# Supporting Information: Proton-Coupled Conformational Activation of SARS Coronavirus Main Proteases and Opportunity for Designing Small-Molecule Broad-Spectrum Targeted Covalent Inhibitors

Neha Verma, Jack A. Henderson, and Jana Shen*

*Department of Pharmaceutical Sciences, University of Maryland School of Pharmacy, Baltimore, Maryland 21201*

E-mail: Jana.Shen@rx.umaryland.edu

# Supplemental Tables

Table S1: Calculated p$K_a$'s of SARS-CoV-2/CoV Mpros from the CpHMD titration simulations[a].

| Residue | | SARS-CoV-2 Mpro 6y2g(A) | SARS-CoV-2 Mpro 6y2g(B) | SARS-CoV Mpro 1uk2(A) | SARS-CoV Mpro 1uk2(B) |
|---|---|---|---|---|---|
| Cys[b] | 22 | 7.5 | 6.8 | 6.8 | 7.0 |
| | 44 | 7.0 | 9.2 | 4.2 | 5.8 |
| | 85 | ∼9.5 | ∼9.0 | neutral | neutral |
| | 117 | neutral | ∼9.2 | neutral | 8.3 |
| | 128 | ∼9.4 | ∼9.1 | neutral | neutral |
| | 145 | neutral | ∼9.4 | neutral | neutral |
| | 300 | 8.4 | neutral | neutral | ∼8.8 |
| His[c] | 41 | 6.6 | 6.7 | 6.2 | 6.5 |
| | 64 | 6.1 | 6.1 | 6.4 | 6.4 |
| | 80 | 6.3 | 6.2 | 6.5 | 6.4 |
| | 134[d] | N/A | N/A | 6.9 | 7.0 |
| | 172 | 6.6 | 6.6 | 6.6 | 7.7 |
| | 246 | ∼4.3 | ∼4.9 | 5.3 | 5.4 |

[a] The p$K_a$'s of both protomers are given. Calculations are based on the crystal structures of SARS-CoV-2 Mpro (PDB: 6y2g[S1]) and SARS-CoV Mpro (PDB: 1uk2[S2]). For residues that did not have a complete titration curves, their estimated p$K_a$'s are given and indicated with ∼. For residues that did not titrate in the entire simulation pH range, their charge states are given. The p$K_a$'s of all Asp and Glu sidechains are not listed, as most of them remained deprotonated (charged) in the entire simulation pH range and a few of them (Glu166A/B, Glu240 B, Asp263 A in SARS-CoV-2 Mpro; Glu166A and Asp216A in SARS-CoV Mpro) have estimated p$K_a$'s 5–5.8. All Lys sidechains remained protonated (charged) in the entire pH range.

[b] The p$K_a$'s of Cys16, Cys38, Cys156, Cys160, and Cys265 are not listed, as they remained protonated (neutral) in the entire pH range.

[c] The p$K_a$'s of His163 and His164 are not listed, as they remained singly protonated (neutral) in the entire pH range.

[d] SARS-CoV-2 Mpro does not have His134.

# Supplemental Figures

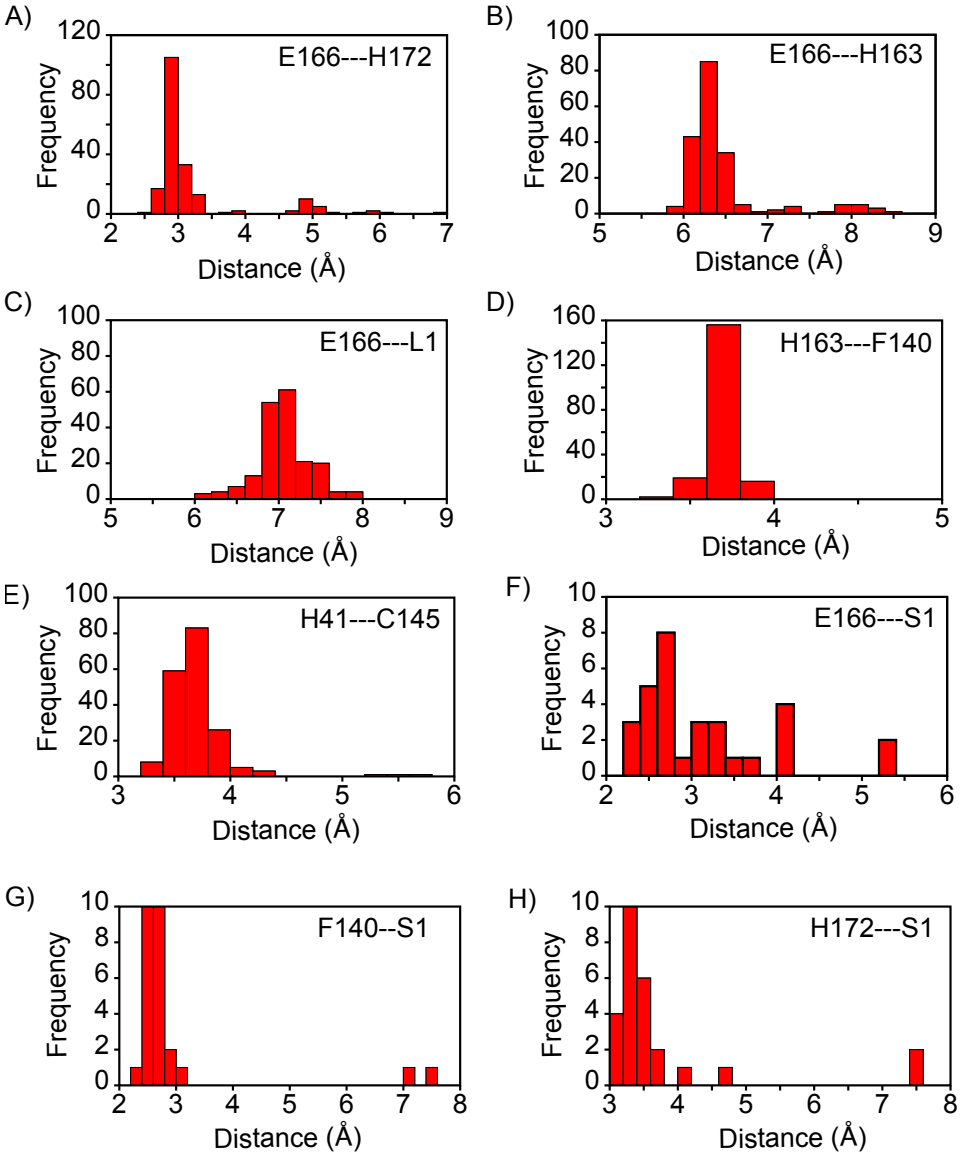


Figure S1: **Relevant distances from all available X-ray crystal structures of SARS-CoV-2 Mpro from Protein Data Bank.** A) Shortest distance between Glu166 (OE1, OE2) and His172 (NE2, ND1). The structures (PDB: 6M2N at pH 5.8, PDB: 5RGK, 5RGX, 5RGZ, 7BRP at pH 6.5 and 6WTJ, 7C8U, 7JYC at pH 7.0, 6WNP, 6YT8 at pH 7.5, 6YZ6 at pH 7.8, 6XKH, 6Y2G(B) at pH 8.5) lack the hydrogen bond between Glu166 and His172. B) Shortest distance between Glu166 (OE1/OE2) and His163 (NE2/ND1). C) Distance between the center of mass (COM) of Glu166 (OE1/OE2) and L1 loop (C$\alpha$ atoms of residues 138-145). D) Distance between the COM of imidazole ring of His163 and phenyl ring of Phe140. E) Shortest distance between the catalytic Cys145(S) and His41 (NE2/ND1). F) Shortest distance between Glu166 (OE1/OE2) and Ser1 (backbone N) of the opposite protomer. G) Distance between Phe140 (O) and the N-terminal amino nitrogen of the opposite protomer. H) Shortest distance between His172 (NE2/ND1) and Ser1 (O) of the opposite protomer. F), G), H) Calculations used the structures containing two independent protomers. The structure (PDB: 7BRP, pH 6.5) lacks the hydrogen bond between Glu166 and Ser1. In another structure (PDB: 6XBI, pH 5.6), Ser1 lacks all interactions with the S1 binding pocket residues (Glu166, His172, and Phe140). The structure (PDB 6XBI, pH 5.6) contains two extra residues before Ser1.

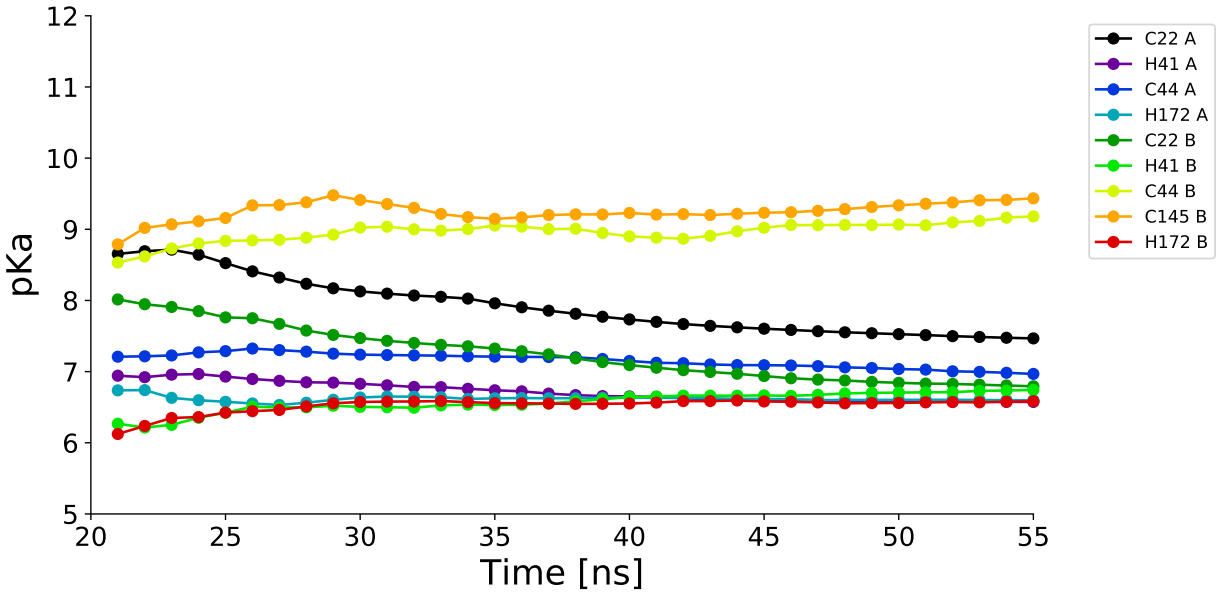


Figure S2: Time series of the calculated p$K_{\mathrm{a}}$ values of the relevant Cys and His residues in SARS-CoV-2 Mpro. Indicated time is for each replica, and data from the first 20 ns (per replica) was discarded. The p$K_{\mathrm{a}}$'s were calculated at every 1 ns using the cumulative data.

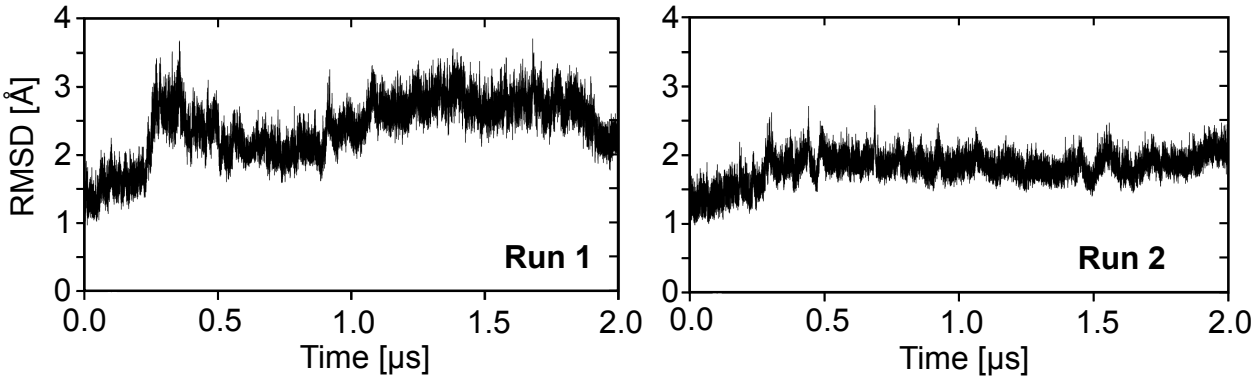


Figure S3: Time course of the C$\alpha$ root mean square deviation (RMSD) of SARS-CoV-2 Mpro dimer from the starting structure of the production run. Two simulation runs are shown.

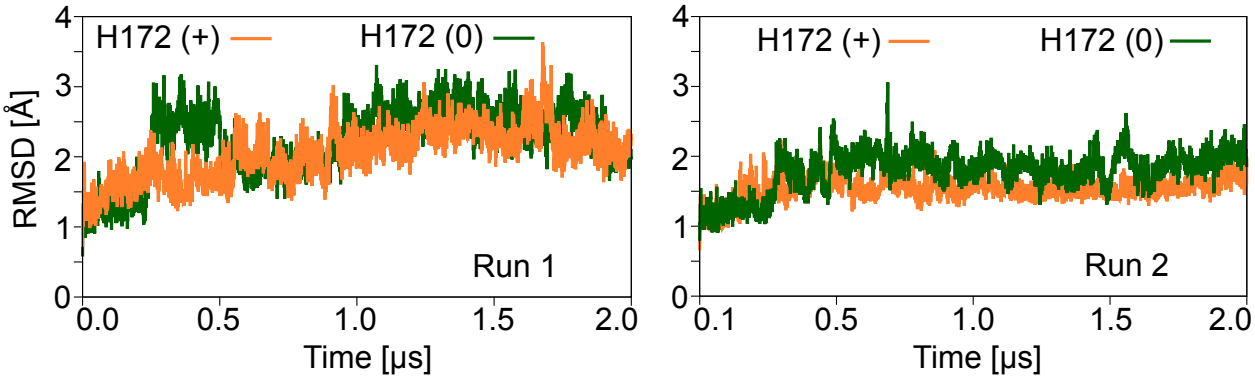


Figure S4: Time course of the C$\alpha$ RMSD of the individual protomer of SARS-CoV-2 Mpro dimer from the starting structure of the production run. Data for the protomer with neutral and charged His172 are shown in green and orange, respectively. Two simulation runs are shown.

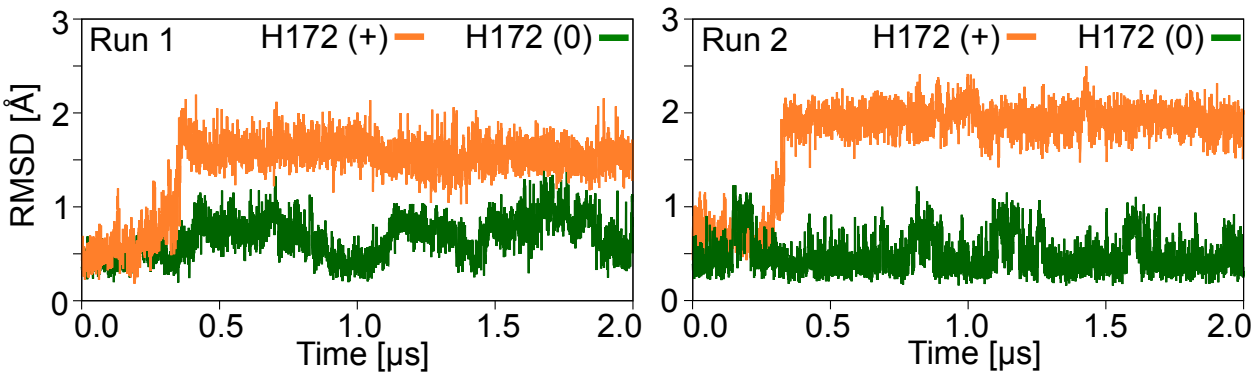


Figure S5: Time course of the C$\alpha$ RMSD of the oxyanion loop (L1, residues 138-145) in SARS-CoV-2 Mpro dimer from the starting structure of the production run. Data for the protomer with neutral and charged His172 are shown in green and orange, respectively. Two simulation runs are shown.

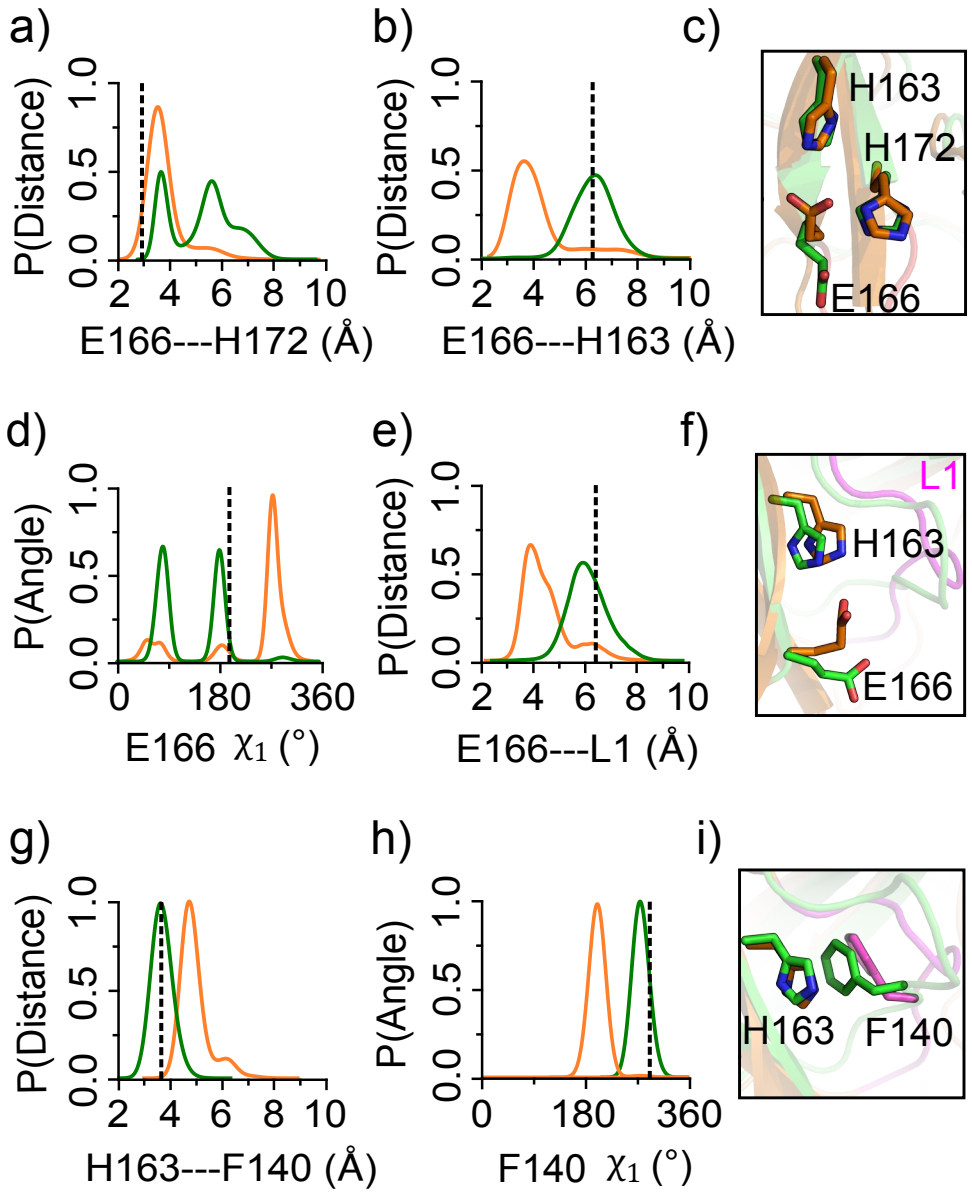


Figure S6: **Conformational changes in the S1 pocket of SARS-CoV-2 Mpro dimer in the simulation run 2.** a), b) Probability distribution of the minimum distance between carboxylate oxygens of Glu166 and imidazole nitrogens of His172 (a) and His163 (b). d) Distribution of the $\chi_1$ angle of Glu166. e) Distribution of the distance between the center-of-mass (COM) of Glu166 (carboxylate oxygens) and the oxyanion loop (C$\alpha$ atoms of residue 138–145). g) Distribution of the distance between the center of mass of the aromatic rings of His163 and Phe140. h) Distribution of the $\chi_1$ angle of Phe140. Data for the protomer with the neutral and charged and His172 are colored green and orange, respectively. All calculations were based on the last 1-$\mu$s trajectory. The dashed vertical lines indicate the distances and angles in the X-ray structure PDB 6Y2G (protomer A). c), f) and i) Snapshots showing the conformational differences between the two protomers with neutral (green) and charged (orange) His172. The oxyanion loop is colored magenta.

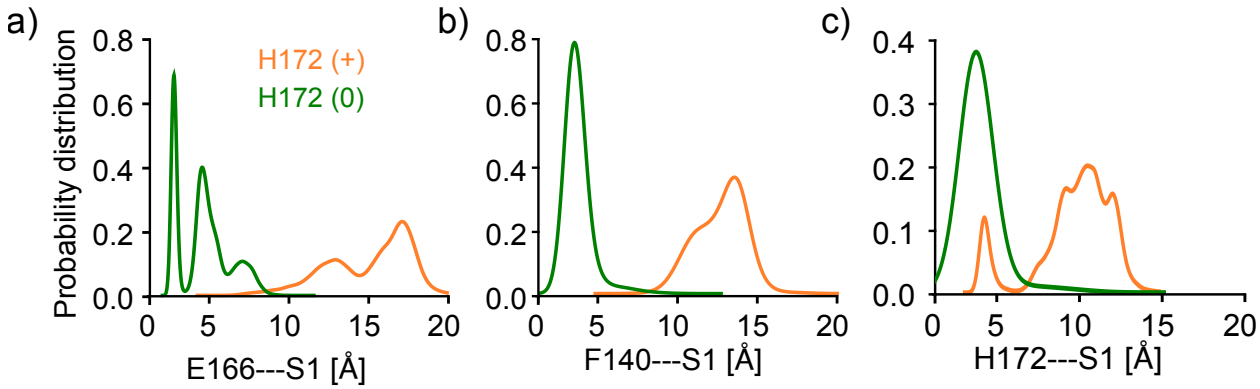


Figure S7: **N-finger interactions are disrupted in the simulation run 2 of SARS-CoV-2 Mpro dimer.** a) Probability distribution of the distance between Glu166$^{A}$ and Ser1$^{B}$ (orange) and between Glu166$^{B}$ and Ser1$^{A}$ (green). The distance is measured between the N-terminal amino nitrogen of Ser1 and the nearest carboxyl oxygen of Glu166. b) Probability distribution of the distance between Phe140$^{A}$ and Ser1$^{B}$ (orange) and between Phe140$^{B}$ and Ser1$^{A}$ (green). The distance is measured between the backbone carbonyl oxygen of Phe140 and the N-terminal amino nitrogen of Ser1. c) Probability distribution of the distance between His172$^{A}$ and Ser1$^{B}$ (orange) and between His172$^{B}$ and Ser1$^{A}$ (green). The distance is measured between the nearest imidazole nitrogen of His172 and the backbone carbonyl oxygen of Ser1. His172 is charged in protomer A and neutral in protomer B.

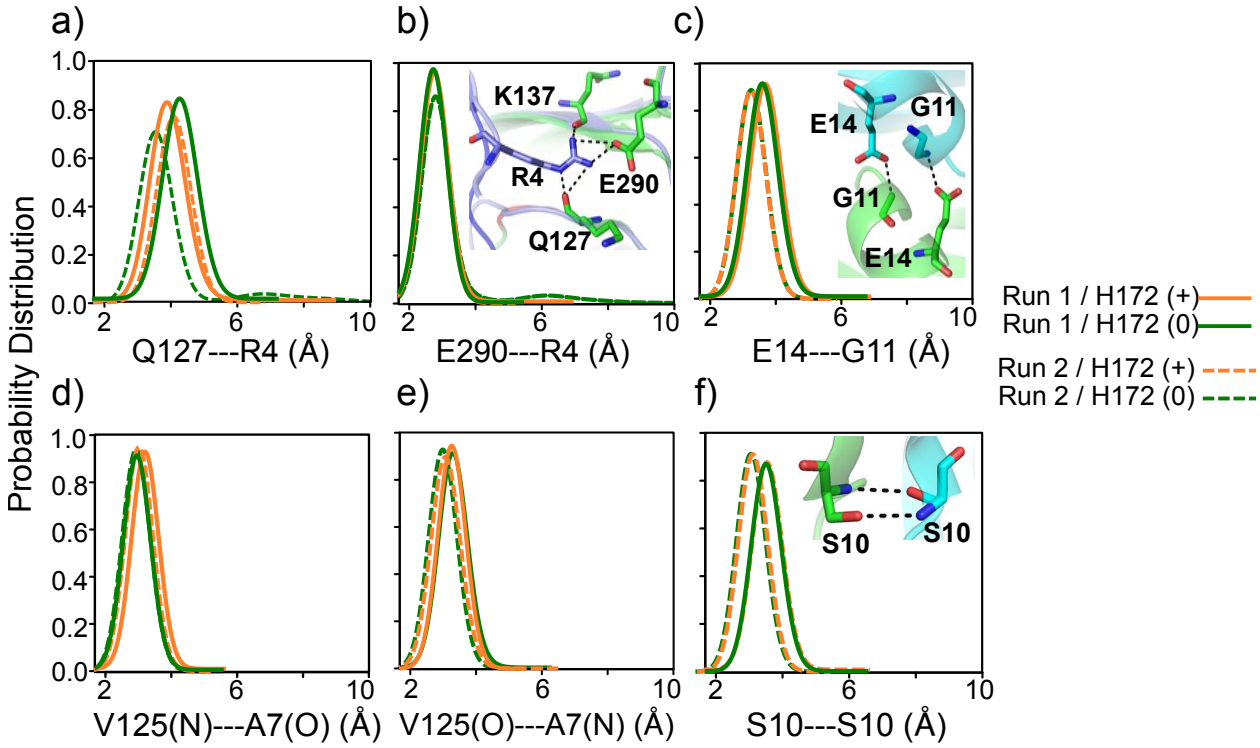


Figure S8: **Important dimer interface interactions in the two simulation runs of SARS-CoV-2 Mpro.** a) Probability distribution of the minimum distance between Gln127$^{A/B}$(backbone carbonyl oxygen) and Arg4$^{B/A}$(guanidinium nitrogen). b) Distribution of the minimum distance between Glu290$^{A/B}$ (carboxylate oxygens) and Arg4$^{B/A}$ (guanidinium nitrogens) c) Distribution of the distance between Glu14$^{A/B}$ (carboxylate oxygens) and Gly11$^{B/A}$ (backbone amide nitrogen). d) Distribution of the distance between Val125$^{A/B}$ (backbone amide nitrogen) and Ala7$^{B/A}$ (backbone carbonyl oxygen). e) Distribution of the distance between Val125$^{A/B}$ (backbone carbonyl oxygen) and Ala7$^{B/A}$ (backbone amide nitrogen). f) Probability distribution of the distance between Ser10$^{A/B}$ (backbone carbonyl oxygen) and Ser$^{B/A}$ (backbone amide nitrogen). Data from run 1 and run 2 are represented by solid and dashed lines, respectively. In run 1, His172 is neutral in protomer A (green) and charged in protomer B (orange). In run 2, His172 is charged in protomer A (orange) and neutral in protomer B (green).

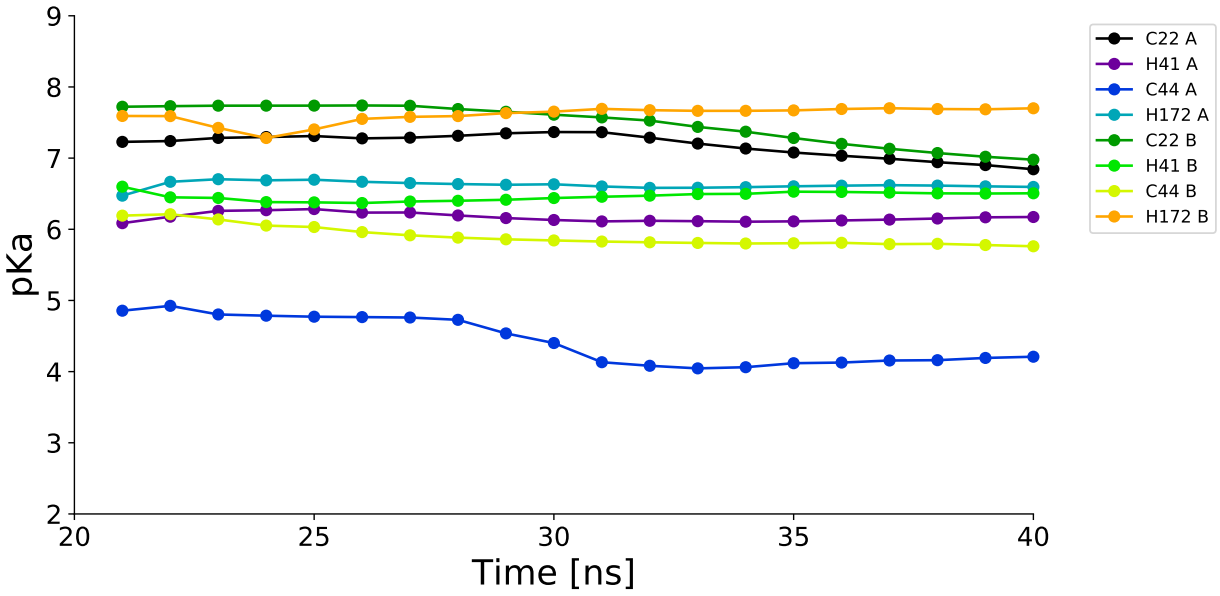


Figure S9: Time series of the calculated p$K_{\mathrm{a}}$ values of the relevant Cys and His residues in SARS-CoV Mpro. Indicated time is for each replica, and data from the first 20 ns (per replica) was discarded. The p$K_{\mathrm{a}}$'s were calculated at every 1 ns using the cumulative data.

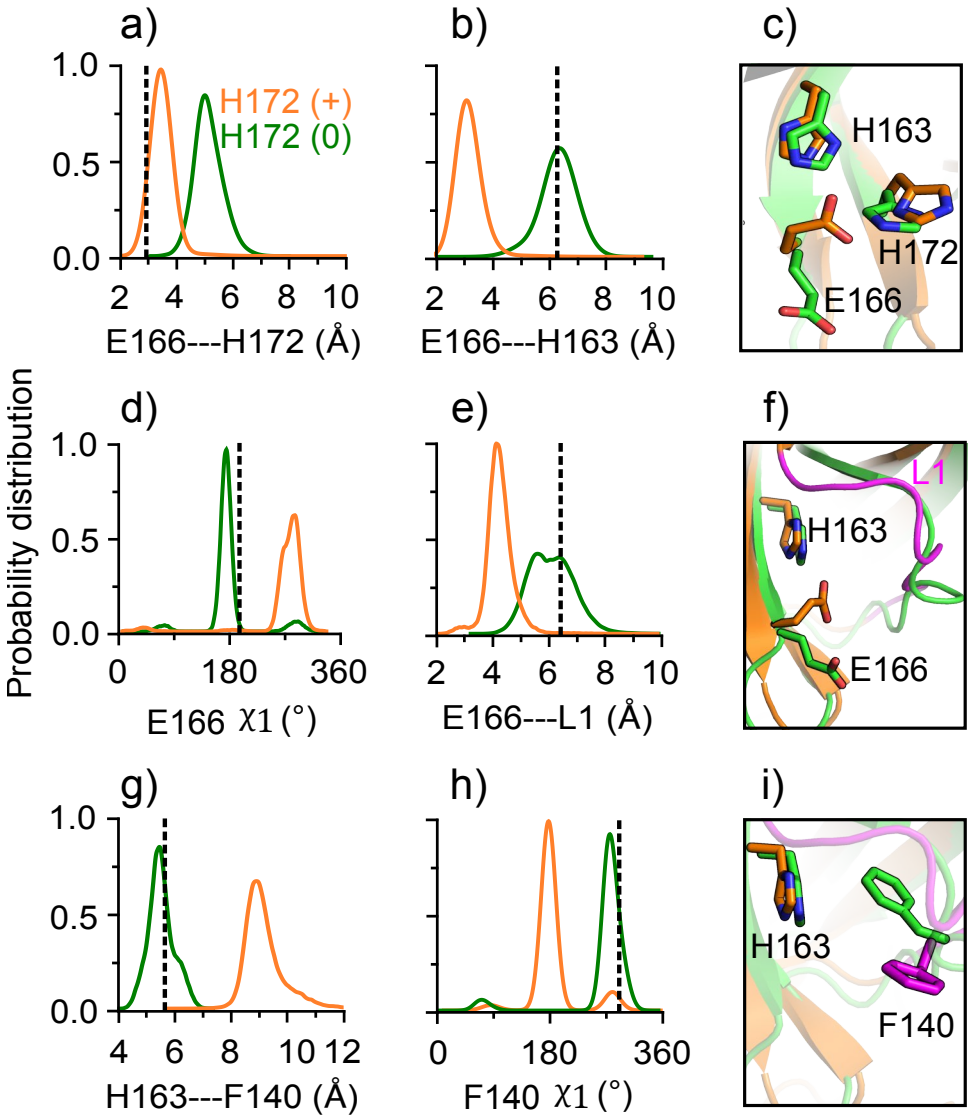


Figure S10: **Conformational changes in the S1 pocket of SARS-CoV Mpro.** a), b) Probability distribution of the minimum distance between carboxylate oxygens of Glu166 and imidazole nitrogens of His172 (a) and His163 (b). d) Distribution of the $\chi_1$ angle of Glu166. e) Distribution of the distance between the center-of-mass (COM) of Glu166 (carboxylate oxygens) and the oxyanion loop (C$\alpha$ atoms of residue 138–145). g) Distribution of the distance between the center of mass of the aromatic rings of His163 and Phe140. h) Distribution of the $\chi_1$ angle of Phe140. Data for the protomer with the neutral and charged and His172 are colored green and orange, respectively. All calculations were based on the last 1-$\mu$s trajectory. The dashed vertical lines indicate the distances and angles in the X-ray structure PDB 6Y2G (protomer A). c), f) and i) Snapshots showing the conformational differences between the two protomers with neutral (green) and charged (orange) His172. The oxyanion loop is colored magenta.

# References

(S1) Zhang, L.; Lin, D.; Sun, X.; Curth, U.; Drosten, C.; Sauerhering, L.; Becker, S.; Rox, K.; Hilgenfeld, R. Crystal Structure of SARS-CoV-2 Main Protease Provides a Basis for Design of Improved $\alpha$-Ketoamide Inhibitors. *Science* **2020**, *368*, 409–412.

(S2) Yang, H.; Yang, M.; Ding, Y.; Liu, Y.; Lou, Z.; Zhou, Z.; Sun, L.; Mo, L.; Ye, S.; Pang, H.; Gao, G. F.; Anand, K.; Bartlam, M.; Hilgenfeld, R.; Rao, Z. The Crystal Structures of Severe Acute Respiratory Syndrome Virus Main Protease and Its Complex with an Inhibitor. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 13190–13195.