Biophysics and Physicobiology

Note

Microsecond molecular dynamics suggest that a non-synonymous mutation, frequently observed in patients with mild symptoms in Tokyo, alters dynamics of the SARS-CoV-2 main protease

Daisuke Kuroda[1,2,3] and Kouhei Tsumoto[1,2,3,4]

[1] *Department of Bioengineering, School of Engineering, The University of Tokyo, Tokyo 108-8639, Japan*
[2] *Medical Device Development and Regulation Research Center, School of Engineering, The University of Tokyo, Tokyo 108-8639, Japan*
[3] *Department of Chemistry and Biotechnology, School of Engineering, The University of Tokyo, Tokyo 113-8656, Japan*
[4] *Laboratory of Medical Proteomics, The Institute of Medical Science, The University of Tokyo, Tokyo 108-8639, Japan*

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which causes the coronavirus disease 2019 (COVID-19), spread rapidly around the globe. The main protease encoded by SARS-CoV-2 is essential for processing of the polyproteins translated from the viral RNA genome, making this protein a potential drug target. A recently reported mutation in the protease, P108S, may be responsible for milder symptoms observed in COVID-19 patients in Tokyo. Starting from a crystal structure of the SARS-CoV-2 main protease in the dimeric form, we performed triplicate 5.0-µs molecular dynamics simulations of the wild-type and P108S mutant. Our computational results suggest a link between the mutation P108S and dynamics of the catalytic sites in the main protease: The catalytic dyad become considerably inaccessible to substrates in the P108S mutant. Our results also demonstrate the potential of molecular dynamics simulations to complement experimental techniques and other computational methods, such as protein design calculations, which predict effects of mutations based on static crystal structures. Further studies are certainly necessary to quantitatively understand the relationships between the P108S mutation and physical properties of the main protease, but the results of our study will immediately inform development of new protease inhibitors.

Key words: virus, molecular dynamics simulation

Corresponding author: Kouhei Tsumoto, Medical Proteomics Laboratory, Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan. e-mail: tsumoto@bioeng.t.u-tokyo.ac.jp

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the causative agent of the global pandemic of coronavirus disease 2019 (COVID-19) [1]. SARS-CoV-2 is an RNA virus, and, despite a genetic

◀ *Significance* ▶

Based on a crystal structure of the SARS-CoV-2 main protease, we performed triplicate 5.0-µs molecular dynamics simulations of the wild-type and mutant P108S protease. This mutation might cause milder symptom observed in COVID-19 patients in Tokyo. Our simulations suggest that the mutation makes substrates less accessible to the catalytic dyad of the protease through rigidification of a loop structure at the entrance of the catalytic sites, which may then lead to milder manifestation of COVID-19. Although further experimental characterization is necessary, this is the first report that links the mutation P108S to the dynamics of the dimeric SARS-CoV-2 main protease.

proofreading mechanism [2] that limits genetic diversity [3], the persistence of the pandemic may allow for selection of rare but favorable mutations [4].

Since the first report of the virus [5], numerous papers have been published in peer-reviewed journals [6] that detail results of medical and biochemical studies [7,8] and structural studies [9,10]. A potential target for therapy is the SARS-CoV-2 main 3C-like protease, which is essential for processing of polyproteins translated from the viral genome [11]. Previous computational studies of the main protease focused on interactions between the protease and potential inhibitors [12–18]. One of the most popular computational methods for analysis of interactions between drugs and proteins is molecular dynamics (MD) simulation, which is frequently employed to complement experimental measurements [19–26]. For instance, Suarez and Diaz characterized dynamics of the main protease in the dimeric and monomeric forms with and without a peptide substrate, respectively, through 2.3-μs MD simulations, revealing instability of the unbound-state of the main protease [27]. Similarly, Grottesi *et al*. conducted 2.3-μs MD simulations of the main protease in the monomeric and dimeric forms without substrates, quantifying conformational dynamics of loop regions near the catalytic site [28]. These computational studies provided insight into dynamic characteristics of the main protease, which could not be obtained by visual inspection of static crystal structures.

The mortality rates due to COVID-19 appear to be lower in Japan than in Western countries (https://covid19.who.int). Recently, based on statistical analyses of viral genome sequences derived from COVID-19 patients hospitalized in Tokyo, Abe *et al*. hypothesized that the accumulation of mutations may have contributed to the decrease in clinical virulence [29]. Their phylogenetic analysis identified four non-synonymous mutations in the viral genome sequence inversely correlated with COVID-19 severity, and subsequent sequence comparison and structure-based prediction suggested that two mutations, P108S and P151L, in the main protease and the nucleocapsid protein, respectively, contribute to a milder clinical course of COVID-19. The authors further demonstrated *in vitro* that the P108S mutation reduced the enzymatic activity.

In this study, we employed triplicate 5.0-μs MD simulations to investigate effects of the non-synonymous mutation, P108S, on dynamics of the SARS-CoV-2 main protease. The main protease consists of three structural domains (Fig. 1). The catalytic dyad (His41 and Cys145) is formed by association of domains I (10–99) and II (100–182), whereas the domain III (198–303) is important for the self-dimerization [30]. The mutation P108S is located in domain II. In a previous study, Amamuddy *et al*. analyzed dynamics of the wild-type (WT) main protease and 50 natural variants, including P108S, in the dimeric

form through a single 100-ns MD simulation of each mutant [31]. They used various metrics, such as root mean square deviations (RMSDs), residue fluctuations, and geometrical calculations, to suggest structural variations caused by mutations. Based on visual inspection, the authors argued that mutations located in a hinge region that connects rigid and flexible regions, such as the P108S mutation, control regulatory motions of the main protease. However, the impacts of the P108S mutation on dynamics and substrate recognition of the main protease were unclear.

In this study, our computational results demonstrated that the systems studied are stabilized only after 4.0 μs of simulation, after which different dynamics between the WT protease and the P108S mutant emerged. We quantified the difference in dynamics based on 5.0-μs MD trajectories and discuss possible effects of the mutation on the substrate binding. Although further studies are certainly necessary, our results suggest that the symptoms caused by COVID-19 may become milder due to emergent rigidification of a loop structure at the entrance of the catalytic site of the main protease, as is caused by the mutation at position 108.

Material and Methods

Preparation of initial structure for simulations

The initial structure for simulations of the main protease was retrieved from a previously determined crystal structure [PDB ID: 7BRO] [32]. Missing residues and the mutation P108S were generated from the WT structure with Modeller [33]. The protonation states of histidine residues were assigned using a structure-based prediction method, PROPKA [34], through the PDB2PQR website [35], with a crystal structure of the WT protease. Thus, the protonation states of histidine residues were the same in the WT and the mutant in this study. Protein structures were visualized with UCSF Chimera [36].
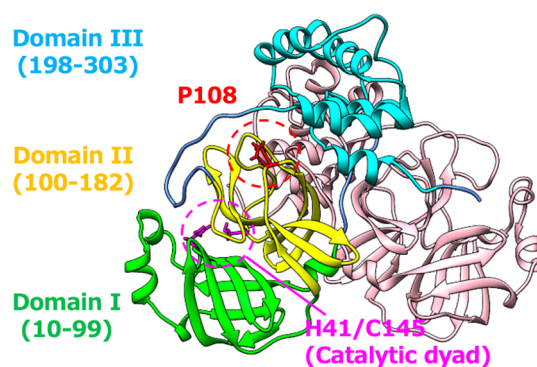

Figure 1   Structure of the SARS-CoV-2 main protease and the position of the P108S mutation (PDB ID: 7BRO) [32]. Protein structures were generated with UCSF Chimera [36].

Molecular dynamics simulations

MD simulations of the WT main protease and the P108S mutant were performed using GROMACS 2018.6 [37] with the CHARMM36m force field [38]. The structures were solvated with TIP3P water [39] in a rectangular box such that the minimum distance to the edge of the box was 15 Å under periodic boundary conditions. The protein charge was neutralized with added Na^+ or Cl^-, and additional ions were added to imitate a salt solution of concentration 0.15 M. Each system was energy-minimized by the steepest descent algorithm (5,000 steps) and was subsequently heated from 50 K to 298 K during 150 ps. The simulations were continued by 200 ps with NVT ensemble and further by 300 ps with NPT ensemble; during these steps, protein atoms were held fixed, whereas non-protein atoms freely moved. Further simulations were performed with NPT ensemble at 298 K for 5.0 µs without any restraints other than the LINCS algorithm to constrain bonds involving hydrogen atoms [40]. The time step was set to 2 fs throughout the simulations. A cutoff distance of 12 Å was used for Coulomb and van der Waals interactions. Long-range electrostatic interactions were evaluated by means of the particle mesh Ewald method [41]. A snapshot was saved every 100 ps. We repeated the same calculations three times with different initial velocities for 30-µs simulation time in total.

Analysis of MD trajectories

Trajectory analysis of each 5.0-µs simulation was performed based on the last 0.5 µs. RMSDs and distances between atoms were computed with the GROMACS package [37]. Cross-correlation analysis and principal component analysis (PCA) were performed for trajectories, in which a snapshot was saved every 500 ps, with the Prody Python library [42]. After removal of the rotational and translational motions, a positional covariance matrix of Cα-atoms of the whole protease was calculated and diagonalized to obtain the PCs. To remove random fluctuations at the terminal regions, we removed the first and last 10 residues, respectively, of each chain for PCA. PCs were visualized through Normal Mode Wizard in VMD [42,43]. Fraction of native contacts or Q-values were computed with the MDTraj Python library [44] using the equation below [45]:

$$Q(X) = \frac{1}{N} \sum_{(i,j) \in s} \frac{1}{1 + \exp[\beta(r_{ij}(X) - \lambda r_{ij}^0)]}$$

where the sum runs over the N pairs of native contacts (i,j), $r_{ij}(X)$ is the distance between i and j in configuration X, r_{ij}^0 is the distance between i and j in the native state, β is a smoothing parameter taken to be 5 Å^{-1}, and the factor λ accounts for fluctuations when the contact is formed, taken to be 1.8 for the all-atom simulations [45].
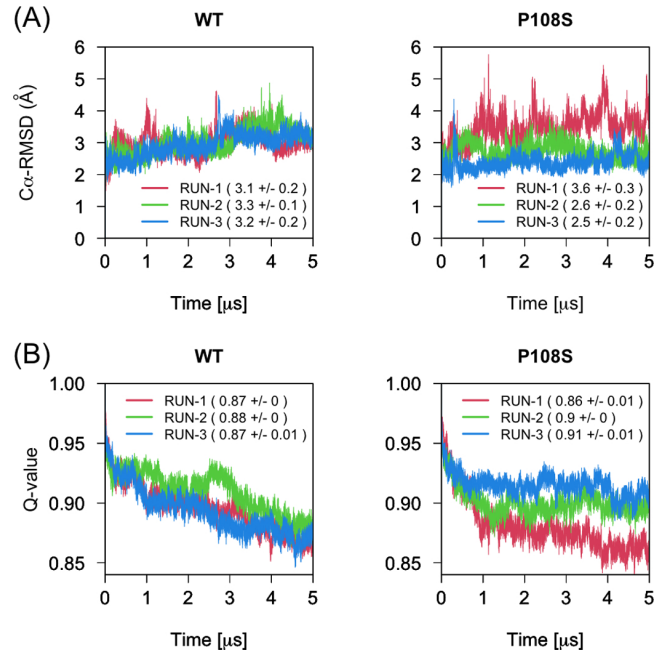

Figure 2   Convergence of the MD trajectories. (A) RMSDs of Cα atoms in the main protease and (B) Q-values as function of simulation time for simulations with three different initial velocities for each protein.

Results

The P108S mutation suppresses plasticity of catalytic dyad

We began by calculating Cα-RMSDs of the protease to evaluate convergence of our simulations (Fig. 2A). The standard deviations of the Cα-RMSDs converged within 0.5 Å during the simulations. However, when the fraction of the native contacts was computed based on the crystal structure as the reference, we found that the Q-value decreased until about 4.0 µs (Fig. 2B). Although the previous work of mutant protease relied on a 100-ns simulation time [31], our Q-value calculations implied that µs-scale simulations were necessary to obtain insight into dynamics of the main protease. Therefore, we decided to use the trajectories from the last 0.5 µs of our 5.0-µs simulations for subsequent analyses. In addition, one of the mutant simulations (RUN-1) appeared to behave differently from the other two simulations (Fig. 2), suggesting that the simulation was trapped in a different state. Therefore, our discussion of the mutant analyses below is mainly based on the simulations of the other converged trajectories.

To examine more specific dynamics, we computed Cα-RMSDs of each domain. Although the main protease is a dimeric protein consisting of chains A and B, for brevity we report here the results obtained for chain A unless otherwise noted. Results obtained from analysis of chain B are provided in the Supplementary Figures S1 and S2.
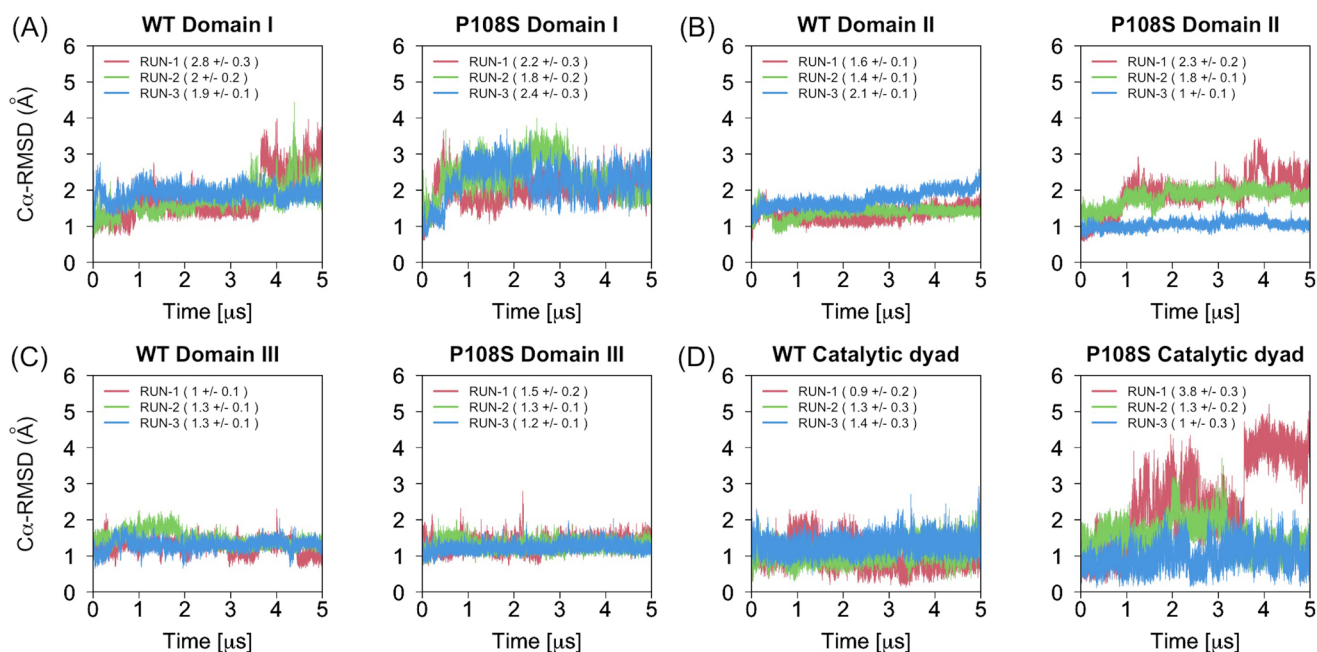
Figure 3 RMSDs of Cα atoms in the chain A main protease (A) domain I (amino acids 10–99), (B) domain II (amino acids 100–182), (C) domain III (amino acids 198–303), and (D) catalytic residues (41His/145Cys) of WT and P108S mutant as function of simulation time.

Since we performed the simulations of each protein three times with different initial velocities, we observed some variations of the results even in the same system. There were more variations in Cα-RMSDs of domains I and II (Figs. 3A and 3B) than in Cα-RMSDs of domain III, which was quite stable (Fig. 3C). The catalytic residues, 41His and 145Cys, are located in domain I and in domain II, respectively. Therefore, we also computed Cα-RMSDs of the catalytic residues. When we superposed Cα-atoms of the domains I and II and computed Cα-RMSDs of the catalytic residues, we observed larger variation after 1.0 μs in the P108S catalytic residues than in those of the WT protease (Fig. 3D). Interestingly, the variation became small at around 4.0 μs, and two of the mutant simulations appeared to reach stable states. Therefore, it is tempting to speculate that the mutation P108S in the main protease, frequently observed in patients with mild symptoms in Tokyo, affects dynamics of the catalytic residues, thereby causing milder symptoms.

To further quantify the difference in dynamics between the WT and P108S mutant, we computed the distances between center of gravity of the catalytic residues. The distance fluctuated along the simulation time with clear distinctions between the WT and the P108S mutant (Fig. 4). In the simulations of the WT protease, distributions of the distances between the catalytic dyad were wider in range (Fig. 4C), than observed in the simulations of the mutant (Fig. 4D). Although there are some overlaps of the distributions among three trajectories
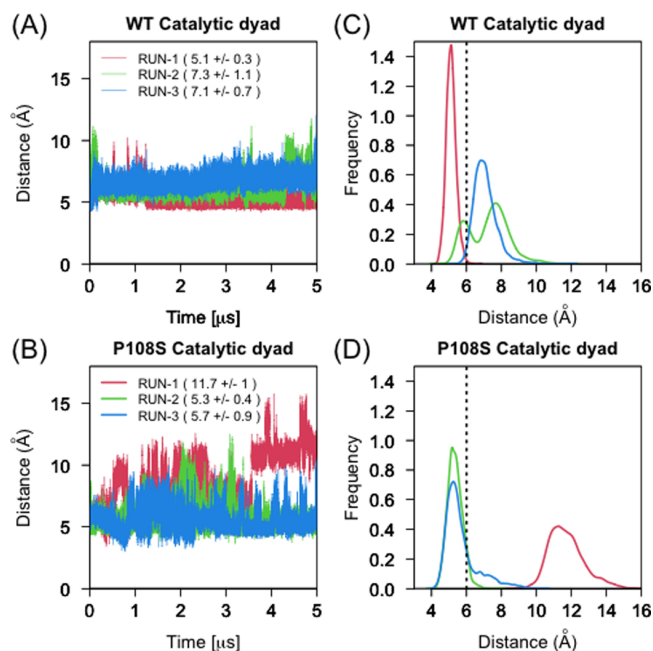

Figure 4 Distance between the catalytic residues is less variable in the WT protease than in the P108S mutant. (A and B) Distance of the center of gravity of side chain atoms between the catalytic residues (41His and 145Cys) in A) WT and B) P108S mutant in the main protease (chain A) as a function of time. (C and D) Frequency at indicated distance for C) WT and D) P108S mutant. Vertical dotted line corresponds to average distance observed in the simulations of the WT.
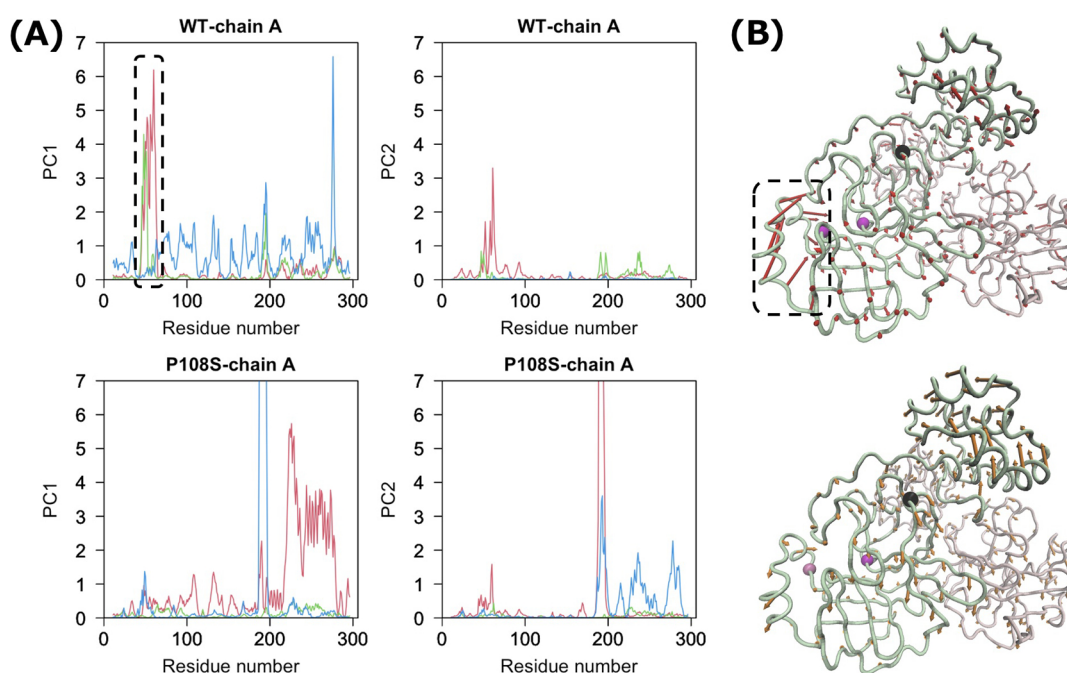
Figure 5   Principal component analysis of protein ensembles generated by MD simulations. (A) First and second principal components along the protein chain. Red, green, and blue lines correspond to the PCs of RUN-1, RUN-2, and RUN-3 of the simulations, respectively. (B) First PC visualized through Normal Mode Wizard in VMD [42,43].

of the WT protease, there appeared to be positional variability of the catalytic residues. On the other hand, in two of the three mutant simulations, the distributions almost entirely overlapped. These observations suggest that the catalytic dyad need to show plasticity to recognize and catalyze substrates.

Principal component analysis suggests rigidification of a loop structure at the entrance of the catalytic site

To further quantify the effect of the P108S mutation on the catalytic dyad, we performed PCA of the MD trajectories based on Cα atom of the protein ensembles. In the WT simulations, the first PC correspond to the motion of a loop structure (44–53) at the entrance of the catalytic site either in the chain A or the chain B (Fig. 5 and Supplementary Fig. S3). On the other hand, in the mutant simulations, the fluctuation of the corresponding loop region was much smaller, suggesting rigidification of the loop structure upon the mutation. Therefore, it is most likely that substrates could not access the catalytic dyad due to steric hindrance of the rigidified loop region. In addition, the first and second PCs of the mutant simulations tend to be the motions of domain III (198–303) (Fig. 5 and Supplementary Fig. S3).

To more directly examine correlations of motions between the P108S mutation and other residues, we also evaluated cross-correlation of residues in the MD trajectories. When residue-wise correlations were evaluated,

fluctuations of the residue at the position 108 were not correlated well with fluctuations of the catalytic residues (correlations <0.5, Supplementary Fig. S4). However, the fluctuations of the residue at the position 108, which is in domain II, were correlated with many residues in the domains I and III and even in the other chain (Supplementary Fig. S4). In addition, motions of catalytic residues appeared to be correlated with each other, and with many other residues along the chains (Supplementary Figs. S5 and S6). These results support our hypothesis that the mutation at position 108 affects dynamics of the entire protein. Together with the fact that domain III contributes to the self-association of the two chains in the main protease [30] and the major PCs of the mutant simulations correspond to motions in domain III, lower residue correlations in domain III of the mutant relative to the WT protease may be one of the reasons that COVID-19 patients infected by virus with the mutation P108S have milder symptoms.

Discussion and Conclusion

Based on microsecond-scale MD simulations, we identified a possible link between a non-synonymous mutation, P108S, frequently observed in patients with mild COVID-19 symptoms in Tokyo, and dynamics of the catalytic site in the main protease. With only data from static crystal structures, effects of mutations on dynamics of

protein structures are difficult to infer. Further, our finding that the catalytic dyad became less accessible upon the P108S mutation were visible only after simulations longer than 4 μs.

It is worth noting that Abe *et al.* [29] conducted hydrogen deuterium exchange mass spectroscopy (HDX-MS) to examine difference in dynamics between WT and the P108S mutant protease. They reported the structural perturbations of two regions, 128–141 and 161–176. On the other hand, the rigidified loop region we identified in this study is the region of 44–53. This difference can be rationalized by the fact that the HDX-MS was performed for the monomeric state of the protease, whereas our simulations were based on the dimeric state; dynamics will differ between the monomeric and dimeric states. In addition, the reported reduction of the catalytic activity was mainly due to the increase of K_m, which became twice upon the mutation while K_{cat} decrease only a little [29]. Dominant effects of K_m suggest that the P108S mutation reduced the enzymatic activity by interfering substrate binding rather than preventing release of the products, in agreement with our results that the loop region at the entrance of the catalytic site stiffened, making the catalytic dyad less accessible.

There have been many studies of drug binding to the main protease of SARS-CoV-2. Considering the differences in flexibility of the loop region at the entrance of the catalytic sites we observed between the WT and P108S mutant, the mechanism of drug binding might be different in the mutant from that of the WT main protease. Therefore, considering the prevalence of the mutation in patients, drug binding to the P108S mutant protease should be studied. It is also worth mentioning that our calculations dealt only with the P108S mutation, and we did not evaluate other mutations, which were predicted to be functionally neutral, found in patients in a previous study [46]. Together with P108S, those mutations may also affect physicochemical properties of the main protease. Further analyses could be done, for instance, by using physicochemical measurements to experimentally characterize the mutant and by using hybrid quantum mechanics/molecular mechanics simulations to understand catalytic mechanism of the mutant, as has been done with the WT protease [47].

Our results also suggest the importance of repeating MD simulations and of using μs-scale simulation times. Here, we repeated the same calculations three times with different initial velocities for each system, and the results of our simulations were highly variable even in the same system. Therefore, caution should be taken when conclusions are drawn from single trajectories of MD simulations, even when the results seem plausible. Further, Q-values in our simulations decreased until about 4.0 μs indicating that ns-scale simulations may not provide insight into protein dynamics. We also demonstrated the potential of MD simulations to complement experimental techniques as well as the other computational methods, such as protein design calculations [48–51], where effects of mutations are predicted based on static crystal structures with side chain rotamer sampling. Further studies are certainly necessary to quantitively understand relationships between the mutation P108S and physical properties of the main protease, but our findings will immediately inform design of new protease inhibitors.

Acknowledgements

Conflict of Interests

None declared.

Author Contributions

K.T. conceived the study. D.K. designed and conducted the computational study, analyzed the data, and wrote the manuscript. All the authors approved the final version of the manuscript.

References

[1] Johns Hopkins Coronavirus Resource Center. COVID-19 Map. In: Johns Hopkins Coronavirus Resource Center-Johns Hopkins University & Medicine [Internet]. 2020. Available: https://coronavirus.jhu.edu/map.html

[2] Sevajol, M., Subissi, L., Decroly, E., Canard, B. & Imbert, I. Insights into RNA synthesis, capping, and proofreading mechanisms of SARS-coronavirus. *Virus Res.* **194**, 90–99 (2014). DOI: 10.1016/j.virusres.2014.10.008

[3] Fauver, J. R., Petrone, M. E., Hodcroft, E. B., Shioda, K., Ehrlich, H. Y., Watts, A. G., *et al.* Coast-to-Coast Spread of SARS-CoV-2 during the Early Epidemic in the United States. *Cell* **181**, 990–996.e5 (2020). DOI: 10.1016/j.cell.2020.04.021

[4] Korber, B., Fischer, W. M., Gnanakaran, S., Yoon, H., Theiler, J., Abfalterer, W., *et al.* Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell* **182**, 812–827.e19 (2020). DOI: 10.1016/j.cell.2020.06.043

[5] Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J.,

et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N. Engl. J. Med.* **382**, 727–733 (2020). DOI: 10.1056/NEJMoa2001017

[6] Brainard, J. Scientists are drowning in COVID-19 papers. Can new tools keep them afloat? *Science* (2020). DOI: 10.1126/science.abc7839

[7] Chen, G., Wu, D., Guo, W., Cao, Y., Huang, D., Wang, H., *et al.* Clinical and immunological features of severe and moderate coronavirus disease 2019. *J. Clin. Invest.* **130**, 2620–2629 (2020). DOI: 10.1172/JCI137244

[8] Hu, B., Guo, H., Zhou, P. & Shi, Z.-L. Characteristics of SARS-CoV-2 and COVID-19. *Nat. Rev. Microbiol.* **19**, 141–154, 2020. DOI: 10.1038/s41579-020-00459-7

[9] Berman, H., Henrick, K. & Nakamura, H. Announcing the worldwide Protein Data Bank. *Nat. Struct. Mol. Biol.* **10**, 980 (2003). DOI: 10.1038/nsb1203-980

[10] Subramaniam, S. COVID-19 and cryo-EM. *IUCrJ* **7**, 575–576 (2020). DOI: 10.1107/S2052252520008799

[11] Gordon, D. E., Jang, G. M., Bouhaddou, M., Xu, J., Obernier, K., White, K. M., *et al.* A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**, 459–468 (2020). DOI: 10.1038/s41586-020-2286-9

[12] Kumar, Y., Singh, H. & Patel, C. N. In silico prediction of potential inhibitors for the main protease of SARS-CoV-2 using molecular docking and dynamics simulation based drug-repurposing. *J. Infect. Public Health.* **13**, 1210–1223 (2020). DOI: 10.1016/j.jiph.2020.06.016

[13] Yoshino, R., Yasuo, N. & Sekijima, M. Identification of key interactions between SARS-CoV-2 main protease and inhibitor drug candidates. *Sci. Rep.* **10**, 12493 (2020). DOI: 10.1038/s41598-020-69337-9

[14] Komatsu, T. S., Okimoto, N., Koyama, Y. M., Hirano, Y., Morimoto, G., Ohno, Y., *et al.* Drug binding dynamics of the dimeric SARS-CoV-2 main protease, determined by molecular dynamics simulation. *Sci. Rep.* **10**, 16986 (2020). DOI: 10.1038/s41598-020-74099-5

[15] Macchiagodena, M., Pagliai, M. & Procacci, P. Identification of potential binders of the main protease 3CLpro of the COVID-19 via structure-based ligand design and molecular modeling. *Chem. Phys. Lett.* **750**, 137489 (2020). DOI: 10.1016/j.cplett.2020.137489

[16] Gao, K., Nguyen, D. D., Chen, J., Wang, R. & Wei, G.-W. Repositioning of 8565 Existing Drugs for COVID-19. *J. Phys. Chem. Lett.* **11**, 5373–5382 (2020). DOI: 10.1021/acs.jpclett.0c01579

[17] Choudhary, S., Malik, Y. S. & Tomar, S. Identification of SARS-CoV-2 Cell Entry Inhibitors by Drug Repurposing Using *in silico* Structure-Based Virtual Screening Approach. *Front. Immunol.* **11**, 1664 (2020). DOI: 10.3389/fimmu.2020.01664

[18] Li, Z., Li, X., Huang, Y.-Y., Wu, Y., Liu, R., Zhou, L., *et al.* Identify potent SARS-CoV-2 main protease inhibitors via accelerated free energy perturbation-based virtual screening of existing drugs. *Proc. Natl. Acad. Sci. USA* **117**, 27381–27387 (2020). DOI: 10.1073/pnas.2010470117

[19] Dror, R. O., Dirks, R. M., Grossman, J. P., Xu, H. & Shaw, D. E. Biomolecular Simulation: A Computational Microscope for Molecular Biology. *Annu. Rev. Biophys.* **41**, 429–452 (2012). DOI: 10.1146/annurev-biophys-042910-155245

[20] Kondo, H. X., Kiribayashi, R., Kuroda, D., Kohda, J., Kugimiya, A., Nakano, Y., *et al.* Effects of a remote mutation from the contact paratope on the structure of CDR-H3 in the anti-HIV neutralizing antibody PG16. *Sci. Rep.* **9**, 19840 (2019). DOI: 10.1038/s41598-019-56154-y

[21] Miyanabe, K., Akiba, H., Kuroda, D., Nakakido, M., Kusano-Arai, O., Iwanari, H., *et al.* Intramolecular H-bonds govern the recognition of a flexible peptide by an antibody. *J. Biochem.* **164**, 65–76 (2018). DOI: 10.1093/jb/mvy032

[22] Kasahara, K., Shiina, M., Higo, J., Ogata, K. & Nakamura, H. Phosphorylation of an intrinsically disordered region of Ets1 shifts a multi-modal interaction ensemble to an auto-inhibitory state. *Nucleic Acids Res.* **46**, 2243–2251 (2018). DOI: z0.1093/nar/gkx1297

[23] Sugawa, M., Okazaki, K., Kobayashi, M., Matsui, T., Hummer, G., Masaike, T., *et al.* F1-ATPase conformational cycle from simultaneous single-molecule FRET and rotation measurements. *Proc. Natl. Acad. Sci. USA* **113**, E2916–E2924 (2016). DOI: 10.1073/pnas.1524720113

[24] Ikura, T., Yonezawa, Y. & Ito, N. Mutational effects of Cys113 on structural dynamics of Pin1. *Biophys. Physicobiol.* **16**, 452–465 (2019). DOI: 10.2142/biophysico.16.0_452

[25] Kawade, R., Kuroda, D. & Tsumoto, K. How the protonation state of a phosphorylated amino acid governs molecular recognition: insights from classical molecular dynamics simulations. *FEBS Lett.* **594**, 903–912 (2020). DOI: 10.1002/1873-3468.13674

[26] Hollingsworth, S. A. & Dror, R. O. Molecular Dynamics Simulation for All. *Neuron* **99**, 1129–1143 (2018). DOI: 10.1016/j.neuron.2018.08.011

[27] Suárez, D. & Díaz, N. SARS-CoV-2 Main Protease: A Molecular Dynamics Study. *J. Chem. Inf. Model.* **60**, 5815-5831 (2020). DOI: 10.1021/acs.jcim.0c00575

[28] Grottesi, A., Bešker, N., Emerson, A., Manelfi, C., Beccari, A. R., Frigerio, F., *et al.* Computational Studies of SARS-CoV-2 3CLpro: Insights from MD Simulations. *Int. J. Mol. Sci.* **21**, 5346 (2020). DOI: 10.3390/ijms21155346

[29] Abe, K., Kabe, Y., Uchiyama, S., Iwasaki, Y. W., Ishizu, H., Uwamino, Y., *et al.* Pro108Ser mutant of SARS-CoV-2 3CL[pro] Reduces the Enzymatic Activity and Ameliorates COVID-19 Severity in Japan. *medRxiv* (2020). DOI: 10.1101/2020.11.24.20235952

[30] Shi, J., Wei, Z. & Song, J. Dissection Study on the Severe Acute Respiratory Syndrome 3C-like Protease Reveals the Critical Role of the Extra Domain in Dimerization of the Enzyme. *J. Biol. Chem.* **279**, 24765–24773 (2004). DOI: 10.1074/jbc.M311744200

[31] Sheik Amamuddy, O., Verkhivker, G. M. & Tastan Bishop, Ö. Impact of Early Pandemic Stage Mutations on Molecular Dynamics of SARS-CoV-2 M pro. *J. Chem. Inf. Model.* **60**, 5080–5102 (2020). DOI: 10.1021/acs.jcim.0c00634

[32] Fu, L., Ye, F., Feng, Y., Yu, F., Wang, Q., Wu, Y., *et al.* Both Boceprevir and GC376 efficaciously inhibit SARS-CoV-2 by targeting its main protease. *Nat. Commun.* **11**, 4417 (2020). DOI: 10.1038/s41467-020-18233-x

[33] Webb, B. & Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Protein Sci.* **86**, 2.9.1-2.9.37 (2016). DOI: 10.1002/cpps.20

[34] Olsson, M. H. M., Søndergaard, C. R., Rostkowski, M. & Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical p K a Predictions. *J. Chem. Theory Comput.* **7**, 525–537 (2011). DOI: 10.1021/ct100578z

[35] Dolinsky, T. J., Czodrowski, P., Li, H., Nielsen, J. E., Jensen, J. H., Klebe, G., *et al.* PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.* **35**, W522–W525 (2007). DOI: 10.1093/nar/gkm276

[36] Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S.,

Greenblatt, D. M., Meng, E. C., *et al.* UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004). DOI: 10.1002/jcc.20084

[37] Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015). DOI: 10.1016/j.softx.2015.06.001

[38] Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., de Groot, B. L., *et al.* CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71–73 (2017). DOI: 10.1038/nmeth.4067

[39] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983). DOI: 10.1063/1.445869

[40] Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997). DOI: 10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H

[41] Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993). DOI: 10.1063/1.464397

[42] Bakan, A., Meireles, L. M. & Bahar, I. ProDy: Protein Dynamics Inferred from Theory and Experiments. *Bioinformatics* **27**, 1575–1577 (2011). DOI: 10.1093/bioinformatics/btr168

[43] Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996). DOI: 10.1016/0263-7855(96)00018-5

[44] McGibbon, R. T., Beauchamp, K. A., Harrigan, M. P., Klein, C., Swails, J. M., Hernández, C. X., *et al.* MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophys. J.* **109**, 1528–1532 (2015). DOI: 10.1016/j.bpj.2015.08.015

[45] Best, R. B., Hummer, G. & Eaton, W. A. Native contacts determine protein folding mechanisms in atomistic simulations. *Proc. Natl. Acad. Sci. USA* **110**, 17874–17879 (2013). DOI: 10.1073/pnas.1311599110

[46] Abe, K., Kabe, Y., Uchiyama, S., Iwasaki, Y. W., Ishizu, H., Uwamino, Y., *et al.* Pro108Ser mutant of SARS-CoV-2 3CL[pro] reduces the enzymatic activity and ameliorates COVID-19 severity in Japan. *medRxiv* (2021). DOI: 10.1101/2020.11.24.20235952

[47] Świderek, K. & Moliner, V. Revealing the molecular mechanisms of proteolysis of SARS-CoV-2 M pro by QM/MM computational methods. *Chem. Sci.* **11**, 10626–10630 (2020). DOI: 10.1039/D0SC02823A

[48] Kuhlman, B. & Bradley, P. Advances in protein structure prediction and design. *Nat. Rev. Mol. Cell Biol.* **20**, 681–697 (2019). DOI: 10.1038/s41580-019-0163-x

[49] Kuroda, D. & Tsumoto, K. Engineering Stability, Viscosity, and Immunogenicity of Antibodies by Computational Design. *J. Pharm. Sci.* **109**, 1631–1651 (2020). DOI: 10.1016/j.xphs.2020.01.011

[50] Kuroda, D., Shirai, H., Jacobson, M. P. & Nakamura, H. Computer-aided antibody design. *Protein Eng. Des. Sel.* **25**, 507–521 (2012). DOI: 10.1093/protein/gzs024

[51] Leman, J. K., Weitzner, B. D., Lewis, S. M., Adolf-Bryfogle, J., Alam, N., Alford, R. F., *et al.* Macromolecular modeling and design in Rosetta: recent methods and frameworks. *Nat. Methods* **17**, 665–680 (2020). DOI: 10.1038/s41592-020-0848-2

(*Edited by Akio Kitao*)