $\mu$ M (Fig 2A and 2H). To determine if M<sup>Pro</sup> from other coronaviruses could be studied similarly, we tested the recent Omicron variant, M<sup>Pro</sup> P132H, which is



**Fig 2. Yeast growth assays identify nirmatrelvir resistant M<sup>Pro</sup> mutants.** A) Total growth of cultures after 72 hours expressing M<sup>Pro</sup> in the presence of increasing doses of nirmatrelvir normalized to growth of yeast carrying empty vector (EV) are plotted. Growth is restored by nirmatrelvir in a dose dependent manner. B) No growth effects are observed in cells treated up to 200μM nirmatrelvir. C) Yeast expressing substitutions E166D and E166N grow as well as EV but E166R, P132H, and N142A results in significant growth reduction comparable to wild-type M<sup>Pro</sup>. D) Western analysis shows that mutants and wild-type M<sup>Pro</sup> are expressed at comparable levels. Ratios of Flag:GAPDH signals relative to M<sup>Pro</sup> WT is shown at the bottom of each lane. E—G) Cells expressing P132H and N142A remain sensitive to nirmatrelvir, indicated by growth recovery, but E166R appears to be resistant as there is a lack of growth even when treated with 200μM of nirmatrelvir. H) RC<sub>50</sub> measurements of each mutant in response to nirmatrelvir treatment. For all experiments, at least three biological and three technical replicates were performed. Error bars represent standard deviations. (\*, p<0.01; \*\*\*, p<0.001) indicates differences compared to EV.

<https://doi.org/10.1371/journal.ppat.1011592.g002>



currently the dominant form of M<sup>Pro</sup>, and M<sup>Pro</sup> from SARS-CoV-1 and Bat-CoV-HKU9. We observed that in all cases M<sup>Pro</sup> conferred a significant growth reduction (Figs 2C and S3). Nirmatrelvir has been reported to have broad M<sup>Pro</sup> specificity [7,9]. Consistent with this work, we observed that nirmatrelvir could restore growth in yeast expressing M<sup>Pro</sup> from all three forms of M<sup>Pro</sup> (Figs 2E, 2H, and S3).

### Characterization of potential nirmatrelvir resistant mutations in M<sup>Pro</sup>

We tested if growth of yeast expressing M<sup>Pro</sup> could be used as an approximation for M<sup>Pro</sup> activity. Thus, providing a system to rapidly determine the effects of mutations on M<sup>Pro</sup> activity and drug resistance. A variety of interactions (H-bonds, salt-bridges, van der Waals) mediate binding between the catalytic site of M<sup>Pro</sup> and inhibitors [24–26]. While knowledge of the residues in contact with the inhibitor can inform predictions that may compromise inhibitor binding it is not obvious what amino acid substitutions would maintain M<sup>Pro</sup> activity toward substrate while compromising inhibitor interactions. With our yeast system we can easily test the effect of substitution mutations and rapidly determine if the mutations alter catalytic activity and sensitivity to inhibitor(s) by following growth phenotypes. To determine the feasibility of this approach we focused on E166, and N142 as these two residues form direct interactions with inhibitors and substrates [24,27].

We tested substitutions of E166 with three different amino acids that are yet to be dominant in the population. The following mutants predicted to be conserved (E166D), as the negative charge is maintained but with one less carbon in the side-chain; non-conserved (E166N), as asparagine is uncharged and has one less side chain carbon; and another non-conserved (E166R) substitution in which the arginine side chain is longer and positively charged.

We observed that all three substitution mutants were expressed at levels comparable to wild-type M<sup>Pro</sup> (Fig 2C). Expression of M<sup>Pro</sup> E166D and M<sup>Pro</sup> E166N did not cause a reduction in growth and grew as well as empty vector controls (Figs 2C, 2D and S4A). These results indicate that M<sup>Pro</sup> E166D and M<sup>Pro</sup> E166N may have defects in their enzymatic activities. However, the M<sup>Pro</sup> E166R mutant conferred a growth reduction that matched the wild-type M<sup>Pro</sup>, suggesting that its catalytic activity was intact (Figs 2C and S4B). While this manuscript was in preparation it was reported from *in vitro* SARS-CoV-2 evolution experiments that E166A and E166V substitutions conferred nirmatrelvir resistance [14,16,17]. Thus, we tested these two mutants in yeast and observed that both mutants were expressed at comparable levels and conferred a growth reduction similar to wild-type M<sup>Pro</sup> (S5A and S5B Fig).

Next, we challenged cells expressing M<sup>Pro</sup> E166R, E166A, and E166V with increasing concentrations of nirmatrelvir (25 μM, 50 μM, 100 μM, or 200 μM) and observed no significant improvement in growth remaining nearly identical to the untreated cultures (Figs 2F, S4B, S5C and S5D). Based on these experiments, the RC<sub>50</sub> for nirmatrelvir is >200 μM compared to wild-type M<sup>Pro</sup> (Figs 2H, S5C and S5D). These results suggest that the E166R, E166A, and E166V mutations confer resistance to nirmatrelvir.

We constructed a substitution at position N142, which is known to contribute to inhibitor and substrate binding [7] and is yet to be present in the population. To inform on the specific substitution to make we used a distantly related M<sup>Pro</sup> from the gamma-coronavirus, IBV, which is conserved but displays slight divergence from SARS-CoV-2 M<sup>Pro</sup> [9]. We replaced N142 with alanine (M<sup>Pro</sup> N142A), as alanine is found in the IBV M<sup>Pro</sup> at the homologous site [28]. We observed M<sup>Pro</sup> N142A was expressed at levels comparable to wild-type and conferred a similar reduction in growth (Fig 2C and S4C) showing that it remained active. The RC<sub>50</sub> for nirmatrelvir increased modestly by ~1.5-fold (Fig 2G and 2H). These results show that the substitution mutants E166A, E166V, and E166R lead to nirmatrelvir resistance while N142A

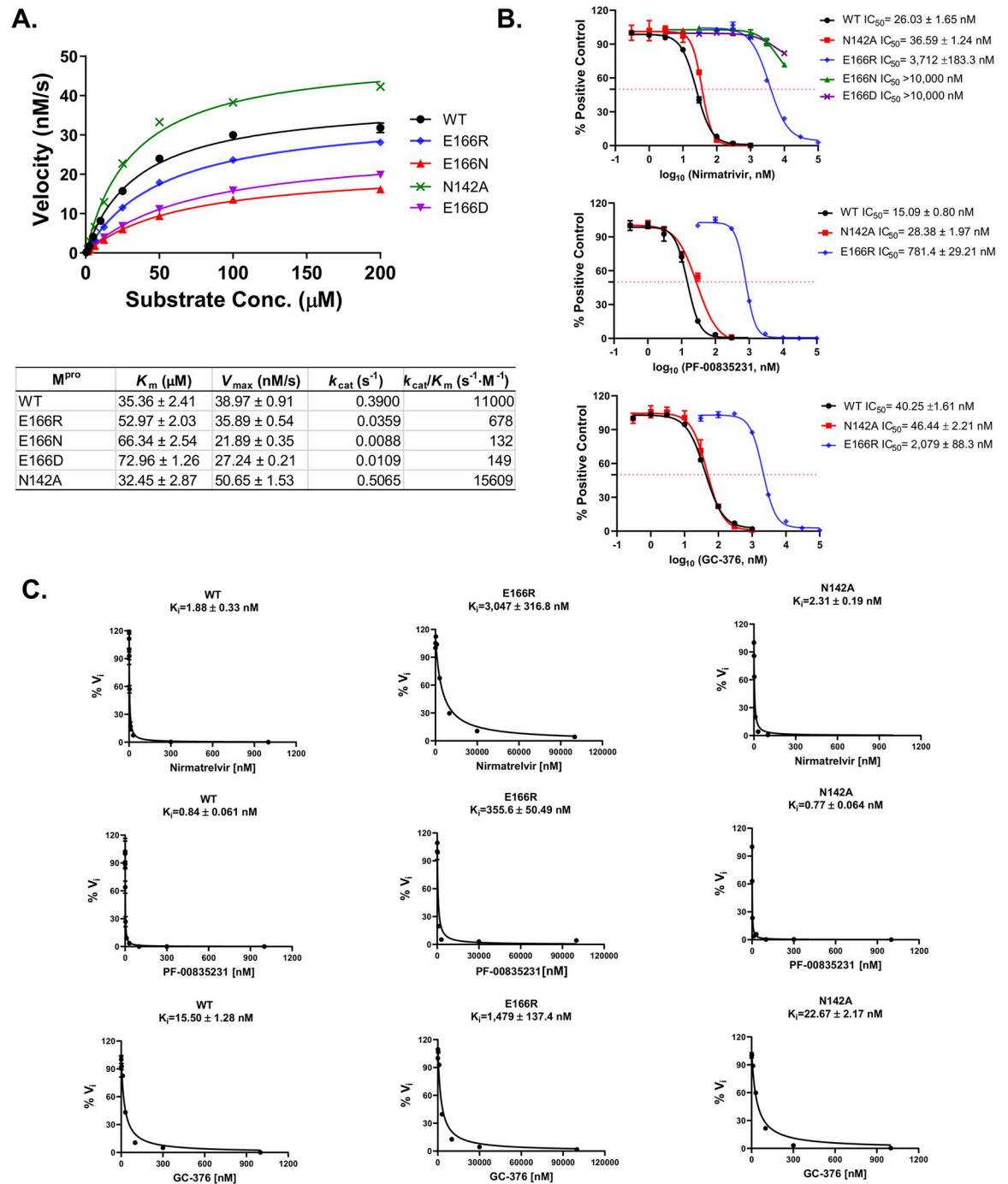
results in little difference from wild-type and E166N and E166D cause a significant loss in activity.

### ***In vitro* protease assays confirm that M<sup>Pro</sup> E166R, E166A, E166V are highly resistant to nirmatrelvir**

To determine how well yeast growth assays correlated with standard enzymatic assays we directly measured proteolytic activity using recombinant M<sup>Pro</sup> WT, E166N, E166D, E166R, E166A, E166V, and N142A. Compared to wild-type M<sup>Pro</sup>, the catalytic efficiencies ( $k_{cat}/K_m$ ) of M<sup>Pro</sup> E166R, E166A, and E166V were decreased by ~16-fold, ~8-fold, and ~10-fold, respectively (Figs 3A and S6A). In contrast, E166N and E166D displayed severe reductions in catalytic efficiencies of ~84- and ~74-fold compared to WT, respectively (Fig 3A). On the other hand, M<sup>Pro</sup> N142A displayed a slight increase in catalytic efficiency of ~1.4-fold compared to WT (Fig 3A). To determine the response of the mutants to inhibitors we performed IC<sub>50</sub> and K<sub>i</sub> measurements. We observed for M<sup>Pro</sup> E166R, E166A, and E166V, increases in IC<sub>50</sub>'s of ~143-, ~6-, >300-fold for nirmatrelvir compared to WT, respectively (Figs 3B, 3C, S6B and S6C). On the other hand, M<sup>Pro</sup> N142A, only minor increases were observed in IC<sub>50</sub>'s of ~1.4-fold for nirmatrelvir (Fig 3B). The K<sub>i</sub> values for the inhibitors in assays with M<sup>Pro</sup> E166R, E166A, and E166V were increased by >1600-, >47-, and >5000-fold for nirmatrelvir, ~423-, ~27-, ~790-fold for PF-0085231, and ~37, ~9-, ~38-fold for GC-376, respectively. (Figs 3C and S6C). Nearly no difference in K<sub>i</sub> values was observed between wild-type M<sup>Pro</sup> and M<sup>Pro</sup> N142A, ~1.2-fold for nirmatrelvir, ~0.9-fold for PF-0085231, ~1.5-fold for GC-376) (Fig 3C). The enzymatic assays confirm the results from the yeast assays showing that M<sup>Pro</sup> E166R, E166A, and E166V are resistant to nirmatrelvir and also show that there is cross-resistance to PF-0085231 and GC-376 (Figs 3B, 3C and S6C). Similarly, results from yeast assays of M<sup>Pro</sup> N142A mutant appears to correspond well to the *in vitro* assays as both show minor to no increases in resistance (Fig 3B and 3C). Furthermore, the M<sup>Pro</sup> E166N and E166D mutants, which is not predicted to be catalytically active from the yeast assay, displayed >70-fold decreases in activity compared to wild-type in the *in vitro* assays. This result is completely consistent with observing no growth reduction when expressed in yeast. Taken together there is good correlation between the enzyme and yeast assays.

### **Crystal structure of M<sup>Pro</sup> E166R reveals a loss of interactions leading to drug resistance**

We were particularly interested in how replacing glutamate at position 166 with arginine led to a >1000-fold increase in resistance to nirmatrelvir while a substitution with asparagine led to an 83.5-fold decrease in enzymatic activity even though E166 is not known to be directly involved in catalysis. Toward addressing both questions, we solved the crystal structure of apo M<sup>Pro</sup> E166N and the complex structure of M<sup>Pro</sup> E166R with GC-376 at 2.3 and 2.1 Å resolution, respectively (Fig 4). Both proteins were crystallized in the C2 space group with one M<sup>Pro</sup> molecule per asymmetric unit, and the biological dimer can be generated through crystallographic symmetry. The conformations of the protein and ligand are therefore exactly the same between the two protomers of the dimer in these structures. Our efforts to obtain nirmatrelvir complex failed due to the relatively low compound solubility and the reduced binding for the mutant. But both nirmatrelvir and GC-376 have the same pyrrolidone side chain placed in the S1 site where E166 is located. In previous WT complex structures (PDB codes, 7RFW (nirmatrelvir), 6WTT(GC-376)), the protein and the pyrrolidone side chain adopted identical conformations in this area. The GC-376 complex can thus offer important insights into how E166 mutations may impact the binding of ligands with similar chemical structures, including nirmatrelvir.

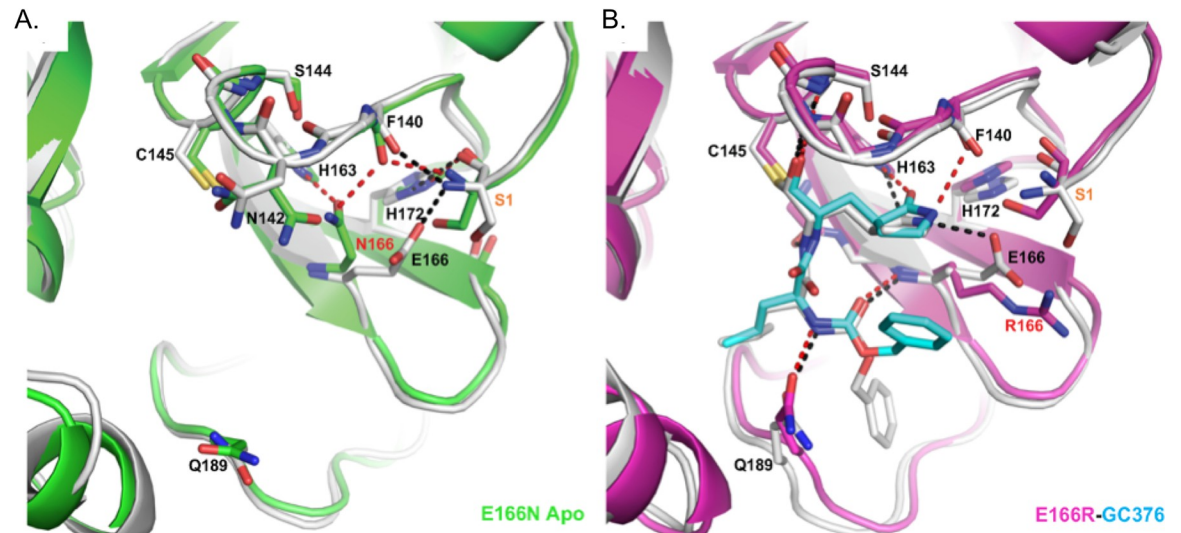


**Fig 3. Enzymatic assays demonstrate that E166N and E166D have severe defects in catalytic activity and M<sup>Pro</sup> E166R is highly resistant to nirmatrelvir.** A) Michaelis–Menten plot of M<sup>Pro</sup> and its mutants with various concentrations of FRET substrate. The K<sub>m</sub>, V<sub>max</sub>, k<sub>cat</sub>, and k<sub>cat</sub>/K<sub>m</sub> values are shown in the table. B) The IC<sub>50</sub> plots of nirmatrelvir, PF-00835231, and GC-376 against M<sup>Pro</sup> WT, N142A, E166R, E166N, and E166D. C) K<sub>i</sub> plots of nirmatrelvir, GC-376, and PF-00835231 against M<sup>Pro</sup>, M<sup>Pro</sup> E166R, and M<sup>Pro</sup> N142A.

<https://doi.org/10.1371/journal.ppat.1011592.g003>

In the M<sup>Pro</sup> E166N mutant structure, N166 forms a hydrogen bond (HB) with H163, an interaction not observed between E166 and H163 in the wild-type M<sup>Pro</sup> structure (Fig 4A). This new HB prevents H163 from hydrogen bonding with the glutamine side chain of the substrate, an interaction crucial to substrate binding. The binding of the substrate would therefore





**Fig 4. Crystal structures reveals structural basis for E166R resistance and E166N inactivity.** A) Apo M<sup>PRO</sup> WT (white, PDB 7JP1) aligned with apo M<sup>PRO</sup> E166N (green, PDB 8DDI). B) M<sup>PRO</sup> WT GC-376 complex (white, PDB 6WTT) aligned with M<sup>PRO</sup> E166R GC-376 complex (magenta, PDB 8DDM). WT hydrogen bonds are shown as black dashes, and mutant hydrogen bonds are shown as red dashes. GC-376 is shown in white for the WT structure and cyan for the mutant structure. Mutations are indicated with red text. Ser1 from an adjacent protomer is indicated with orange text.

<https://doi.org/10.1371/journal.ppat.1011592.g004>

require N166 to adopt a different conformation, breaking the HB with H163 and increasing the energetic cost. These observations are consistent with the drastic decrease of activity in the E166N mutant and lack of toxicity when expressed in yeast (Fig 2C) bringing to light how residues outside of the catalytic core can influence substrate binding.

In contrast, the longer and positively charged R166 side chain in the M<sup>PRO</sup> E166R mutant does not interact with H163, but rather extends into the solvent (Fig 4B). Therefore, the S1 site is open for substrate binding. However, the E166R mutation does affect ligand binding in several aspects. The negatively charged E166 side chain forms two crucial HBs, one with the N-terminus of the neighboring M<sup>PRO</sup> protomer in the biological dimer, and the other with the pyrrolidone side chain of inhibitors (in both nirmatrelvir and GC-376) or with the glutamine side chain of the substrate as described above. The E166R mutation would abolish this direct HB with the substrate or inhibitor, resulting in the pyrrolidone ring of GC-376 forming an alternative weak HB with F140 backbone carbonyl group (3.1 Å in length) in the mutant complex structure (Fig 4B). In addition, the N-terminus of the enzyme interacts with both E166 and the backbone carbonyl group of F140, and plays an important role in maintaining the structural stability of the enzyme active site. The E166R mutation eliminates the salt bridge with the N-terminus of the adjacent protomer, and further introduces electrostatic repulsion leading to small yet significant changes in the N-terminus conformation. Consequently, the distance between the N-terminal amine group and the F140 carbonyl group increased from 2.6 Å in the WT to 3.7 Å in the M<sup>PRO</sup> E166R mutant, diminishing the HB. This in turn may destabilize the loop that F140 resides on and also contains other important structural features involved in enzyme catalysis and ligand binding, including the backbone amide groups of Gly143 and Ser144 that form part of the oxyanion hole to stabilize the reaction transition state. This loop also contains the peptide bond between Leu141 and Asn142 that interacts with the two extra carbon atoms of the inhibitor pyrrolidone ring, but not with the substrate glutamine side chain. Destabilization of the region near F140 may increase the entropic cost of binding to the rigid pyrrolidone ring of nirmatrelvir and GC-376, more than the smaller and more

flexible substrate glutamine side chain. For similar entropic reasons, the HB between the pyrrolidone ring and E166 might contribute more to inhibitor binding than that between the more flexible glutamine side chain and E166 (Fig 4B). Consequently, the E166R mutation may have a stronger effect on binding to inhibitors such as nirmatrelvir versus substrate.

## Discussion

In sum, we demonstrate that using yeast growth as an approximation for M<sup>Pro</sup> activity can be a reliable indicator of the effects that mutations in M<sup>Pro</sup> can have on its activity and potential for drug resistance. Yeast assays indicated that E166 substituted with R, V, and A were all resistant to nirmatrelvir and *in vitro* enzyme assays confirmed this observation, revealing a ~47 to ~5000-fold increase in  $K_i$ . Furthermore, the C145A catalytic mutant and E166N and E166D mutants did not cause a growth reduction in yeast and enzyme assays showed that the E166D or N substitution confers a dramatic ~74- or ~84-fold decrease in activity, respectively. In yeast assays the N142A mutant displayed minor differences in drug sensitivity compared to wild-type ( $RC_{50}$  ~1.5-fold more than WT), which was confirmed by our *in vitro* enzyme assays. Similarly, the P132H mutant remained sensitive to nirmatrelvir based on our yeast assay, potentially even more sensitive with an  $RC_{50}$  ~2.8-fold less than WT. This is consistent with previous reports showing that the P132H mutant remains sensitive to nirmatrelvir in *in vitro* enzyme assays [29–32]. It appears that M<sup>Pro</sup> mutants (i.e. E166R) that have a decrease in catalytic efficiencies of up to 16-fold compared to WT are still able to confer a marked reduction in yeast growth. This is important as resistant mutants are likely to reduce protein fitness [33,34]. However, the yeast assay is unable to detect enhanced M<sup>Pro</sup> activity (e.g., M<sup>Pro</sup> N142A), which we observed in *in vitro* assays. This may have been due to the relatively small increase (1.4-fold). However, the enhanced activity associated with N142A suggests that M<sup>Pro</sup> can evolve to be a more active enzyme. It is possible that mutants which enhance M<sup>Pro</sup> activity can improve protein fitness when combined with resistant mutants that on their own may have reduced activity [15]. Recent reports from *in vitro* evolution experiments suggest that single mutants such as E166A and E166V, which do provide resistance, have significant reductions in proteolytic activity and require compensatory mutations to improve fitness levels necessary for infection at least *in vitro* [14,16,17]. While the yeast system does not identify the necessary compensatory mutation(s), it nonetheless can point to mutations that confer resistance and to prioritize mutants that are of interest for further studies. The crystal structure of E166R with GC-376 revealed loss of key hydrogen bonds with the pyrrolidone ring of GC-376 which can explain the increase in resistance to nirmatrelvir containing the same functional group. The E166A and E166V mutations would have a similar effect by abolishing the hydrogen bonds involving the E166 side chain, therefore reducing direct contacts with the substrate/inhibitor and destabilizing the active site. On the other hand, the E166N mutant which could be considered a more conserved change than E166R decreased activity by ~83-fold and did not confer a growth reduction in the yeast assays. In turn the crystal structure shows that the asparagine prevents substrate binding through a new hydrogen bond with H163, providing a mechanism to explain the significant reduction in activity. The additional mutants at E166 that are associated with *in vitro* viral evolution experiments along with what we show here highlight the importance of this site in playing a role in nirmatrelvir resistance. Our crystal structure illuminates a structural mechanism to help explain how substitutions at E166 can either lead to loss of activity versus gain of resistance.

While the drug doses used with yeast are in the micromolar versus nanomolar range that is more typical of *in vitro* enzymatic or viral assays, we observed good correlations between the yeast and enzymatic assays for nearly all of the mutants tested. The higher concentrations of

drug may be needed even though we deleted the major efflux pump, Pdr5, as yeast harbor a range of efflux activities [35], or possibly differences in permeability as a result of lipid composition differences from human cells, as well as potential drug interactions with the yeast cell wall [36]. Additional differences observed between the yeast and enzymatic assays may be a result of having multiple substrates in yeast, additional complexity of the cellular proteome, differences in pH, salt, and oxidation levels. Analysis of the top 153 proteins ( $p < 0.05$ ) with reduced abundance relative to strains expressing M<sup>PRO</sup> C145A revealed that 16% (25) have at least one M<sup>PRO</sup> site (LQ/S, LQ/A, or LQ/G) (S2 Table) suggesting that these may be direct M<sup>PRO</sup> substrates [22,23]. There were another 36% (55) of proteins that have non-canonical M<sup>PRO</sup> sites in which LQ is followed by any amino acid besides S, A, or G. A few reports identifying human host substrates suggest that M<sup>PRO</sup> might recognize such sites [37,38]. Thus, the total possible substrates may be greater than 16% of the total proteins we found to be reduced when M<sup>PRO</sup> is present. However, it is likely that these 36% of proteins are reduced due to a secondary consequence of the primary yeast substrates with canonical sites. We acknowledge that this may also suggest that drug resistance phenotypes in yeast may in some cases not directly indicate resistance relative to the viral cleavage sites. However, finding mutants (e.g. E166A and E166V) known to be compatible with virus replication and drug resistance to also display resistance in yeast suggests that it can be a good proxy for mutants of interest.

Taken together, these results demonstrate that a non-pathogenic, rapid, inexpensive and highly accessible yeast-based method can be used to characterize mutants for both their effects on M<sup>PRO</sup> activity and their responses to inhibitor compounds. There are reports using yeast as a tool to screen for M<sup>PRO</sup> inhibitors or that use deep mutational scanning of M<sup>PRO</sup> to identify high and low tolerant sites [16,39,40]. Unlike these previous reports we show that nirmatrelvir resistant mutants can be tested directly in yeast using growth as an approximation for M<sup>PRO</sup> drug resistance. The qualitative results from the yeast assays can be an important tool to help prioritize mutants of interest before moving ahead to more demanding viral based experiments. As more inhibitors are used in the general population there will be increasing selection pressures for drug resistant mutations that will go beyond the current set of mutants that are potentially drug resistant [41,42]. The yeast system reported here promises to be an invaluable tool in helping to combat future drug resistant mutations to stem the tide of COVID-19 infections.

## Materials and methods

### Strains, media, and chemicals

All yeast strains carried a *pdr5::G418* deletion in the BY4741 background (*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*). Yeast were grown in liquid synthetic complete (SC) media (0.17% yeast nitrogen base, 0.5% ammonium sulfate, amino acid mix with appropriate drop out as noted, 2% glucose) or on solid SC media containing 2% agar at 30°C. Media and reagents for culturing yeast were from United States Biological (Salem, MA). M<sup>PRO</sup> and PL<sup>PRO</sup> inhibitors were from MedChemExpress (Monmouth Junction, NJ) and Selleck Chemicals (Houston, TX). All other chemicals were from Sigma Aldrich (St. Louis, MO) or VWR (Radnor, PA).

### Expression of SARS-CoV-2 genes in yeast and mutagenesis

The indicated SARS-CoV-2 genes were codon optimized for yeast, tagged at the 3' with a 3X-Flag epitope, carried on high copy plasmids and genes were under the control of the Gal1 promoter (see S3 Table). Site directed mutagenesis was performed using In-Fusion Cloning Kit (Takara). Primers used for mutagenesis can be found in S2 Table. The M<sup>PRO</sup> gene was sequenced to confirm that mutations were incorporated successfully.

### Yeast transformation

A single yeast colony was used to inoculate 5ml liquid YPD (1% yeast extract, 1% yeast bacto-peptone, 2% glucose) and grown overnight at 30°C. The next day cells were washed and resuspended in 1ml lithium acetate/TE solution (100 mM lithium acetate, 10 mM Tris-HCl, 1 mM EDTA, pH 7.5). Cells were aliquoted (60 µl) into microcentrifuge tubes, followed by the addition of denatured salmon sperm DNA (50µg), 0.2µg of plasmid, 1ml polyethylene glycol (PEG) lithium acetate solution (40% (w/v) PEG 4000, 100 mM lithium acetate, 10 mM Tris-HCl, 1 mM EDTA, pH 7.5), and incubated for 45min at 30°C. This was followed by a 20min incubation at 42° and chilled for 2min on ice. Cells were washed and resuspended in 100µl H<sub>2</sub>O and plated on selective SC agar plates, incubated for ~3 days at 30°C.

### Protein extraction and western analysis

Cells were grown overnight in 5 ml SC-Ura, 2% raffinose at 30°C. The next day, fresh cultures were started with optical density OD<sub>600</sub> of 0.5 in 20 ml SC-Ura, 2% galactose at 30° for 6 hrs. Cells were then harvested, frozen in liquid nitrogen, and stored at -80°. For total protein extract, trichloroacetic acid was performed as described previously [43] and protein concentration was determined by BCA protein assay kit (Thermo scientific). Protein samples were separated by 4–12% gradient SDS-PAGE (GenScript) and blotted onto nitrocellulose or PVDF membranes. The following primary antibodies were used at 1:5000 dilution: anti-FLAG antibody (GenScript), and anti-GAPDH antibody (Proteintech). Secondary anti-mouse IgG HRP antibody was used at 1:7000 dilution (Promega). ChemiDoc (Bio-Rad) imaging system was used to detect chemiluminescence signals from western blots. Relative quantification of protein bands ( $\alpha$ -Flag,  $\alpha$ -GAPDH) from western blots was performed using Fiji [44].

### Cell growth assays and RC<sub>50</sub> measurements

Cells were grown overnight in 5ml SC-Ura, 2% raffinose at 30°C. The next day, fresh cultures were started with an OD<sub>600</sub> of 0.1 in SC-Ura, 2% galactose, with or without inhibitors and transferred to 96-well plates, incubated at 30°C on a rotary shaker. Three independent transformants were used to test each form of M<sup>Pro</sup>. Each transformant was sampled three times for each assay. The plate was transferred to a Tecan Infinite 200 PRO plate reader, and OD<sub>600</sub> measurements were taken at 0, 24, 48 and 72 hours, with 5 flashes per well. Excel (Microsoft) was used to analyze the raw data. As a measure of inhibitory activity of nirmatrelvir we calculated a Recovery Concentration (RC<sub>50</sub>). The slopes from the dose responses were calculated and used to estimate the concentration of inhibitor that improves growth to half-maximal relative to empty vector control after 72 hours of growth. As we only tested up to 200µM of nirmatrelvir, in cases where the RC<sub>50</sub> is beyond this concentration we indicate as >200µM.

### Yeast proteomics

Cells were grown overnight in 10ml SC-Ura + 2% Raffinose media at 30°C. The next day, fresh cultures were started with OD<sub>600</sub> of 0.1 in 100ml SC-Ura + 2% galactose at 30°C for 6 hr. Cells were then harvested, frozen in liquid nitrogen, and stored at -80°. Protein extraction was performed as described previously [43] and protein concentration was determined using Pierce BCA protein assay kit (Thermo scientific).

To determine changes in the proteome associated with expression of M<sup>Pro</sup> versus M<sup>Pro</sup> C145A, in-solution tryptic digestion was performed as described [45] followed by desalting with a Pierce Peptide Desalting Spin Columns per the manufacturer's protocol (ThermoFisher

Scientific, cat no. 89852) and the peptides were dried by vacuum centrifugation. 600 ng of the final sample was analyzed by mass spectrometry. HPLC-ESI-MS/MS was performed as previously described [46]. In brief, MS/MS was performed in positive ion mode on a Thermo Scientific Orbitrap Fusion Lumos tribrid mass spectrometer fitted with an EASY-Spray Source (Thermo Scientific, San Jose, CA). NanoLC was performed using a Thermo Scientific UltiMate 3000 RSLCnano System with an EASY Spray C18 LC column (Thermo Scientific).

Tandem mass spectra were extracted from Xcalibur 'RAW' files and charge states were assigned using the ProteoWizard 2.1.x msConvert script using the default parameters(23). The fragment mass spectra were then searched against the *Saccharomyces cerevisiae* (strain ATCC 204508 / S288c) (Baker's yeast) UniProt database (6067 entries) using Mascot (Matrix Science, London, UK; version 2.6) using the default probability cut-off score. Cross-correlation of Mascot search results with X! Tandem was accomplished with Scaffold (version Scaffold\_4.8.7; Proteome Software, Portland, OR, USA). Probability assessment of peptide assignments and protein identifications were made through the use of Scaffold. Only peptides with  $\geq 95\%$  probability were considered. Progenesis QI for proteomics software (version 2.4, Nonlinear Dynamics Ltd., Newcastle upon Tyne, UK) was used to perform ion-intensity based label-free quantification similar to as previously described [46]. Principal component analysis and unbiased hierarchical clustering analysis (heat map) was performed in Perseus [47,48]. Gene ontology and KEGG pathway enrichment analysis was performed with DAVID [49]. Proteomics data has been deposited to ProteomeXchange. Project accession: PXD036325 and Project DOI:10.6019/PXD036325.

### Recombinant M<sup>PTo</sup> and proteolytic activity assays

SARS-CoV-2 M<sup>PTo</sup> mutants were generated with QuikChange II Site-Directed Mutagenesis Kit from Agilent (Catalog #200524), using plasmid pE-SUMO-M<sup>PTo</sup> as the template. The plasmid produces tag-free M<sup>PTo</sup> protein with no extra residue at either N- or C-terminus upon removal of the SUMO tag by SUMO protease digestion [21].

SARS-CoV-2 M<sup>PTo</sup> mutant proteins were expressed and purified as previously described [21,50] with minor modifications. Plasmids were transformed into *E. coli* BL21(DE3) competent cells and bacterial cultures overexpressing the target proteins were grown in LB (Luria-Bertani) medium containing 50  $\mu\text{g}/\text{mL}$  of kanamycin at 37°C, and expression of the target protein was induced at an optical density (A600) of 0.6–0.8 by the addition of isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM. The cell culture was incubated at 18°C for 12–16 hrs. Bacterial cultures were harvested by centrifugation (8,000  $\times g$ , 10 min, 4°C) and resuspended in lysis buffer containing 25 mM Tris (pH 8.0), 750 mM NaCl, 2 mM DTT, 0.5 mg/mL lysozyme, 0.5 mM phenylmethylsulfonyl fluoride (PMSF) and 0.02 mg/mL DNase I. Bacterial cells were lysed by alternating sonication (30% amplitude, 1 s on/1 s off) and homogenization using a tissue grinder. The lysed cell suspension was clarified by centrifugation (18,000  $\times g$ , 30min, 4°C) and the supernatant was incubated with Ni-NTA resin for over 2 hrs at 4°C on a rotator. The Ni-NTA resin was thoroughly washed with 20 mM imidazole in washing buffer containing 50mM Tris (pH 8.0), 150mM NaCl, 2 mM DTT, and SUMO-M<sup>PTo</sup> protein was eluted with elution buffer containing 50 to 300mM imidazole, 50mM Tris (pH 8.0), 150mM NaCl, 2mM DTT. Fractions containing SUMO-M<sup>PTo</sup> proteins greater than 90% homogeneity were pooled and subjected to dialysis (two times) against a buffer containing 50mM Tris (pH 8.0), 150mM NaCl, 2mM DTT and 10% glycerol. SUMO protease digestion was carried out at 30°C for 1 hr to remove SUMO tag. Following digestion, SUMO Protease and SUMO tag were removed by Ni-NTA resin. The purified tag-free SARS-CoV-2 M<sup>PTo</sup> mutant proteins were fast frozen in liquid nitrogen and stored at -80°C.



For measurement of  $K_m/V_{max}$  of SARS-CoV-2 M<sup>Pro</sup> mutants, proteolytic reactions were carried out with optimized concentrations of the mutant proteins and a series of concentrations of FRET substrate, Dabcyl-KTSAVLQ/SGFRKME (Edans), ranging from 0 to 200  $\mu$ M in 100 $\mu$ L of reaction buffer containing 20mM HEPES (pH 6.5), 120mM NaCl, 0.4mM EDTA, 4mM DTT, and 20% glycerol at 30°C in a BioTek Cytation 5 imaging reader (Agilent) with filters for excitation at 360/40 nm and emission at 460/40 nm. The SARS-CoV-2 M<sup>Pro</sup> FRET substrate used in this study is Dabcyl-KTSAVLQ/SGFRKME (Edans), which corresponds to the sequence between viral polypeptide NSP4-NSP5 junction from SARS-CoV-2 [21,51]. Reactions were monitored every 90s, and the initial velocity of the proteolytic activity was calculated by linear regression for the first 15min of the kinetic progress curves. The initial velocity was plotted against the FRET substrate concentrations using the classic Michaelis-Menten equation in Prism 8 software.

For IC<sub>50</sub> measurements, optimized concentrations of the mutant proteins were incubated with series concentrations of GC-376, PF-00835231 or nirmatrelvir (PF-07321332) in 100 $\mu$ L of reaction buffer at 30°C for 15 min, and the reaction was initiated by adding 10 $\mu$ M FRET substrate. The reaction was monitored for 1 hr, and the initial velocity was calculated for the first 15min by linear regression. The IC<sub>50</sub> was determined by plotting the initial velocity against various concentrations of the compounds using log (inhibitor) vs response-variable slope in Prism 8 software.

For  $K_i$  measurements, optimized concentrations of the mutant proteins were added to 20 $\mu$ M FRET substrate with various concentrations of GC-376, PF-00835231 or nirmatrelvir (PF-07321332) in 200 $\mu$ L of reaction buffer at 30°C to initiate the proteolytic reaction. The reaction was monitored for 2 hrs and the initial velocity was calculated for the first 90 min by linear regression. The  $K_i$  was calculated by plotting the initial velocity against various concentrations of the compounds using Morrison plot (tight binding) in Prism 8 software.

### M<sup>Pro</sup> crystallization and structure determination

SARS-CoV-2 M<sup>Pro</sup> E166N/R was diluted to 5 mg/mL in protein buffer (50 mM Tris pH 7.0, 150 mM NaCl, 4 mM DTT). Protein for complex determination was incubated overnight at 4°C with 2mM GC-376. No precipitation was observed after incubation, and centrifugation was not necessary. Apo and complex crystals were grown using 1.5  $\mu$ L:1.5  $\mu$ L (protein:well solution) hanging drops and a well solution of 0.1 M MgCl<sub>2</sub>, 20% PEG 3350, 10% 1,6-hexanediol, 0.1 M HEPES pH 7.5, and 0.1 M LiSO<sub>4</sub>. E166N/R crystals grew overnight at 20°C. Crystals were cryoprotected using the well solution supplemented with 20% glycerol, and then flash-frozen in liquid nitrogen.

X-ray diffraction data (S4 Table) were collected at the Southeast Regional Collaborative Access Team (SER-CAT) 22-BM beamline at the Advanced Photon Source (APS) in Argonne, IL, and processed with HKL2000 and CCP4. PHASER was used for molecular replacement using a previously solved SARS-CoV-2 M<sup>Pro</sup> structure (PDB ID: 7LYH) as a reference model. The CCP4 suite, (23) Coot, (24) and the PDB REDO server ([pdb-redo.eu](https://pdb-redo.eu)) (25) were used to complete the model building and refinement. The PyMOL Molecular Graphics System (Schrödinger, LLC) was used to generate all images.

### Supporting information

**S1 Table. List of yeast proteins that displayed significant changes in abundance when the catalytic mutant M<sup>Pro</sup> was expressed versus wild-type M<sup>Pro</sup>.**

(XLSX)

**S2 Table. Presence or absence of canonical M<sup>Pro</sup> cleavage site(s) in the top 153 proteins (p<0.05) with reduced abundance in strains expressing wild-type M<sup>Pro</sup>.**  
(XLSX)

**S3 Table. Plasmids and primers used in this study.**  
(DOCX)

**S4 Table. X-ray Data Collection and Refinement Statistics Data Collection.**  
(DOCX)

**S1 Fig. Expression of SARS-CoV-2 Genes in yeast.** A) SARS-CoV-2 genes were cloned into a high copy yeast plasmid and regulated by a galactose inducible promoter (Gal1). Growth curve of yeast carrying EV, M<sup>Pro</sup> WT, or M<sup>Pro</sup> C145A in glucose (Glu) versus galactose (Gal). B) Spot assays were performed on glucose (left) or galactose (right) containing plates to determine the growth effects conferred by expression of the indicated genes. Two biological replicates were performed for each (two rows) EV and indicated gene. Samples (3ul) of a 5-fold serial dilution are spotted in each row. EV refers to yeast carrying an empty vector. M, E, S, N are structural proteins and all others are non-structural proteins (NSP).  
(TIFF)

**S2 Fig. Determining potential yeast substrates of M<sup>Pro</sup>.** Schematic of the work flow to perform proteomics from yeast strains expressing wild-type M<sup>Pro</sup> or catalytically inactive M<sup>Pro</sup> C145A.  
(TIFF)

**S3 Fig. Effects of nirmatrelvir on yeast expressing M<sup>Pro</sup> from SARS-CoV-1 and Bat-CoV-HKU9.** Growth of yeast expressing M<sup>Pro</sup> from the indicated coronavirus is plotted in the presence or absence of nirmatrelvir. M<sup>Pro</sup> from both viruses remain responsive to nirmatrelvir as indicated by improved growth (nearly matching the empty vector (EV) control) in the presence of nirmatrelvir. (\*, p<0.01; \*\*\*, p<0.001) indicates differences compared to EV.  
(TIFF)

**S4 Fig. Growth curves for yeast expressing SARS-CoV-2 M<sup>Pro</sup>.** A) Indicated M<sup>Pro</sup> mutants were expressed in yeast. Wild-type M<sup>Pro</sup>, M<sup>Pro</sup> E166R, M<sup>Pro</sup> P132H, and M<sup>Pro</sup> N142A all confer a strong growth defect compared to empty vector (EV) control. On the other hand M<sup>Pro</sup> E166N and M<sup>Pro</sup> E166D do not confer a growth defect and grow as well as EV control. B-D) Growth curves for M<sup>Pro</sup> E166R, M<sup>Pro</sup> P132H, and M<sup>Pro</sup> N142A treated with 0, 25, 50, 100, 200mM of nirmatrelvir. E) Western blot showing levels of M<sup>Pro</sup> WT, E166R, and N142A.  
(TIFF)

**S5 Fig. Yeast growth assays show M<sup>Pro</sup> mutants E166A and E166V are resistant to nirmatrelvir.** A) Total growth of cultures after 72 hours expressing M<sup>Pro</sup> WT, E166A, and E166V. B) Western analysis shows that mutants and wild-type M<sup>Pro</sup> are expressed at comparable levels. C-D) E166A and E166V appear to be resistant as there is a lack of growth even when treated with 200mM of nirmatrelvir. (\*\*\*) indicates differences (p<0.001) compared to EV. Ratios of Flag:GAPDH signals relative to M<sup>Pro</sup> WT is shown at the bottom of each lane.  
(TIFF)

**S6 Fig. Enzymatic assays demonstrate that E166A and E166V are resistant to nirmatrelvir and other M<sup>Pro</sup> inhibitors.** A) Michaelis–Menten plot of M<sup>Pro</sup> and its mutants with various concentrations of FRET substrate. The Km, Vmax, kcat, and kcat/Km values are shown in the table. B) The IC50 plots of nirmatrelvir against M<sup>Pro</sup> E166A and E166V. C) Ki plots of

nirmatrelvir, GC-376, and PF-00835231 against M<sup>PRO</sup> E166A and E166V. (TIFF)

## Acknowledgments

JSC thanks members of the Choy lab for their support and special thanks to Drs. Stephen J. Kron and Sean Palacek for critical discussions related to this work. We are grateful to the staff and scientists at SER-CAT for assistance with X-ray diffraction data collection. JSC dedicates this work to the loving memory of Yin Y. Choy.

## Author Contributions

**Conceptualization:** Jun Wang, Yu Chen, John S. Choy.

**Formal analysis:** Jin Ou.

**Funding acquisition:** Jun Wang, Yu Chen, John S. Choy.

**Investigation:** Jin Ou, Eric M. Lewandowski, Yanmei Hu, Austin A. Lipinski, Ali Aljasser, Mariliz Colon-Ascanio, Ryan T. Morgan, Lian M. C. Jacobs, Xiujun Zhang, Melissa J. Bikowitz, Paul R. Langlais, Haozhou Tan.

**Project administration:** Jun Wang, Yu Chen, John S. Choy.

**Supervision:** Jun Wang, Yu Chen, John S. Choy.

**Writing – original draft:** John S. Choy.

**Writing – review & editing:** Jin Ou, Jun Wang, Yu Chen, John S. Choy.

## References

1. Ng TI, Correia I, Seagal J, DeGoey DA, Schrimpf MR, Hardee DJ, et al. Antiviral Drug Discovery for the Treatment of COVID-19 Infections. *Viruses*. 2022; 14(5). Epub 2022/05/29. <https://doi.org/10.3390/v14050961> PMID: 35632703.
2. Jin Z, Du X, Xu Y, Deng Y, Liu M, Zhao Y, et al. Structure of M(pro) from SARS-CoV-2 and discovery of its inhibitors. *Nature*. 2020; 582(7811):289–93. Epub 2020/04/10. <https://doi.org/10.1038/s41586-020-2223-y> PMID: 32272481.
3. Chia CSB, Xu W, Shuyi Ng P. A Patent Review on SARS Coronavirus Main Protease (3CL(pro)) Inhibitors. *ChemMedChem*. 2022; 17(1):e202100576. Epub 2021/10/16. <https://doi.org/10.1002/cmde.202100576> PMID: 34651447; PubMed Central PMCID: PMC8653044.
4. Hung HC, Ke YY, Huang SY, Huang PN, Kung YA, Chang TY, et al. Discovery of M Protease Inhibitors Encoded by SARS-CoV-2. *Antimicrob Agents Chemother*. 2020; 64(9). Epub 2020/07/17. <https://doi.org/10.1128/AAC.00872-20> PMID: 32669265; PubMed Central PMCID: PMC7449189.
5. Morse JW, Gladkikh TM, Hackenburg DM, Gould RK. COVID-19 and human-nature relationships: Vermonters' activities in nature and associated nonmaterial values during the pandemic. *PLoS One*. 2020; 15(12):e0243697. Epub 2020/12/12. <https://doi.org/10.1371/journal.pone.0243697> PMID: 33306716; PubMed Central PMCID: PMC7732125.
6. FDA. Fact sheet for healthcare providers: Emergency use authorization for Paxlovid 2021. Available from: <https://www.fda.gov/media/155050/download>.
7. Owen DR, Allerton CMN, Anderson AS, Aschenbrenner L, Avery M, Berritt S, et al. An oral SARS-CoV-2 M(pro) inhibitor clinical candidate for the treatment of COVID-19. *Science*. 2021; 374(6575):1586–93. Epub 2021/11/03. <https://doi.org/10.1126/science.abl4784> PMID: 34726479.
8. Hammond J, Leister-Tebbe H, Gardner A, Abreu P, Bao W, Wisemandle W, et al. Oral Nirmatrelvir for High-Risk, Nonhospitalized Adults with Covid-19. *N Engl J Med*. 2022; 386(15):1397–408. Epub 2022/02/17. <https://doi.org/10.1056/NEJMoa2118542> PMID: 35172054; PubMed Central PMCID: PMC8908851.
9. Boras B, Jones RM, Anson BJ, Arenson D, Aschenbrenner L, Bakowski MA, et al. Preclinical characterization of an intravenous coronavirus 3CL protease inhibitor for the potential treatment of COVID19. *Nat*

- Commun. 2021; 12(1):6055. Epub 2021/10/20. <https://doi.org/10.1038/s41467-021-26239-2> PMID: 34663813; PubMed Central PMCID: PMC8523698.
10. Cully M. A tale of two antiviral targets—and the COVID-19 drugs that bind them. *Nat Rev Drug Discov.* 2022; 21(1):3–5. Epub 2021/12/04. <https://doi.org/10.1038/d41573-021-00202-8> PMID: 34857884.
  11. Mukae H, Yotsuyanagi H, Ohmagari N, Doi Y, Imamura T, Sonoyama T, et al. A Randomized Phase 2/3 Study of Ensitrelvir, a Novel Oral SARS-CoV-2 3C-Like Protease Inhibitor, in Japanese Patients with Mild-to-Moderate COVID-19 or Asymptomatic SARS-CoV-2 Infection: Results of the Phase 2a Part. *Antimicrob Agents Chemother.* 2022; 66(10):e0069722. Epub 2022/09/14. <https://doi.org/10.1128/aac.00697-22> PMID: 36098519; PubMed Central PMCID: PMC9578433.
  12. CDC. Information for laboratories about Coronavirus (COVID-19). Centers for Disease Control and Prevention; 2020. Available from: <https://www.cdc.gov/coronavirus/2019-ncov/lab/lab-biosafety-guidelines.html>.
  13. Zhu W, Xu M, Chen CZ, Guo H, Shen M, Hu X, et al. Identification of SARS-CoV-2 3CL Protease Inhibitors by a Quantitative High-Throughput Screening. *ACS Pharmacol Transl Sci.* 2020; 3(5):1008–16. Epub 2020/10/17. <https://doi.org/10.1021/acsptsci.0c00108> PMID: 33062953; PubMed Central PMCID: PMC7507806.
  14. Jochmans D, Liu C, Donckers K, Stoycheva A, Boland S, Stevens SK, 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. *mBio.* 2023; 14(1):e0281522. Epub 2023/01/11. <https://doi.org/10.1128/mbio.02815-22> PMID: 36625640; PubMed Central PMCID: PMC9973015.
  15. Zhou Y, Gammeltoft KA, Ryberg LA, Pham LV, Fahnøe U, Binderup A, et al. Nirmatrelvir Resistant SARS-CoV-2 Variants with High Fitness in Vitro. *bioRxiv.* 2022:2022.06.06.494921. <https://doi.org/10.1101/2022.06.06.494921>
  16. Iketani S, Hong SJ, Sheng J, Bahari F, Culbertson B, Atanaki FF, et al. Functional map of SARS-CoV-2 3CL protease reveals tolerant and immutable sites. *Cell Host Microbe.* 2022; 30(10):1354–62 e6. Epub 2022/08/28. <https://doi.org/10.1016/j.chom.2022.08.003> PMID: 36029764; PubMed Central PMCID: PMC9365866.
  17. Iketani S, Mohri H, Culbertson B, Hong SJ, Duan Y, Luck MI, et al. Multiple pathways for SARS-CoV-2 resistance to nirmatrelvir. *Nature.* 2023; 613(7944):558–64. Epub 2022/11/10. <https://doi.org/10.1038/s41586-022-05514-2> PMID: 36351451; PubMed Central PMCID: PMC9849135 development of inhibitors against the SARS-CoV-2 3CL protease. D.D.H. is a cofounder of TaiMed Biologics and RenBio, consultant to WuXi Biologics and Bria Biosciences and board director for Vicarious Surgical.
  18. V'Kovski P, Kratzel A, Steiner S, Stalder H, Thiel V. Coronavirus biology and replication: implications for SARS-CoV-2. *Nat Rev Microbiol.* 2021; 19(3):155–70. Epub 2020/10/30. <https://doi.org/10.1038/s41579-020-00468-6> PMID: 33116300; PubMed Central PMCID: PMC7592455.
  19. Frieman M, Basu D, Matthews K, Taylor J, Jones G, Pickles R, et al. Yeast based small molecule screen for inhibitors of SARS-CoV. *PLoS One.* 2011; 6(12):e28479. Epub 2011/12/14. <https://doi.org/10.1371/journal.pone.0028479> PMID: 22164298; PubMed Central PMCID: PMC3229576.
  20. Kneller DW, Phillips G, O'Neill HM, Jedrzejczak R, Stols L, Langan P, et al. Structural plasticity of SARS-CoV-2 3CL M(pro) active site cavity revealed by room temperature X-ray crystallography. *Nat Commun.* 2020; 11(1):3202. Epub 2020/06/26. <https://doi.org/10.1038/s41467-020-16954-7> PMID: 32581217; PubMed Central PMCID: PMC7314768.
  21. Sacco MD, Ma C, Lagarias P, Gao A, Townsend JA, Meng X, 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.* 2020; 6(50). Epub 2020/11/08. <https://doi.org/10.1126/sciadv.abe0751> PMID: 33158912; PubMed Central PMCID: PMC7725459.
  22. Dai W, Zhang B, Jiang XM, Su H, Li J, Zhao Y, et al. Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. *Science.* 2020; 368(6497):1331–5. Epub 2020/04/24. <https://doi.org/10.1126/science.abb4489> PMID: 32321856; PubMed Central PMCID: PMC7179937.
  23. Hegyi A, Ziebuhr J. Conservation of substrate specificities among coronavirus main proteases. *J Gen Virol.* 2002; 83(Pt 3):595–9. Epub 2002/02/14. <https://doi.org/10.1099/0022-1317-83-3-595> PMID: 11842254.
  24. Shaqra AM, Zvornicanin SN, Huang QYJ, Lockbaum GJ, Knapp M, Tandeske L, et al. Defining the substrate envelope of SARS-CoV-2 main protease to predict and avoid drug resistance. *Nat Commun.* 2022; 13(1):3556. Epub 2022/06/22. <https://doi.org/10.1038/s41467-022-31210-w> PMID: 35729165.
  25. Matthew AN, Leidner F, Lockbaum GJ, Henes M, Zephyr J, Hou S, et al. Drug Design Strategies to Avoid Resistance in Direct-Acting Antivirals and Beyond. *Chem Rev.* 2021; 121(6):3238–70. Epub 2021/01/08. <https://doi.org/10.1021/acs.chemrev.0c00648> PMID: 33410674; PubMed Central PMCID: PMC8126998.

26. Xiong M, Su H, Zhao W, Xie H, Shao Q, Xu Y. What coronavirus 3C-like protease tells us: From structure, substrate selectivity, to inhibitor design. *Med Res Rev.* 2021; 41(4):1965–98. Epub 2021/01/19. <https://doi.org/10.1002/med.21783> PMID: 33460213; PubMed Central PMCID: PMC8014231.
27. MacDonald EA, Frey G, Namchuk MN, Harrison SC, Hinshaw SM, Windsor IW. Recognition of Divergent Viral Substrates by the SARS-CoV-2 Main Protease. *ACS Infect Dis.* 2021; 7(9):2591–5. Epub 2021/08/27. <https://doi.org/10.1021/acscinfecdis.1c00237> PMID: 34437808; PubMed Central PMCID: PMC8424689.
28. Xue X, Yu H, Yang H, Xue F, Wu Z, Shen W, et al. Structures of two coronavirus main proteases: implications for substrate binding and antiviral drug design. *J Virol.* 2008; 82(5):2515–27. Epub 2007/12/21. <https://doi.org/10.1128/JVI.02114-07> PMID: 18094151; PubMed Central PMCID: PMC2258912.
29. Greasley SE, Noell S, Plotnikova O, Ferre R, Liu W, Bolanos B, et al. Structural basis for the in vitro efficacy of nirmatrelvir against SARS-CoV-2 variants. *J Biol Chem.* 2022; 298(6):101972. Epub 2022/04/25. <https://doi.org/10.1016/j.jbc.2022.101972> PMID: 35461811; PubMed Central PMCID: PMC9023115.
30. Sacco MD, Hu Y, Gongora MV, Meilleur F, Kemp MT, Zhang X, 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.* 2022; 32(5):498–500. Epub 2022/03/17. <https://doi.org/10.1038/s41422-022-00640-y> PMID: 35292745; PubMed Central PMCID: PMC8923085.
31. Ullrich S, Ekanayake KB, Otting G, Nitsche C. Main protease mutants of SARS-CoV-2 variants remain susceptible to nirmatrelvir. *Bioorg Med Chem Lett.* 2022; 62:128629. Epub 2022/02/20. <https://doi.org/10.1016/j.bmcl.2022.128629> PMID: 35182772; PubMed Central PMCID: PMC8856729.
32. Li P, Wang Y, Lavrijsen M, Lamers MM, de Vries AC, Rottier RJ, et al. SARS-CoV-2 Omicron variant is highly sensitive to molnupiravir, nirmatrelvir, and the combination. *Cell Res.* 2022; 32(3):322–4. Epub 2022/01/22. <https://doi.org/10.1038/s41422-022-00618-w> PMID: 35058606; PubMed Central PMCID: PMC8771185.
33. Chang MW, Torbett BE. Accessory mutations maintain stability in drug-resistant HIV-1 protease. *J Mol Biol.* 2011; 410(4):756–60. Epub 2011/07/19. <https://doi.org/10.1016/j.jmb.2011.03.038> PMID: 21762813; PubMed Central PMCID: PMC3139113.
34. De Luca A. The impact of resistance on viral fitness and its clinical implications. In: Geretti AM, editor. *Antiretroviral Resistance in Clinical Practice.* London2006.
35. Sa-Correia I, Godinho CP. Exploring the biological function of efflux pumps for the development of superior industrial yeasts. *Curr Opin Biotechnol.* 2022; 74:32–41. Epub 2021/11/16. <https://doi.org/10.1016/j.copbio.2021.10.014> PMID: 34781103.
36. Mukhopadhyay K, Kohli A, Prasad R. Drug susceptibilities of yeast cells are affected by membrane lipid composition. *Antimicrob Agents Chemother.* 2002; 46(12):3695–705. Epub 2002/11/19. <https://doi.org/10.1128/AAC.46.12.3695-3705.2002> PMID: 12435664; PubMed Central PMCID: PMC132749.
37. Meyer B, Chiaravalli J, Gellenoncourt S, Brownridge P, Bryne DP, Daly LA, et al. Characterising proteolysis during SARS-CoV-2 infection identifies viral cleavage sites and cellular targets with therapeutic potential. *Nat Commun.* 2021; 12(1):5553. Epub 2021/09/23. <https://doi.org/10.1038/s41467-021-25796-w> PMID: 34548480; PubMed Central PMCID: PMC8455558.
38. Pablos I, Machado Y, de Jesus HCR, Mohamud Y, Kappelhoff R, Lindskog C, et al. Mechanistic insights into COVID-19 by global analysis of the SARS-CoV-2 3CL(pro) substrate degradome. *Cell Rep.* 2021; 37(4):109892. Epub 2021/10/22. <https://doi.org/10.1016/j.celrep.2021.109892> PMID: 34672947; PubMed Central PMCID: PMC8501228.
39. Alalam H, Sigurdardottir S, Bourgard C, Tiukova I, King RD, Grotli M, et al. A Genetic Trap in Yeast for Inhibitors of SARS-CoV-2 Main Protease. *mSystems.* 2021; 6(6):e0108721. Epub 2021/11/24. <https://doi.org/10.1128/mSystems.01087-21> PMID: 34812651; PubMed Central PMCID: PMC8609969.
40. Flynn JM, Samant N, Schneider-Nachum G, Bakan DT, Yilmaz NK, Schiffer CA, et al. Comprehensive fitness landscape of SARS-CoV-2 M(pro) reveals insights into viral resistance mechanisms. *Elife.* 2022;11. Epub 2022/06/21. <https://doi.org/10.7554/eLife.77433> PMID: 35723575.
41. Padhi AK, Tripathi T. 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. *bioRxiv.* 2022:2022.07.12.499687. <https://doi.org/10.1016/j.bbrc.2022.09.010> PMID: 36113178
42. Hu Y, Lewandowski EM, Tan H, Morgan RT, Zhang X, Jacobs LMC, et al. Naturally occurring mutations of SARS-CoV-2 main protease confer drug resistance to nirmatrelvir. *bioRxiv.* 2022:2022.06.28.497978. <https://doi.org/10.1101/2022.06.28.497978> PMID: 36119652
43. Shah SB, Parmiter D, Constantine C, Elizalde P, Naldrett M, Karpova TS, et al. Glucose Signaling Is Connected to Chromosome Segregation Through Protein Kinase A Phosphorylation of the Dam1 Kinetochores Subunit in *Saccharomyces cerevisiae*. *Genetics.* 2019; 211(2):531–47. Epub 2018/12/14.



- <https://doi.org/10.1534/genetics.118.301727> PMID: 30546002; PubMed Central PMCID: PMC6366914.
44. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods*. 2012; 9(7):676–82. Epub 2012/06/30. <https://doi.org/10.1038/nmeth.2019> PMID: 22743772; PubMed Central PMCID: PMC3855844.
  45. Pendleton AL, Antolic AT, Kelly AC, Davis MA, Camacho LE, Doubleday K, et al. Lower oxygen consumption and Complex I activity in mitochondria isolated from skeletal muscle of fetal sheep with intra-uterine growth restriction. *Am J Physiol Endocrinol Metab*. 2020; 319(1):E67–E80. Epub 2020/05/13. <https://doi.org/10.1152/ajpendo.00057.2020> PMID: 32396498; PubMed Central PMCID: PMC7468780.
  46. Parker SS, Krantz J, Kwak EA, Barker NK, Deer CG, Lee NY, et al. Insulin Induces Microtubule Stabilization and Regulates the Microtubule Plus-end Tracking Protein Network in Adipocytes. *Mol Cell Proteomics*. 2019; 18(7):1363–81. Epub 2019/04/26. <https://doi.org/10.1074/mcp.RA119.001450> PMID: 31018989; PubMed Central PMCID: PMC6601206.
  47. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods*. 2016; 13(9):731–40. Epub 2016/06/28. <https://doi.org/10.1038/nmeth.3901> PMID: 27348712.
  48. Tyanova S, Cox J. Perseus: A Bioinformatics Platform for Integrative Analysis of Proteomics Data in Cancer Research. *Methods Mol Biol*. 2018; 1711:133–48. Epub 2018/01/19. [https://doi.org/10.1007/978-1-4939-7493-1\\_7](https://doi.org/10.1007/978-1-4939-7493-1_7) PMID: 29344888.
  49. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009; 4(1):44–57. Epub 2009/01/10. <https://doi.org/10.1038/nprot.2008.211> PMID: 19131956.
  50. Hu Y, Ma C, Szeto T, Hurst B, Tarbet B, Wang J. Boceprevir, calpain inhibitors II and XII, and GC-376 have broad-spectrum antiviral activity against coronaviruses in cell culture. *bioRxiv*. 2020. Epub 2020/11/04. <https://doi.org/10.1101/2020.10.30.362335> PMID: 33140049; PubMed Central PMCID: PMC7605558.
  51. Ma C, Sacco MD, Hurst B, Townsend JA, Hu Y, Szeto T, et al. Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease. *Cell Res*. 2020; 30(8):678–92. Epub 2020/06/17. <https://doi.org/10.1038/s41422-020-0356-z> PMID: 32541865; PubMed Central PMCID: PMC7294525 identified compounds for COVID-19.