JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JACS

Article

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

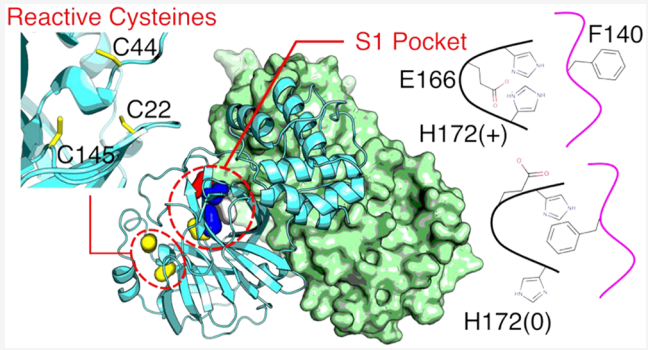
Neha Verma, Jack A. Henderson,[‡] and Jana Shen*

Cite This: *J. Am. Chem. Soc.* 2020, 142, 21883−21890

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The SARS coronavirus 2 (SARS-CoV-2) main protease (Mpro) is an attractive broad-spectrum antiviral drug target. Despite the enormous progress in structure elucidation, the Mpro's structure−function relationship remains poorly understood. Recently, a peptidomimetic inhibitor has entered clinical trial; however, small-molecule orally available antiviral drugs have yet to be developed. Intrigued by a long-standing controversy regarding the existence of an inactive state, we explored the proton-coupled dynamics of the Mpros of SARS-CoV-2 and the closely related SARS-CoV using a newly developed continuous constant pH molecular dynamics (MD) method and microsecond fixed-charge all-atom MD simulations. Our data supports a general base mechanism for Mpro's proteolytic function. The simulations revealed that protonation of His172 alters a conserved interaction network that upholds the oxyanion loop, leading to a partial collapse of the conserved S1 pocket, consistent with the first and controversial crystal structure of SARS-CoV Mpro determined at pH 6. Interestingly, a natural flavonoid binds SARS-CoV-2 Mpro in the close proximity to a conserved cysteine (Cys44), which is hyper-reactive according to the CpHMD titration. This finding offers an exciting new opportunity for small-molecule targeted covalent inhibitor design. Our work represents a first step toward the mechanistic understanding of the proton-coupled structure−dynamics−function relationship of CoV Mpros; the proposed strategy of designing small-molecule covalent inhibitors may help accelerate the development of orally available broad-spectrum antiviral drugs to stop the current pandemic and prevent future outbreaks.



INTRODUCTION

The ongoing COVID-19 pandemic has claimed more than 1.5 million human lives worldwide, but an effective therapeutic intervention has yet to be developed. COVID-19 is caused by a new coronavirus called severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2, GenBank accession code: MN908947.3[1]). Belonging to the genus *Betacoronavirus*, SARS-CoV-2 is closely related to several bat SARS-like coronaviruses, and SARS-CoV which caused an outbreak in 2002−2003, as well as MERS-CoV (Middle East Respiratory Syndrome coronavirus) which caused an outbreak in 2012.[1] The related genus *Alphacoronavirus* includes human coronaviruses that cause the common cold.[2] The replicase gene of coronaviruses encodes two polyproteins (pp1a and pp1ab), which are processed into functional nonstructural proteins to form the viral replication/transcription complex by two cysteine proteases:[3] the main protease (Mpro, also known as 3CLpro due to the similarity to picornavirus 3C protease) and the papain-like protease (PLpro).[3] The Mpro cleaves the larger polyprotein pp1ab at 11 sites, with the sequence of (small)-X-(Leu/Phe/Met)-Gln↓(Gly/Ala/Ser), where X denotes any amino acid and ↓ indicates the cleavage site. Note, Gln in the P1 position is absolutely required by the Mpro, and the substrate specificity pattern is not shared by any human protease.[3] The S1′, S1, and S2 substrate binding pockets are conserved among the Mpros of SARS-, MERS-, and SARS-like bat-CoVs.[3] Thus, the Mpro is an attractive broad-spectrum antiviral drug target for combating the COVID-19 pandemic and preventing future outbreaks through the zoonotic transmission of SARS-like bat coronaviruses to human.[3,4]

SARS-CoV-2/CoV Mpros have been well characterized by crystallography. To date, Protein Data Bank (PDB) contains 54 and 188 entries of X-ray crystal structures for SARS-CoV and SARS-CoV-2 Mpros, respectively, most of which are inhibitor-bound complexes.[5−10] The rapid structure determi-

Received: October 11, 2020
Published: December 15, 2020

https://dx.doi.org/10.1021/jacs.0c10770
J. Am. Chem. Soc. 2020, 142, 21883−21890

nation led to the discovery of the first broad-spectrum lead inhibitor against SARS, MERS, and the related coronavirus Mpros.[2,5] Most recently, Pfizer initiated a clinical trial of PF-07304814, a prodrug which metabolizes into a ketone-based peptidomimetic inhibitor PF-00835231 against the Mpros of SARS-CoV-2/CoV and other coronaviruses.[11] Despite the progress, the structure−function relationship and conformational plasticity of the Mpros remain poorly understood. Such information is urgently needed to advance broad-spectrum antiviral drug design.

SARS-CoV-2/CoV Mpros function as a dimer.[2,3] With 96% overall sequence identity and 100% identity in the active site, their structures are nearly superimposable.[5] The Mpro protomer is composed of three domains (Figure 1a). The

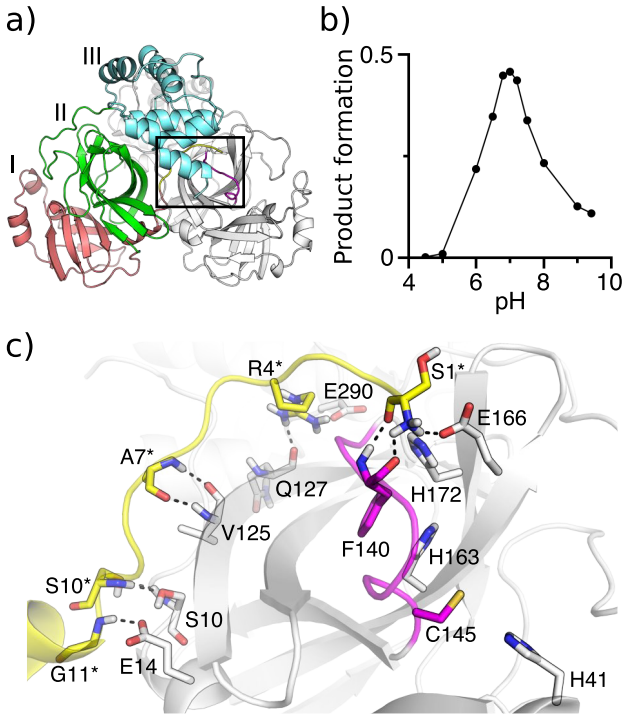


Figure 1. Crystal structure and pH-dependent activity for SARS-CoV-2/CoV Mpros. (a) X-ray crystal structure of SARS-Cov-2 Mpro dimer (PDB: 6y2g[5]). Protomer A is colored gray, while domains I, II, and III in protomer B are colored red, green, and cyan, respectively. The oxyanion loop L1 (residues 138−145) is colored magenta and the N-finger loop (residues 1−10) as well as G11 are colored yellow. (b) pH-activity profile of SARS-CoV Mpro determined by the HPLC-based peptide cleavage assay.[17] (c) Zoomed-in view of the S1 specificity pocket in protomer A and the N-finger loop in protomer B. Residues involved in the dimer interface interactions are explicitly shown (those in protomer B denoted with an asterisk). The conserved S1 pocket residues Phe140, His163, Glu166, and His172, and the catalytic dyad Cys145 and His41 are explicitly shown.

chymotrypsin-like β-barrel domains I (residues 10−99) and II (residues 100−182) embrace the Cys-His catalytic dyad at the interface, while the α-helical domain III (residues 198−303) is connected to domain II via a long loop called L3 (residues 183−197). Dimerization of Mpro is mainly supported by the domain III interactions with domain III and the N-terminal loop (residues 1−10) of the opposite protomer.[12] In particular, residues 1−7 (called N-finger[13]) form hydrogen bonds and electrostatic interactions with the S1 pocket (Figure 1c), and have been considered critical for both dimerization and active site integrity of SARS-CoV Mpro.[13] Deletion or mutation of Arg4[14] and Gly11[15] was shown to completely inactivate SARS-CoV Mpro, while deletion of residues 1−3 maintained 76% enzymatic activity.[14] Deletion of residues 1−5 in the related transmissible gastroenteritis virus Mpro nearly abolished the activity.[16]

Mpro's substrate binding site is composed of the canonical S1−S4 pockets, whereby the S1 pocket is formed by the side chains of His163, Phe140 and the backbones of Met165, Glu166, and His172, all of which are conserved (Figure 1c). In the majority of crystal structures of SARS-CoV-2 Mpro, Glu166 forms a hydrogen bond with His172, while His163 forms an aromatic stacking with Phe140 (Figure 1c and Figure S1). Additionally, the side chain of Glu166 and backbone of Phe140 form hydrogen bonds with the amino group of Ser1 (amino terminus) of the opposite protomer (Figure 1c and Figure S1). The oxyanion hole of the cysteine protease is formed by the backbone amides of Gly143, Ser144, and Cys145,[5,13] and the oxyanion loop (residues 138−145, also called L1) forms a wall of the S1 pocket (Figure 1c).

The proteolytic activity of SARS-CoV Mpro displays a bell-shaped pH profile, with the peak at pH 7.0[17] (or 7.4[18]) and sharp decreases at lower or higher pH (Figure 1b); however, the molecular mechanism of the pH dependence is not understood. Interestingly, a pH 6 crystal structure of SARS-CoV Mpro (PDB: 1uj1[13]), which captured two independent protomers in the asymmetric unit, revealed an inactive conformation for protomer B, in which the oxyanion loop is moved toward the S1 pocket and the hydrogen bond between Glu166 and His172 is broken. Hilgenfeld et al. hypothesized that the pH-dependent activity change arises from the protonation state switches of His163 and His172.[13,17] Based on the crystal structures,[13,17] His163 and His172 were suggested as neutral and charged at physiological pH, respectively, and consequently, the protonation of His163 was hypothesized to induce the aforementioned conformational changes in the S1 pocket.[17] Shortly afterward, the authenticity of the inactive conformation was questioned, by arguing that a fusion tag added to the termini from the recombinant protein expression system may have destabilized the Ser1 interactions with the S1 pocket of the opposite protomer, causing it to collapse.[19]

Recently, an X-ray structure of SARS-CoV-2 Mpro dimer (PDB: 6y2g[5]) revived the debate regarding the existence of an inactive Mpro protomer and its pH dependence. In protomer B of this crystal structure, the Glu166−His172 hydrogen bond is broken, similar to the inactive protomer of the pH 6 structure of SARS-CoV Mpro (PDB: 1UJ1[13]), although the large movement of the oxyanion loop is absent.[5] Testing the hypothesis of an inactive state by crystallography proves challenging, as most crystal structures of Mpros were determined in space group C2 which captures only one protomer in the asymmetric unit which favors a more ordered structure.

To shed light on the above controversy, we conducted a computational study of SARS-CoV-2/CoV Mpros, employing the newly developed GPU-accelerated implicit-solvent continuous constant pH MD (CpHMD)[20,21] to determine the protonation states and microsecond-long conventional fixed-charge MD simulations to characterize the detailed conformational changes. Our data confirms the general base mechanism for Mpro's proteolytic function and suggests that protonation of His172 (and not His163) induces a partial collapse of the S1

https://dx.doi.org/10.1021/jacs.0c10770
J. Am. Chem. Soc. 2020, 142, 21883−21890

pocket, consistent with the aforementioned pH 6 crystal structure of SARS-CoV Mpro. The simulations also predict an alternative Cys residue for targeted covalent inhibitor design. Taken together, our work represents a first step toward a mechanistic understanding of the pH-dependent structure−dynamics−function relationship of SARS-CoV-2/CoV Mpros and contributes urgently needed knowledge for broad-spectrum antiviral drug design.

RESULTS AND DISCUSSION

Protonation States of SARS-CoV-2 Mpro. The Mpro protomer has 12 Cys and 7 His residues, including the catalytic dyad Cys145 and His41 as well as the absolutely conserved His163, His164, and His172 in the S1 pocket. Because the solution pK_a's of Cys and His (8.5 and 6.5, respectively) are only 1 pH unit away from the physiological pH, it is important to determine their protonation states in the protein before conducting a fixed-charged MD study. We performed the pH replica-exchange GBNeck2-CpHMD simulations,[20,21] which have been shown to give accurate pK_a estimates, particularly for Cys[22] and His.[21] The simulations were initiated from the aforementioned crystal structure of SARS-CoV-2 Mpro (PDB: 6y2g[5]), whereby all side chains of Asp/Glu/His/Cys/Lys were allowed to titrate. Nine pH replicas were used over a pH range of 5−9, with a total sampling time of 495 ns. The calculated pK_a's were well converged (Figure S2).

CpHMD showed that the catalytic dyad residues Cys145 and His41 are both neutral at physiological pH, suggesting that the cleavage reaction of the Mpro follows a general base and not an ion-pair mechanism like SARS-CoV-2/CoV PLpros,[23] consistent with the pH-dependent activity data of SARS-CoV Mpro.[18] The latter gave the pK_a's of 8.3 and 6.4 for Cys145 and His41, respectively.[18] The calculated pK_a of His41 (6.6/6.7 for the two protomers, Table 1) is in excellent agreement with experiment, while the calculated pK_a of Cys145 is overestimated by about 2 units. Trajectory snapshots showed that the deprotonated form of Cys145 requires stabilization by several hydrogen bonds that are formed after extensive conformational sampling, consistent with our previous findings regarding the deprotonation of cysteines.[22,24] We defer the detailed investigation of this topic to a future study.

Table 1. Calculated pK_a's of Relevant Cys and His Residues in the Two Protomers of SARS-CoV-2/CoV Mpros[a]

	SARS-CoV-2		SARS-CoV	
Residue	A	B	A	B
C22	7.5	6.8	6.8	7.0
C44	7.0	9.2	4.2	5.8
C145	neutral	∼9.4[b]	neutral	neutral
H41	6.6	6.7	6.2	6.5
H163	neutral	neutral	neutral	neutral
H164	neutral	neutral	neutral	neutral
H172	6.6	6.6	6.6	7.7[c]

[a]Calculations based on the crystal structures of SARS-CoV-2 (PDB: 6y2g[5] and SARS-CoV (PDB: 1uk2[13]) Mpros. A complete list of pK_a's is given in SI Table 1. For residues that did not titrate in the simulation pH range, their protonation states are given. [b]Estimated based on partial deprotonation in the pH range 7.0−9.0. [c]See discussion in the main text.

Importantly, CpHMD titration gave the pK_a's of His163 and His164 below 5, whereas the pK_a of His172 is 6.6/6.6 (Table 1). Thus, our data supports the hypothesis that His163 is neutral, but it contradicts the hypothesis that His172 is charged at physiological pH.[5,13,17] Consequently, our data suggests that the protonation state switch of His172 (and not the proposed His163[5,13,17]) is responsible for the conformational changes of the S1 pocket at pH 6.

Conformational Changes in the S1 Pocket of SARS-CoV-2 Mpro. To test the hypothesis that the protonation of His172 leads to a conformational deactivation of SARS-CoV-2 Mpro, we carried out two independent runs of 2 μs fixed-charged MD starting from the crystal structure (PDB: 6y2g[5]). In run 1, His172 was neutral in protomer A and charged in protomer B, while in run 2, His172 was charged in protomer A and neutral in protomer B. All other residues were fixed in the standard protonation states, as determined by the CpHMD titration for physiological pH (see Table S1). During the simulations, the dimer and each protomer remained stable with a Cα root-mean-square deviation (RMSD) below 3 Å (Figure S3 and Figure S4). However, the RMSD of the oxyanion loop in the monomer with the charged His172 increased to about 2 Å within 500 ns, whereas that in the monomer with the neutral His172 stayed below 1 Å throughout the 2 μs simulations (Figure S5).

To understand the impact of the His172 protonation state, we first examine its interaction with Glu166 and the latter interaction with His163 by calculating the probability distributions of the side chain minimum distances. In both simulation runs, the charged His172 forms a salt bridge (weak hydrogen bond) with Glu166 (distance of ∼3.5 Å); however, the interaction between the neutral His172 and Glu166 is flexible, as evident from the distance distribution showing two peaks at 4.5/6 Å in the first run (Figure 2a) and 3.5/6 Å in the second run (Figure S6). Interestingly, in the presence of the charged His172, Glu166 also forms a weak hydrogen bond with His163 (distance of ∼3.5 Å), whereas with the neutral His172, the Glu166−His163 interaction is largely absent, with the most probable distance of 7 Å in the first run (Figure 2b) and 6.2 Å in the second run (Figure S6). Trajectory snapshots and the χ_1 angle distributions revealed that the side chain of Glu166 adopts different rotamer states dependent on the protonation state of His172. In the presence of the charged His172, Glu166 prefers the g^- state (χ_1 of 300°), whereas with the neutral His172, it mainly samples the *trans* state (χ_1 of 180°, similar to the crystal structure value) and occasionally adopts the g^+ state (χ_1 of 90°, Figure 2d and Figure S6). These data suggest that when His172 is charged, Glu166 is locked in the g^- state, interacting with both His172 and His163, whereas when His172 is neutral, Glu166 side chain is more flexible, sampling both *trans* and g^+ states and loosely interacting with His172 and His163.

Movement of the Oxyanion Loop in SARS-CoV-2 Mpro. The protonation state of His172 is also coupled to the conformation of the oxyanion loop L1. The Cα RMSD of L1 increased to 1.5−2 Å in the protomer with charged His172, but it remained below 1 Å in the protomer with neutral His172 (Figure S5). Trajectory snapshots revealed that L1 is moved closer to the S1 pocket when His172 is charged. In the presence of the charged His172, the distance between the center of mass of Glu166 and that of L1 is moved to ∼4 Å, whereas with neutral His172, the distance is ∼6 Å (Figure 2e and Figure S6). The L1 movement is likely due to the change in the interaction between Phe140 on L1 and His163. Crystal structures of SARS-Cov-2/CoV Mpros show an aromatic ring

https://dx.doi.org/10.1021/jacs.0c10770
J. Am. Chem. Soc. 2020, 142, 21883−21890

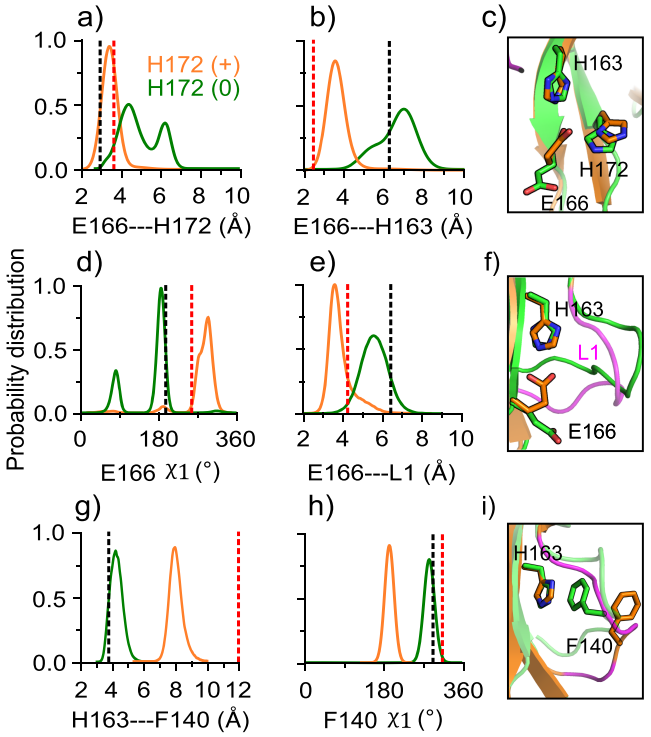


Figure 2. Conformational changes of the S1 pocket in SARS-CoV-2 Mpro is coupled to the switch of His172 protonation state. (a,b) Probability distribution of the minimum distance between the carboxylate oxygens of Glu166 and the imidazole nitrogens of His172 (a) and His163 (b). (d) Distribution of the χ_1 angle of Glu166. (e) Distribution of the distance between the center of mass of Glu166 (carboxylate oxygens) and the oxyanion loop (Cα atoms of residues 138−145). (g) Distribution of the distance between the center of mass of the aromatic rings of His163 and Phe140. (h) Distribution of the χ_1 angle of Phe140. Data for the protomer with neutral and charged His172 are colored green and orange, respectively. All plots were based on simulation run 1 (run 2 data given in Figure S6). The black and red dashed lines indicate the corresponding values from the protomer A of the SARS-CoV-2 Mpro structure (PDB: 6y2g[5]), and the inactive protomer B of the SARS-CoV Mpro structure (PDB: 1uj1[13]). (c,f,i) Snapshots showing the conformational differences between the protomers with neutral (green) and charged (orange) His172. The oxyanion loop is colored magenta.

stacking interaction between them, which is thought to stabilize L1 in the active form.[5,13] Simulations showed that the stacking interaction is maintained when His172 is neutral (distance of ∼4 Å) and disrupted when His172 is charged (distance of ∼8 Å, Figure 2g, Figure S6). Interestingly, the loss in the aromatic stacking is correlated with a change in the side chain conformation of Phe140, which adopts a *trans* rotamer state (χ_1 angle ∼180°) in the protomers with charged His172 and a g^- state (χ_1 angle ∼300°) in the protomers with neutral His172 (Figure 2h,i and Figure S6).

Comparison to Crystal Structures of Active and Inactive Mpro Protomers. The conformational changes of the S1 pocket induced by the protonation of His172 are consistent with the differences between the crystal structures of the active protomer of SARS-CoV-2 Mpro (PDB:6y2g[5]) and the inactive protomer of SARS-CoV Mpro (PDB:1uj1,[13] see Figure 2 dashed lines). Two seeming discrepancies are worth noting. While the Glu166−His172 interaction in the presence of neutral His172 is flexible in our simulations, the aforementioned crystal structure shows a hydrogen bond. This discrepancy may be due to the crystal vs solution condition. In fact, about 10 crystal structures of SARS-CoV-2 Mpro display a corresponding distance of ∼5 Å (Figure S1), in line with the simulation data (Figure 2a and Figure S6). Another seeming discrepancy is in the rotamer state of Phe140. In the simulations, Phe140 switches to a *trans* rotamer (χ_1 of 180°) with the charged His172, whereas in the crystal structure of the inactive protomer B of SARS-CoV Mpro (PDB: 1uj1[13]), the χ_1 angle remains similar to that in the active protomer. This may be explained by the extremely low electron density of the Phe140 side chain atoms in the X-ray structure (B-factor nearly 100).

N-Finger Interactions and Dimer Stability in SARS-CoV-2 Mpro. Having tested our hypothesis that protonation of His172 leads to a partial collapse of the S1 pocket, we turn to its interactions with the N-finger, which were thought to be disrupted at low pH.[5,13] Note, the positions of the first two N-terminal residues were missing in the pH 6 crystal structure of SARS-CoV Mpro (PDB: 1uj1[13]). In most of the crystal structures of SARS-CoV-2 Mpro, the interactions between Phe140/His172$^{A/B}$ and Ser1$^{B/A}$ are stable, while the Glu166$^{A/B}$−Ser1$^{B/A}$ distance varies between 2.2 and 5.3 Å (Figure S1).

In the simulation run 1, the interactions between Glu166^A/His172^A (neutral His172^A) and Ser1^B are largely maintained, and the interaction between Phe140^A and Ser1^B is flexible (Figure 3). In contrast, the interactions between Phe140^B/

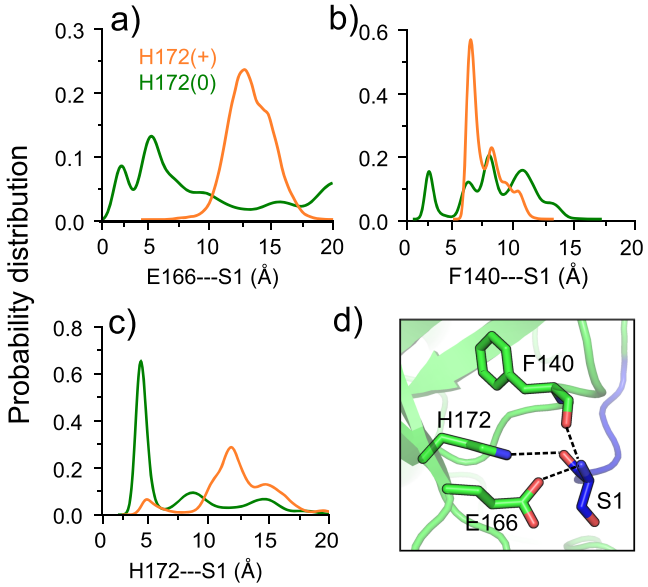


Figure 3. Interactions between the S1 pocket residues and Ser1 of the opposite protomer in SARS-CoV-2 Mpro are disrupted by the protonation of His172. (a,b,c) Probability distributions of the Glu166^A/Phe140^A/His172^A−Ser1^B distances are colored green; distributions of the Glu166^B/Phe140^B/His172^B−Ser1^A distances are colored orange. His172 is neutral in protomer A and charged in protomer B. The Glu166/Phe140−Ser1 distance refers to that between the N-terminal amino nitrogen and the nearest carboxyl oxygen of Glu166 or the carbonyl oxygen of Phe140. The His172−Ser1 distance is refers that between the backbone carbonyl oxygen of Ser1 and the nearest imidazole nitrogen of His172. (d) Zoomed-in view of the interactions between the S1 pocket (protomer A) and Ser1 (protomer B) in the X-ray structure of SARS-CoV-2 Mpro (PDB: 6y2g[5]). Hydrogen bonds are shown in dashed lines.

https://dx.doi.org/10.1021/jacs.0c10770
J. Am. Chem. Soc. 2020, 142, 21883−21890

Glu166^{B}/His172^{B} (charged His172^{B}) and Ser1^{A} are completely abolished (Figure 3). In the simulation run 2, the differences are more pronounced; all three interactions sample the crystal structure values when His172 is neutral but become disrupted when His172 is charged (Figure S7).

To further investigate the effect of His172 protonation on the dimerization stability, we examined several hydrogen bond contacts formed between the two protomers in the crystal structure (PDB: 6y2g[5]). Distance distributions show that all interactions, Gln127$^{A/B}$/ Glu290$^{A/B}$−Arg4$^{B/A}$, Ser10$^{A/B}$−Ser10$^{B/A}$, Glu14$^{A/B}$−Gly11$^{B/A}$, and Val125$^{A/B}$−Ala7$^{B/A}$, remain stable regardless of the His172 protonation state (Figure S8), which is consistent with the dimer stability in both simulation runs (Figure S3).

Protonation States and Proton-Coupled Dynamics of SARS-CoV Mpro. To further corroborate our findings, we conducted CpHMD titration and fixed-charged MD simulations of SARS-CoV Mpro starting from the crystal structure (PDB: 1uk2[13], convergence plots are given in Figure S9). The CpHMD titration results of SARS-CoV Mpro are consistent with those from SARS-CoV-2 Mpro (Table 1 and Table S1). The pK_a's of the catalytic His41 and Cys145 remain neutral at physiological pH (Table 1), supporting a general base mechanism for Mpro's proteolytic function.[18] Consistently, His163 and His164 remain neutral above pH 5.0. However, while the pK_a of His172^{A} is identical to that in SARS-CoV-2 Mpro, the pK_a of His172^{B} is one unit higher. The difference can be attributed to the loss of the interaction with Ser1^{A} in the crystal structure (PDB: 1uk2[13]). Due to the charged amino group, the His172^{B}−Ser1^{A} interaction would destabilize the charged form of His172^{B} and thereby lowering its pK_a.

To test the hypothesis that protonation of His172 leads to the conformational deactivation of the S1 pocket in SARS-CoV Mpro, we performed a 2 μs fixed-charged MD simulation starting from the aforementioned crystal structure (PDB: 1uk2[13]), in which His172 was fixed in neutral and charged states in protomers A and B, respectively. The differences in the S1 pocket interactions between protomer A and B are in complete agreement with those of SARS-CoV-2 Mpro (Figure S10). We did not test the S1 pocket interactions with Ser1, as the hydrogen bonds between the S1 pocket residues of protomer B and Ser1 of protomer A in the X-ray structure of SARS-CoV Mpro are absent (PDB: 1uk2[13]).

New Opportunity for Small-Molecule Broad-Spectrum Covalent Inhibitor Design. Currently, the majority of potent broad-spectrum antiviral inhibitors are large substrate mimetics carrying an electrophilic warhead that covalently binds with the catalytic Cys145 in SARS-CoV-2/CoV Mpros, e.g., the α-ketoamide based inhibitors from the Hilgenfeld lab[5] and Pfizer's ketone-based inhibitor PF-00835231 that has entered clinical trial.[11] The latter is promising, as it also showed additive/synergistic effect in combination with remdesivir which targets the RNA polymerase. However, delivery of large peptidomimetic compounds such as Pfizer's inhibitor[11] often requires intravenous infusion, which necessitates a hospital visit and lowers patient compliance. Thus, orally available small-molecule antiviral drugs are more desirable.

Interestingly, the CpHMD titration revealed that Cys22 and Cys44, which are conserved among the Mpros of SARS-CoV-2/CoV, MERS-CoV, and SARS-like bat-CoV, are significantly more nucleophilic than Cys145, as they are readily deprotonated at physiological pH (Table 1). The hyperreactivity of Cys44 and its promixity to the substrate binding pocket make it a particularly attractive site of covalent linkage. A recent crystal structure showed that galangin (3,5,7-trihydroxyflavone), a naturally occurring flavonoid with antitumor activities,[25,26] is bound to SARS-CoV-2 Mpro in a position with a minimum heavy-atom distance of 3.9 Å to either Cys44 or Cys145 (Figure 4). Thus, we envision the

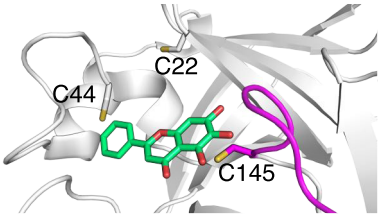


Figure 4. Natural flavonoid compound offers a starting point for designing small-molecule broad-spectrum targeted covalent inhibitors. X-ray structure of SARS-CoV-2 Mpro in complex with galangin (PDB: 6m2n) shows that it is in the proximity of both Cys44 and Cys145.

design of targeted covalent inhibitors of Mpros by installing a weakly electrophilic warhead on ganlangin to enable the covalent bond formation with the Cys44 thiolate. It is also conceivable that two warheads can be installed on both ends of ganlangin to bond with both Cys44 and Cys145, allowing superior potency. These strategies may lead to novel small-molecule broad-spectrum covalent inhibitors that have not been attempted so far.

CONCLUSION

Our simulations showed that the protonation state of His172 profoundly impacts the conformation of the S1 pocket through an interaction network involving His172, Glu166, His163, and Phe140. When His172 is neutral, Glu166 is flexible and does not form a hydrogen bond with His163, which allows His163 to maintain the aromatic stacking with Phe140 and thereby keeping the oxyanion loop in the active open conformation. However, when His172 is charged, it forms a salt bridge with Glu166, which locks Glu166 in the g^- rotameric state, stabilizing the hydrogen bond interaction with His163. The latter results in a disruption of the stacking interaction between His163 and Phe140 and consequently a collapse (closing) of the oxyanion loop toward the S1 pocket. Consistent with the pH 6 crystal structure of SARS-CoV Mpro, our data provides a detailed molecular mechanism of the conformational deactivation and loss of catalytic activity of SARS-CoV-2/CoV Mpros with decreased pH. Our future work will address the mechanism of the Mpro's activity decrease at high pH. We speculate that deprotonation of the amino terminus may play a significant role, as the solution pK_a of the α-amino group is 8.0[27] and in the proximity of the negatively charged Glu166, the pK_a may shift somewhat higher.

The present study has several caveats worth pointing out. In the CpHMD simulations, the N-terminus was acetylated which may have reduced the stability of the dimer interface. In the fixed-charge simulations, only the protonation state of His172 was varied, as it is one of the two residues suggested by experiment to possibly make an impact on the S1 pocket; however, His41 may also become protonated at similar pH (e.g., pH 6), and this effect was neglected in the present study. With regard to the dimerization stability, the fixed-charge

https://dx.doi.org/10.1021/jacs.0c10770
J. Am. Chem. Soc. 2020, 142, 21883−21890

simulations of SARS-CoV-2 Mpro demonstrated that His172 protonation disrupted the S1 pocket interactions with Ser1 but not other dimer interface contacts. The latter could be a result of the limited sampling time.

While a complete elucidation of the pH-dependent conformational mechanisms awaits future investigation, the present work offers a first step in understanding the proton-coupled structure−dynamics−function relationships of coronavirus Mpros. Our finding of the hyper-reactive cysteine and the proposed strategies of designing small molecule targeted covalent inhibitors may help accelerate the development of orally available broad-spectrum antiviral drugs to stop the current pandemic and prevent future outbreaks.

METHODS AND PROTOCOLS

System Preparation for CpHMD Simulations. The coordinates of the X-ray crystal structures of SARS-CoV-2 (PDB: 6y2g,[5] determined at pH 8.5) and SARS-CoV (PDB 1uk2,[13] determined at pH 8) Mpros were retrieved from the PDB. Both structures contain two independent protomers in the asymmetric unit. The N- and C-termini of each protomer were acetylated and amidated, respectively (free in the fixed-charge simulations). Missing hydrogens were added using the HBUILD facility in CHARMM (C38b2);[28] dummy hydrogens for titration of Asp/Glu were added using a custom CHARMM script.[29] A short energy minimization was performed for the hydrogen positions using the steepest decent and Newton−Raphson methods for 20 steps with the heavy atoms constrained. The force field parameters and coordinate files were constructed using the LEAP utility in AMBER.[30] The protein was represented by the ff14sb force field[31] and solvent was represented by the GBNeck2 (igb = 8) implicit-solvent model.[32] The default mbondi3 intrinsic Born radii, except for His and Cys, for which the following radii optimized for CpHMD were used: 1.17 Å for imidazole nitrogen atoms of His[20] and 2.0 Å for sulfur atom of Cys.[33] Following energy minimization using the steepest decent and conjugate-gradient algorithms for 5000 and 1000 steps, respectively, the system underwent restrained equilibration at pH 7 in four stages (2000 MD steps each). The force constant in the heavy-atom restraint was 5, 2, 1, and 0 kcal/mol/Å^2 in the four stages. The final structure was used for the CpHMD titration simulations.

pH Replica-Exchange CpHMD Simulations. The CpHMD titration simulations were performed with the GPU-accelerated GBNeck2-CpHMD method[21] in Amber18.[30] The pH replica-exchange protocol[34] was used to accelerate convergence. To allow for the use of one or two GPU cards, we applied the newly implemented asynchronous replica-exchange scheme.[23] For SARS-CoV-2 Mpro, 9 replicas over the pH range 5−9 with an interval of 0.5 pH unit were used, with an aggregate sampling time of 495 ns. For SARS-CoV Mpro, 9 replicas over the pH range 4.5−8.5 were used, with an aggregate sampling time of 360 ns. All Asp, Glu, His, Cys, and Lys side chains were allowed to titrate, with the CpHMD default model pK_a's of 3.8, 4.2, 6.5, 8.5, and 10.4, respectively.[21] The simulations were run at 300 K with an ionic strength of 0.15 M and an effectively infinite cutoff (999 Å) for nonbonded interactions. The SHAKE algorithm was used to constrain the bonds involving the hydrogens to allow for a 2 fs time step. For pK_a calculations and analysis, data from the first 20 ns per replica were discarded.

Conventional Fixed-Charge All-Atom MD Simulations. The initial structures for SARS-CoV-2/CoV Mpros (PDB: 6y2g[5] and 1uk2[13]) prepared using the LEAP utility in Amber.[30] The termini of each protomer were left in the free form. Hydrogens were added according to the appropriate protonation states (see below). Sodium and chloride ions were added to represent a neutral system with a 150 mM physiological salt concentration. Two 2-μs all-atom fixed-charged MD simulations were performed for SARS-CoV-2/CoV Mpros (PDB: 6y2g[5]), whereby the protonation states (except for His172) were fixed to those as determined by CpHMD for physiological pH, i.e., Asp(−)/Glu(−)/His(0)/Cys(0). In simulation run 1, His172 was neutral in protomer A and charged in protomer B, whereas in simulation run 2, His172 was charged in monomer A and neutral in monomer B. As a comparison, one 2-μs simulation for SARS-CoV Mpro was performed using the same protocol as run 1 of SARS-CoV-2 Mpro.

All simulations were performed with Amber18.[30] The Mpro dimer was placed in a truncated octahedral water box, with a minimum of 11 Å between the protein heavy atoms and the edges of the water box. The protein was represented by the ff14SB force field[31] and water represented by the TIP3P[35] model. Long-range electrostatic interactions were treated with the particle-mesh Ewald method.[36] A nonbonded cutoff of 8 Å was used. SHAKE was turned on to allow a 2 fs time step. Initially, the system was energy minimized by applying 5000 steps of steepest descent followed by 5000 steps of conjugate gradient minimization with a force constant of 25 kcal/mol/Å^2 to the solute heavy atoms. The force constant was then reduced to 5 kcal/mol/Å^2, and the system was heated from 100 to 300 K in the canonical ensemble for 50 ps followed by 250 ps restrained and 100 ps unrestrained equilibration in the isothermal−isobaric ensemble with the isotropic Berendsen barostat.[37] Finally, the production runs of 2 μs each were performed, starting from a different random initial velocity seed for each system. All analysis was performed with the Amber module CPPTRAJ[38] using the last 1 μs of each trajectory.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.0c10770.

Supplemental tables and figures (PDF)

AUTHOR INFORMATION

Corresponding Author

Jana Shen − *Department of Pharmaceutical Sciences, University of Maryland School of Pharmacy, Baltimore, Maryland 21201, United States;* orcid.org/0000-0002-3234-0769; Email: Jana.Shen@rx.umaryland.edu

Authors

Neha Verma − *Department of Pharmaceutical Sciences, University of Maryland School of Pharmacy, Baltimore, Maryland 21201, United States*

Jack A. Henderson − *Department of Pharmaceutical Sciences, University of Maryland School of Pharmacy, Baltimore, Maryland 21201, United States;* orcid.org/0000-0001-6675-7944

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.0c10770

Author Contributions

‡J.A.H. is a joint first author.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Financial support from the National Institutes of Health (GM09888) is acknowledged. JAH acknowledges a fellowship from American Foundation for Pharmaceutical Education.

REFERENCES

(1) Wu, F.; Zhao, S.; Yu, B.; Chen, Y.-M.; Wang, W.; Song, Z.-G.; Hu, Y.; Tao, Z.-W.; Tian, J.-H.; Pei, Y.-Y.; Yuan, M.-L.; Zhang, Y.-L.; Dai, F.-H.; Liu, Y.; Wang, Q.-M.; Zheng, J.-J.; Xu, L.; Holmes, E. C.; Zhang, Y.-Z. A New Coronavirus Associated with Human Respiratory Disease in China. *Nature* **2020**, *579*, 265−269.

(2) Zhang, L.; Lin, D.; Kusov, Y.; Nian, Y.; Ma, Q.; Wang, J.; von Brunn, A.; Leyssen, P.; Lanko, K.; Neyts, J.; de Wilde, A.; Snijder, E. J.; Liu, H.; Hilgenfeld, R. α-Ketoamides as Broad-Spectrum Inhibitors of Coronavirus and Enterovirus Replication: Structure-Based Design, Synthesis, and Activity Assessment. *J. Med. Chem.* **2020**, *63*, 4562−4578.
(3) Hilgenfeld, R. From SARS to MERS: Crystallographic Studies on Coronaviral Proteases Enable Antiviral Drug Design. *FEBS J.* **2014**, *281*, 4085−4096.
(4) Needle, D.; Lountos, G. T.; Waugh, D. S. Structures of the Middle East Respiratory Syndrome Coronavirus 3C-like Protease Reveal Insights into Substrate Specificity. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2015**, *71*, 1102−1111.
(5) 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 α-Ketoamide Inhibitors. *Science* **2020**, *368*, 409−412.
(6) Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; Duan, Y.; Yu, J.; Wang, L.; Yang, K.; Liu, F.; Jiang, R.; Yang, X.; You, T.; Liu, X.; Yang, X.; Bai, F.; Liu, H.; Liu, X.; Guddat, L. W.; Xu, W.; Xiao, G.; Qin, C.; Shi, Z.; Jiang, H.; Rao, Z.; Yang, H. Structure of Mpro from SARS-CoV-2 and Discovery of Its Inhibitors. *Nature* **2020**, *582*, 289−293.
(7) Rathnayake, A. D.; Zheng, J.; Kim, Y.; Perera, K. D.; Mackin, S.; Meyerholz, D. K.; Kashipathy, M. M.; Battaile, K. P.; Lovell, S.; Perlman, S.; Groutas, W. C.; Chang, K.-O. 3C-like Protease Inhibitors Block Coronavirus Replication in Vitro and Improve Survival in MERS-CoV−Infected Mice. *Sci. Transl. Med.* **2020**, *12*, No. eabc5332.
(8) Ma, C.; Sacco, M. D.; Hurst, B.; Townsend, J. A.; Hu, Y.; Szeto, T.; Zhang, X.; Tarbet, B.; Marty, M. T.; Chen, Y.; Wang, J. Boceprevir, GC-376, and Calpain Inhibitors II, XII Inhibit SARS-CoV-2 Viral Replication by Targeting the Viral Main Protease. *Cell Res.* **2020**, *30*, 678−692.
(9) Jin, Z.; Zhao, Y.; Sun, Y.; Zhang, B.; Wang, H.; Wu, Y.; Zhu, Y.; Zhu, C.; Hu, T.; Du, X.; Duan, Y.; Yu, J.; Yang, X.; Yang, X.; Yang, K.; Liu, X.; Guddat, L. W.; Xiao, G.; Zhang, L.; Yang, H.; Rao, Z. Structural Basis for the Inhibition of SARS-CoV-2 Main Protease by Antineoplastic Drug Carmofur. *Nat. Struct. Mol. Biol.* **2020**, *27*, 529−532.
(10) Kneller, D. W.; Phillips, G.; O'Neill, H. M.; Jedrzejczak, R.; Stols, L.; Langan, P.; Joachimiak, A.; Coates, L.; Kovalevsky, A. Structural Plasticity of SARS-CoV-2 3CL Mpro Active Site Cavity Revealed by Room Temperature X-Ray Crystallography. *Nat. Commun.* **2020**, *11*, 3202.
(11) Hoffman, R. L.; Kania, R. S.; Brothers, M. A.; Davies, J. F.; Ferre, R. A.; Gajiwala, K. S.; He, M.; Hogan, R. J.; Kozminski, K.; Li, L. Y.; Lockner, W.; Lou, J.; Marra, M. T.; Jr, L. J. M.; Murray, B. W.; Nieman, A.; Noell, S.; Planken, S. P.; Rowe, T.; Ryan, K.; Smith, G. J.; Solowiej, J. E.; Steppan, C. M.; Taggart, B. The Discovery of Ketone-Based Covalent Inhibitors of Coronavirus 3CL Proteases for the Potential Therapeutic Treatment of COVID-19. *J. Med. Chem.* **2020**, *63*, 12725−12747.
(12) Shi, J.; Sivaraman, J.; Song, J. Mechanism for Controlling the Dimer-Monomer Switch and Coupling Dimerization to Catalysis of the Severe Acute Respiratory Syndrome Coronavirus 3C-Like Protease. *J. Virol.* **2008**, *82*, 4620−4629.
(13) 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. U. S. A.* **2003**, *100*, 13190−13195.
(14) Hsu, W.-C.; Chang, H.-C.; Chou, C.-Y.; Tsai, P.-J.; Lin, P.-I.; Chang, G.-G. Critical Assessment of Important Regions in the Subunit Association and Catalytic Action of the Severe Acute Respiratory Syndrome Coronavirus Main Protease. *J. Biol. Chem.* **2005**, *280*, 22741−22748.
(15) Chen, S.; Hu, T.; Zhang, J.; Chen, J.; Chen, K.; Ding, J.; Jiang, H.; Shen, X. Mutation of Gly-11 on the Dimer Interface Results in the Complete Crystallographic Dimer Dissociation of Severe Acute Respiratory Syndrome Coronavirus 3C-like Protease: *Crystal Structurewith Molecular Dynamics Simulations. J. Biol. Chem.* **2008**, *283*, 554−564.
(16) Anand, K.; Palm, G. J.; Mesters, J. R.; Siddell, S. G.; Ziebuhr, J.; Hilgenfeld, R. Structure of Coronavirus Main Proteinase Reveals Combination of a Chymotrypsin Fold with an Extra α-Helical Domain. *EMBO J.* **2002**, *21*, 3213−3224.
(17) Tan, J.; Verschueren, K. H.; Anand, K.; Shen, J.; Yang, M.; Xu, Y.; Rao, Z.; Bigalke, J.; Heisen, B.; Mesters, J. R.; Chen, K.; Shen, X.; Jiang, H.; Hilgenfeld, R. pH-Dependent Conformational Flexibility of the SARS-CoV Main Proteinase (Mpro) Dimer: Molecular Dynamics Simulations and Multiple X-Ray Structure Analyses. *J. Mol. Biol.* **2005**, *354*, 25−40.
(18) Huang, C.; Wei, P.; Fan, K.; Liu, Y.; Lai, L. 3C-like Proteinase from SARS Coronavirus Catalyzes Substrate Hydrolysis by a General Base Mechanism. *Biochemistry* **2004**, *43*, 4568−4574.
(19) Xue, X.; Yang, H.; Shen, W.; Zhao, Q.; Li, J.; Yang, K.; Chen, C.; Jin, Y.; Bartlam, M.; Rao, Z. Production of Authentic SARS-CoV Mpro with Enhanced Activity: Application as a Novel Tag-Cleavage Endopeptidase for Protein Overproduction. *J. Mol. Biol.* **2007**, *366*, 965−975.
(20) Huang, Y.; Harris, R. C.; Shen, J. Generalized Born Based Continuous Constant pH Molecular Dynamics in Amber: Implementation, Benchmarking and Analysis. *J. Chem. Inf. Model.* **2018**, *58*, 1372−1383.
(21) Harris, R. C.; Shen, J. GPU-Accelerated Implementation of Continuous Constant pH Molecular Dynamics in Amber: pKa Predictions with Single-pH Simulations. *J. Chem. Inf. Model.* **2019**, *59*, 4821−4832.
(22) Harris, R. C.; Liu, R.; Shen, J. Predicting Reactive Cysteines with Implicit-Solvent-Based Continuous Constant pH Molecular Dynamics in Amber. *J. Chem. Theory Comput.* **2020**, *16*, 3689−3698.
(23) Henderson, J. A.; Verma, N.; Harris, R. C.; Liu, R.; Shen, J. Assessment of Proton-Coupled Conformational Dynamics of SARS and MERS Coronavirus Papain-like Proteases: Implication for Designing Broad-Spectrum Antiviral Inhibitors. *J. Chem. Phys.* **2020**, *153*, 115101.
(24) Liu, R.; Yue, Z.; Tsai, C.-C.; Shen, J. Assessing Lysine and Cysteine Reactivities for Designing Targeted Covalent Kinase Inhibitors. *J. Am. Chem. Soc.* **2019**, *141*, 6553−6560.
(25) Chien, S.-T.; Shi, M.-D.; Lee, Y.-C.; Te, C.-C.; Shih, Y.-W. Galangin, a Novel Dietary Flavonoid, Attenuates Metastatic Feature via PKC/ERK Signaling Pathway in TPA-Treated Liver Cancer HepG2 Cells. *Cancer Cell Int.* **2015**, *15*, 15.
(26) Kong, Y.; Feng, Z.; Chen, A.; Qi, Q.; Han, M.; Wang, S.; Zhang, Y.; Zhang, X.; Yang, N.; Wang, J.; Huang, B.; Zhang, Q.; Xiang, G.; Li, W.; Zhang, D.; Wang, J.; Li, X. The Natural Flavonoid Galangin Elicits Apoptosis, Pyroptosis, and Autophagy in Glioblastoma. *Front. Oncol.* **2019**, *9*, 942.
(27) Thurlkill, R. L.; Grimsley, G. R.; Scholtz, J. M.; Pace, C. N. {{\bvphantom}} \evphantompK{}} Values of the Ionizable Groups of Proteins. *Protein Sci.* **2006**, *15*, 1214−1218.
(28) Brooks, B.; Brooks, C.; MacKerell, A.; Nilsson, L.; Petrella, R.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R.; Post, C.; Pu, J.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D.; Karplus, M. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545−1614.
(29) Khandogin, J.; Brooks, C. L. Constant pH Molecular Dynamics with Proton Tautomerism. *Biophys. J.* **2005**, *89*, 141−157.
(30) Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T., III; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Greene, D.; Harris, R.; Homeyer, N.; Huang, Y.; Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.;

Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; Salomon-Ferrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. *AMBER 2018*, 2018.
(31) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696−3713.
(32) Nguyen, H.; Roe, D. R.; Simmerling, C. Improved Generalized Born Solvent Model Parameters for Protein Simulations. *J. Chem. Theory Comput.* **2013**, *9*, 2020−2034.
(33) Liu, R.; Yue, Z.; Tsai, C.-C.; Shen, J. Assessing Lysine and Cysteine Reactivities for Designing Targeted Covalent Kinase Inhibitors. *J. Am. Chem. Soc.* **2019**, *141*, 6553−6560.
(34) Wallace, J. A.; Shen, J. K. Continuous Constant pH Molecular Dynamics in Explicit Solvent with pH-Based Replica Exchange. *J. Chem. Theory Comput.* **2011**, *7*, 2617−2629.
(35) Jorgensen, W. L. Efficient Drug Lead Discovery and Optimization. *Acc. Chem. Res.* **2009**, *42*, 724−733.
(36) Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878−3888.
(37) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(38) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084−3095.

https://dx.doi.org/10.1021/jacs.0c10770
J. Am. Chem. Soc. 2020, 142, 21883−21890