# Extended Data Fig and table

Extended Data Table 1 X-ray crystallography data collection and refinement statistics.

Extended Data Fig. 1: Size exclusion chromatography (SEC) elution profiles of SARS-CoV-2 $M^{pro}$ wild-type and mutants (T45I, Δ23G and Δ23G/T45I).

Extended Data Fig. 2: Electron density map showing the nirmatrelvir-binding site of SARS-CoV-2 $M^{pro}$ mutants Δ23G and Δ23G/T45I.

Extended Data Fig. 3: Structural comparison of SARS-CoV-2 $M^{pro}$ mutants (Δ23G, Δ23G/T45I) and wild-type $M^{pro}$ bound to various inhibitors.

Extended Data Fig. 4: Histopathological assessment of nasal turbinate and lung tissues in Syrian golden hamsters infected with recombinant SARS-CoV-2 variants.

Extended Data Fig. 5: Genetic stability of recombinant SARS-CoV-2 variants after serial passaging in vitro.

Extended Data Fig. 6: Global sequence analysis of SARS-CoV-2 $M^{pro}$ residues 23 and 45.

**Extended Data Table 1 X-ray crystallography data collection and refinement statistics.**

| | SARS-CoV-2 $M^{pro}$-Δ23G | SARS-CoV-2 $M^{pro}$-Δ23G: nirmatrelvir | SARS-CoV-2 $M^{pro}$-Δ23G/T45I: nirmatrelvir | SARS-CoV-2 $M^{pro}$-Δ23G/T45I: nirmatrelvir |
|---|---|---|---|---|
| **Data collection** | | | | |
| Beam Line | | PLS BL-5C | | |
| Wavelength, Å | 0.9793 | | 1.0000 | |
| Space group | *C*2 | *C*2 | *C*2 | $P2_1$ |
| Unit cell parameters | | | | |
| a, b, c (Å) | 114.0; 53.9; 44.4 | 115.3; 53.7; 45.6 | 115.0; 53.8; 45.7 | 45.6; 53.6; 115.5 |
| α, β, γ (°) | 90.0; 101.0; 90.0 | 90.0; 102.3; 90.0 | 90.0; 102.2; 90.0 | 90.0; 101.4; 90.0 |
| [a]Resolution, Å | 50-1.75 (1.81-1.75) | 50-2.10 (2.18-2.10) | 50-2.40 (2.49-2.40) | 50-2.35 (2.43-2.35) |
| No. of total reflections | 468,150 | 403,526 | 184,794 | 730,483 |
| No. of unique reflections | 26,882 | 15,898 | 10,774 | 22,740 |
| No. in asymmetric unit | 1 | 1 | 1 | 2 |
| [a]Completeness, % | 96.2 (81.1) | 95.3 (90.0) | 96.6 (83.2) | 90.2 (81.3) |
| [a]$I/\sigma(I)$ | 21.3 (1.5) | 11.5 (1.6) | 13.5 (1.8) | 13.4 (1.7) |
| [a]Redundancy | 4.6 (2.4) | 4.3 (2.6) | 4.9 (2.9) | 3.9 (2.4) |
| [b]$R_{merge}$, % | 6.5 (42.2) | 6.2 (27.2) | 7.2 (26.1) | 5.1 (23.6) |
| [a]$CC_{1/2}$, % | 99.7 (68.5) | 99.7 (80.0) | 99.5 (85.2) | 99.6 (80.7) |
| **Refinement** | | | | |
| Resolution, Å | 50-1.75 (1.81-1.75) | 50-2.1 (2.18-2.10) | 50-2.4 (2.49-2.40) | 50-2.35 (2.43-2.35) |
| [c]$R_{cryst}/R_{free}$, % | 19.2/23.8 | 17.5/23.0 | 21.1/25.1 | 20.2/24.4 |
| No. of protein atoms | 2324 | 2324 | 2324 | 4648 |
| No. of water molecules | 151 | 130 | 42 | 75 |
| No. of ligand molecules | 0 | 1 | 1 | 2 |

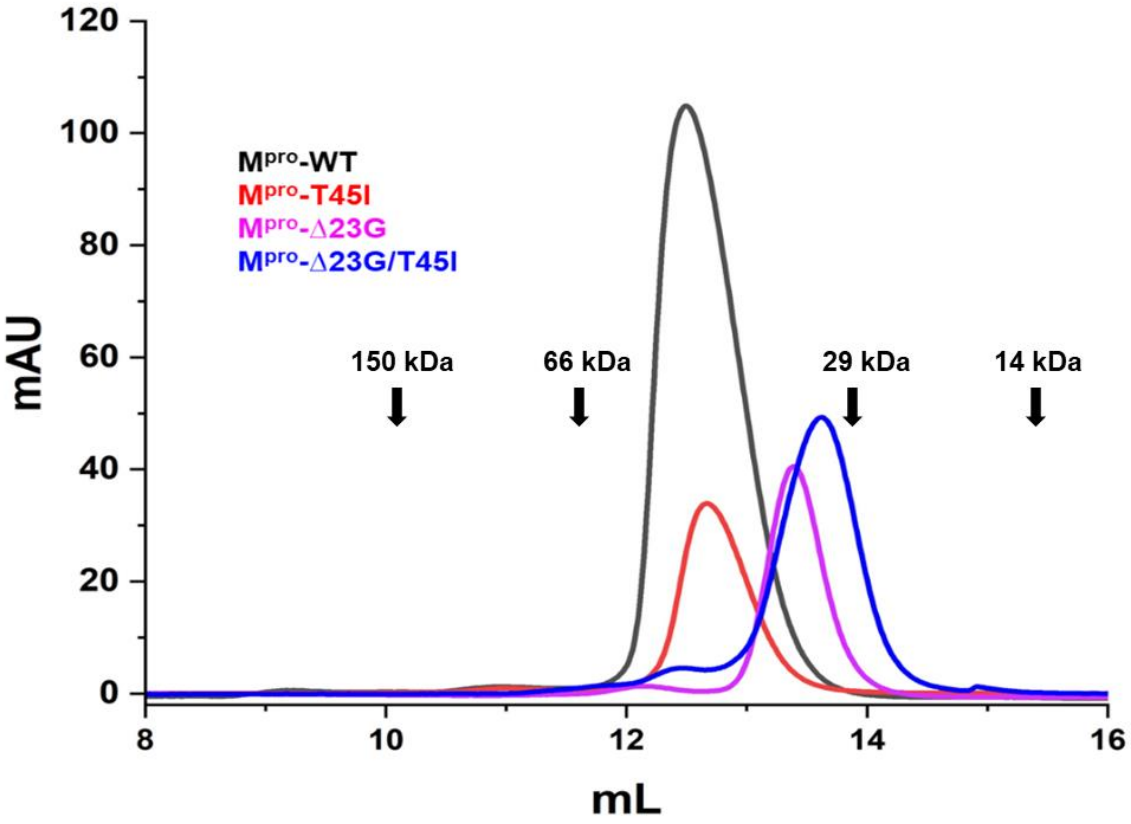
| | | | | |
|---|---|---|---|---|
| Bond lengths, Å | 0.005 | 0.007 | 0.002 | 0.003 |
| Bond angles, º | 0.857 | 0.903 | 0.493 | 0.646 |
| Average *B*-factor, Å$^2$ | 24.7 | 27.6 | 35.0 | 36.6 |
| Ramachandran analysis | | | | |
| Favored, % | 95.3 | 97.6 | 95.6 | 95.2 |
| Allowed, % | 4.7 | 2.4 | 4.4 | 4.8 |
| Outliers | 0 | 0 | 0 | 0 |
| Rotamer outliers, % | 0.7 | 1.9 | 2.3 | 2.1 |
| PDB entry | 9M9N | 9M9R | 9MA3 | 9MA6 |

[a] Values in parentheses represent statistics for the highest-resolution shell.

[b] $R_{merge} = \sum_h \sum_i \left| I(h,i) - \langle I(h) \rangle \right| / \sum_h \sum_i I(h,i)$, where $I(h,i)$ is the intensity of the i$^{th}$ measurement of reflection $h$ and $\langle I(h) \rangle$ is the mean value of $I(h,i)$ for all i measurements.

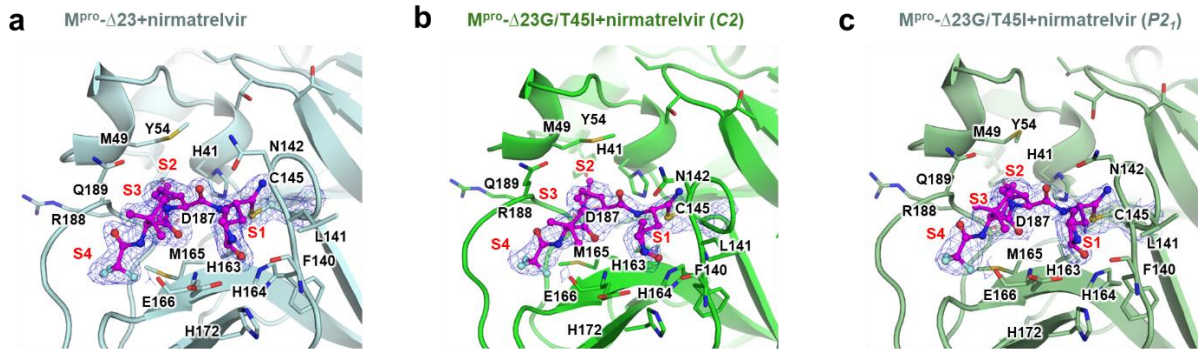
[c] $R_{free}$ was calculated from randomly selected 5% set of reflections not included in the calculation of the $R$ value.

**Extended Data Fig. 1: Size-exclusion chromatography (SEC) elution profiles of SARS-CoV-2 $M^{pro}$ wild-type and mutants (T45I, Δ23G, and Δ23G/T45I).**



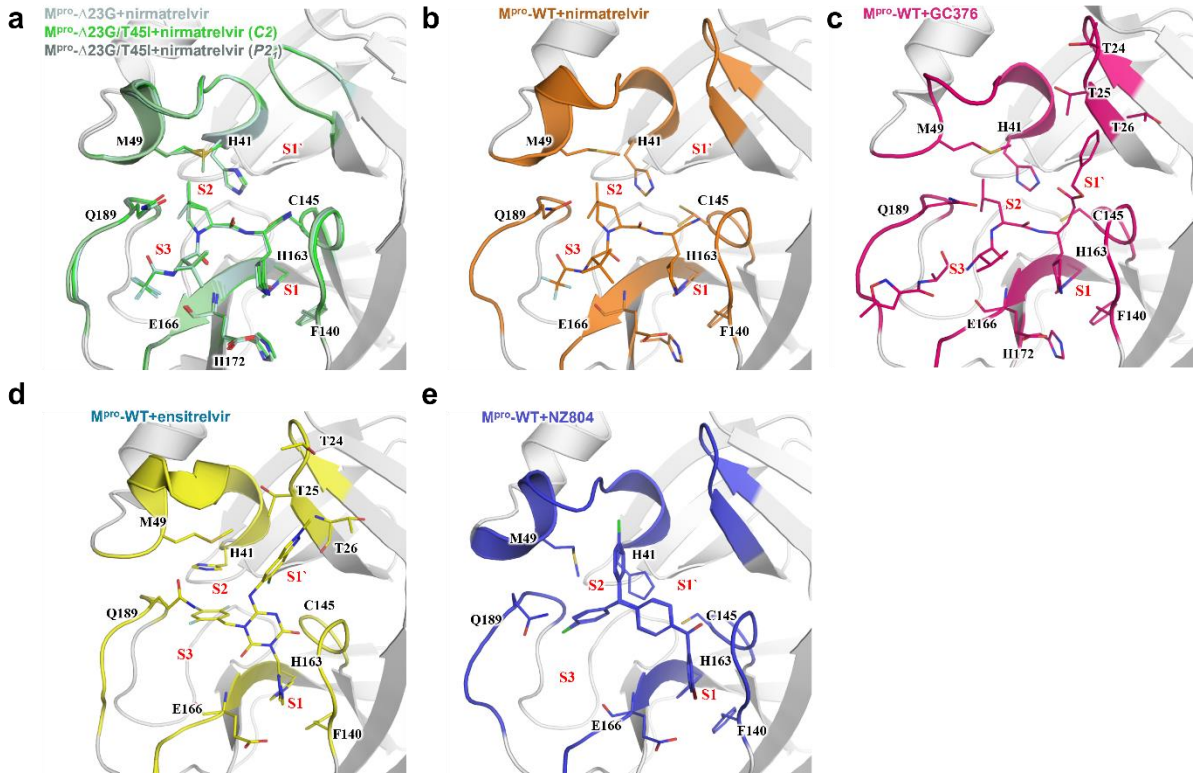

Gel-dfiltration chromatography was performed using a Superdex 75 prep grade (HiLoad 26/60) column. Elution profiles are shown for $M^{pro}$-WT (black), T45I (red), Δ23G (magenta), and Δ23G/T45I (blue). $M^{pro}$-WT and T45I eluted at volumes corresponding to the molecular mass of a dimer, whereas Δ23G and Δ23G/T45I mutants eluted at volumes consistent with a monomer. Molecular mass standards (with elution volumes indicated in parentheses) were alcohol dehydrogenase (150 kDa, V0 = 10.04 mL), bovine serum albumin (66 kDa, 11.32 mL), carbonic anhydrase (29 kDa, 13.9 mL), and α-lactalbumin (14 kDa, 15.5 mL).

**Extended Data Fig. 2: Electron density map showing the nirmatrelvir-binding site of SARS-CoV-2 $M^{pro}$ mutants Δ23G and Δ23G/T45I.**



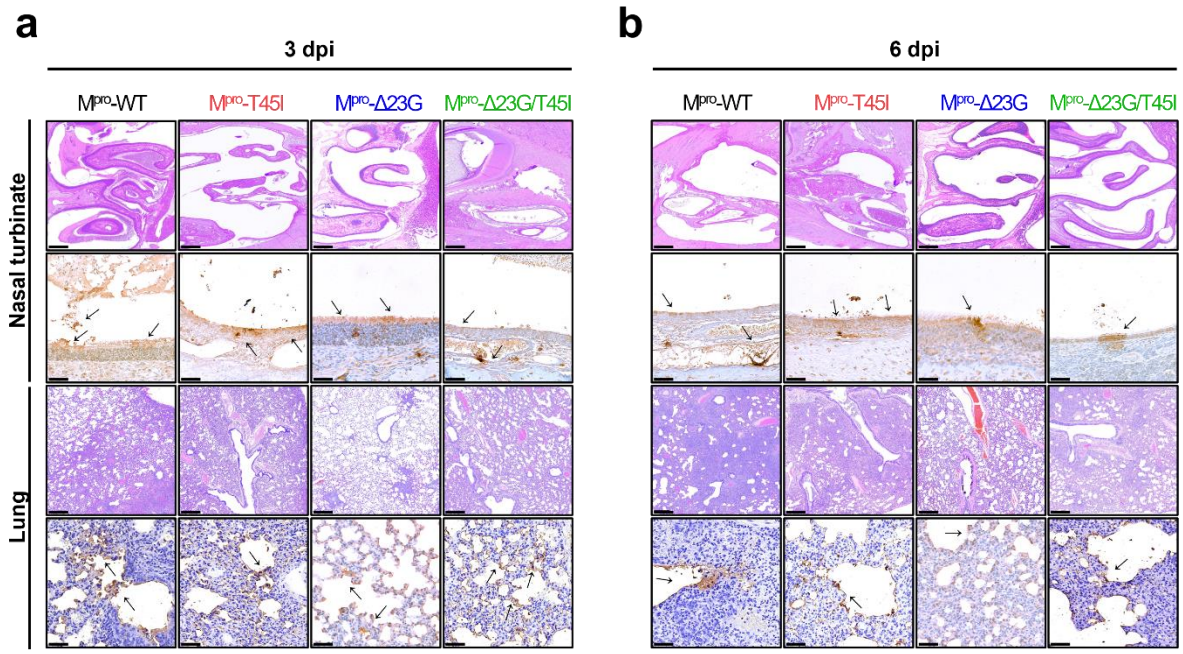

**a**–**c,** Ribbon diagrams illustrate the inhibitor binding sites of $M^{pro}$-Δ23G (**a**, pale cyan), $M^{pro}$-Δ23G/T45I in the C2 space group (**b**, green), and $M^{pro}$-Δ23G/T45I in the $P2_1$ space group (**c**, pale green). Nirmatrelvir is shown as a stick-and-ball model (magenta), overlaid with a 2Fo-Fc electron density map contoured at 1σ. The key substrate-binding pockets (S1–S4) and interacting residues critical for inhibitor binding are labeled for clarity.

**Extended Data Fig. 3: Structural comparison of SARS-CoV-2 M$^{pro}$ mutants (Δ23G, Δ23G/T45I) and wild-type M$^{pro}$ bound to various inhibitors.**



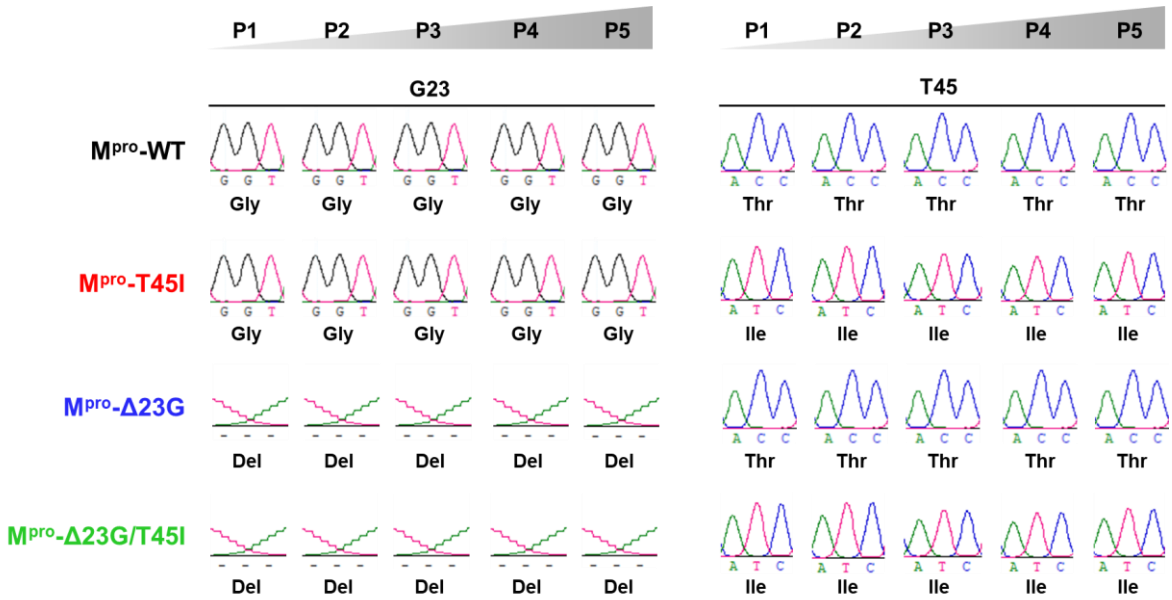

Superposition of crystal structures highlighting inhibitor-binding modes and substrate-binding subsites (S1, S1', S2, and S3). **a**, Structural alignment of M$^{pro}$-Δ23G (pale cyan), M$^{pro}$-Δ23G/T45I in space group C2 (green), and M$^{pro}$-Δ23G/T45I in space group $P2_1$ (pale green) bound to nirmatrelvir, illustrating structural variations among the mutant complexes. **b**–**e**, Wild-type M$^{pro}$ complexes with nirmatrelvir (**b**, oragne, PDB: 7VH8), GC376 (**c**, pink, PDB: 6LU7), ensitrelvir (**d**, yellow, PDB: 8HEF), and NZ804 (**e**, blue, PDB: 8W1T), highlighting critical active-site residues interacting with respective inhibitors. Active-site residues and inhibitors are colored consistently, and key binding pockets are clearly labeled.

**Extended Data Fig. 4: Histopathological assessment of nasal turbinate and lung tissues in Syrian golden hamsters infected with recombinant SARS-CoV-2 variants.**




**a,b** Representative images showing Hematoxylin and Eosin (HE, upper rows) staining for tissue pathology and immunohistochemistry (IHC, lower rows) staining for SARS-CoV-2 nucleocapsid antigen detection (brown signals, indicated by arrows) at 3 days post-infection (dpi) (**a**) and 6 dpi (**b**). Scale bars: nasal turbinate HE, 100 µm; nasal turbinate IHC, 50 µm; lung H&E, 100 µm; lung IHC, 20 µm. Severe tissue damage and abundant viral antigen were observed in nasal turbinate and lung tissues of hamsters infected with $M^{pro}$-WT and $M^{pro}$-T45I variants. Conversely, tissues infected with Δ23G-containing variants ($M^{pro}$-Δ23G and $M^{pro}$-Δ23G/T45I) exhibited milder pathology and reduced antigen presence, highlighting attenuation linked to the Δ23G mutation.

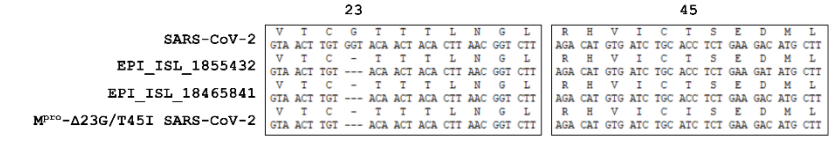
**Extended Data Fig. 5: Genetic stability of recombinant SARS-CoV-2 variants after serial passaging in vitro.**




Recombinant SARS-CoV-2 variants ($M^{pro}$-T45I, $M^{pro}$-Δ23G, and $M^{pro}$-Δ23G/T45I) were serially passaged five times in Vero E6 cells at a multiplicity of infection (MOI) of 0.001 without antiviral pressure. Viral RNA extracted after each passage was subjected to Sanger sequencing to assess stability of introduced mutations. Chromatograms of nucleotide sequences encoding the indicated residues (positions 23 and 45) demonstrate stable retention of the engineered mutations across five consecutive passages, highlighting their genetic stability even in the absence of selective antiviral pressure.

**Extended Data Fig. 6: Global sequence analysis of SARS-CoV-2 M$^{pro}$ residues 23 and 45.**

**a**




**b**

| Strain | Total count |
|---|---|
| SARS-CoV-2 | 16,842,481 |
| M$^{pro}$-T45I | 3,134 |
| M$^{pro}$-Δ23G | 2 |
| M$^{pro}$-Δ23G/T45I | 0 |

**a**, Sequence alignment of M$^{pro}$ amino acid residues at positions 23 and 45 from globally isolated SARS-CoV-2 variants (retrieved from GISAID database, as of March 30, 2025). The alignment highlights nucleotide and corresponding amino acid variations, including identified resistance-associated mutations (Δ23G and T45I). **b,** Summary table showing total counts of globally reported SARS-CoV-2 sequences containing the specific M$^{pro}$ mutations (M$^{pro}$-T45I, M$^{pro}$-Δ23G, and M$^{pro}$-Δ23G/T45I). The data emphasize the limited global occurrence of Δ23G and Δ23G/T45I mutations compared to the more frequently identified T45I mutation.