RESEARCH ARTICLE

PROTEINS WILEY

Molecular interactions and inhibition of the SARS-CoV-2 main protease by a thiadiazolidinone derivative

Jacob Andrzejczyk[1] | Katarina Jovic[2] | Logan M. Brown[2] | Valerie G. Pascetta[2] |
Krisztina Varga[2] | Harish Vashisth[1]

[1]Department of Chemical Engineering, University of New Hampshire, Durham, New Hampshire, USA

[2]Department of Molecular, Cellular, and Biomedical Services, University of New Hampshire, Durham, New Hampshire, USA

Correspondence
Harish Vashisth, Department of Chemical Engineering, University of New Hampshire, Durham, NH 03824, USA.
Email: harish.vashisth@unh.edu

Funding information
National Institutes of Health, Grant/Award Numbers: R35GM138217, P20GM113131; National Science Foundation, Grant/Award Number: OIA-1757371

Abstract

We report molecular interactions and inhibition of the main protease (M^{Pro}) of SARS-CoV-2, a key enzyme involved in the viral life cycle. By using a thiadiazolidinone (TDZD) derivative as a chemical probe, we explore the conformational dynamics of M^{Pro} via docking protocols and molecular dynamics simulations in all-atom detail. We reveal the local and global dynamics of M^{Pro} in the presence of this inhibitor and confirm the inhibition of the enzyme with an IC_{50} value of $1.39 \pm 0.22\ \mu M$, which is comparable to other known inhibitors of this enzyme.

KEYWORDS
docking, main protease, molecular dynamics, SARS-CoV-2, thiadiazolidinone

1 | INTRODUCTION

COVID-19 is caused by SARS-CoV-2, a novel β-coronavirus with a ~30 kilobase genome.[1] The virion contains four main structural proteins: the small envelope (E) glycoprotein, the membrane (M) glycoprotein, the nucleocapsid (N) protein, and the spike (S) glycoprotein.[2] Outside of these structural proteins, the SARS-CoV-2 genome encodes for 16 nonstructural proteins (nsps 1–16) and several accessory proteins.[3] The nsps have a wide range of functions, including suppression of the host immune response, cleavage of polyproteins, promotion of cytokine expression, and proofreading of the viral genome. Two of the nsps, nsp3 and nsp5, also known as the papain-like protease (PL^{Pro}) and the main protease (M^{Pro}),[4,5] are essential for viral replication due to their role in the cleavage of viral polyproteins. Of the two, M^{Pro} is considered the preferred therapeutic target due to its high sequence conservation with proteases of other coronaviruses,[6,7] unique sequence recognition, and the knowledge of its cleavage mechanism.[8,9]

M^{Pro} is genetically encoded as the nsp5 and is responsible for cleaving at least 11 sites in the viral polyproteins (pp1a and pp1b).[10] M^{Pro} is a homodimeric enzyme, with each monomer containing 306 amino acids, divided into three subdomains (Figure 1A): the N-terminal domain I (residues 8–101) and domain II (residues 102–184), and the C-terminal domain III (residues 201–306).[10] The domains

I and II are primarily composed of β-barrels and the domain III is chiefly α-helical.[11] The active site (Figure 1B) is defined by the catalytic dyad formed by a cysteine residue (C145) and a histidine residue (H41), and is located at the interface of domains I and II. It has been proposed that the cleavage mechanism of M^{Pro} is initiated by the polarization of the SH group of C145 by the imidazole group of H41,[8] eventually resulting in a proton transfer from the S_γ atom of C145 to H41, and this mechanism has been further probed via QM/MM studies.[9]

The need for an alternative drug development pipeline, instead of the intensive process of introducing a novel drug to market,[12] is highlighted by the high infection rate of COVID-19, compared to the infection rates of previous coronaviruses.[13] The application of known compounds outside of their original purpose, also known as drug repurposing, offers a way to facilitate the identification of new properties, functions, and utilities of existing compounds, to potentially deliver an effective therapeutic than would otherwise be possible.[14–16] An example of a recently repurposed compound is tideglusib, an irreversible small molecule kinase inhibitor,[17–19] that was shown to also inhibit the activity of M^{Pro}.[20] In addition, another compound known as shikonin, a traditional Chinese medicine derived from the root of *Lithospermum erythrorhizon*, has several biological functions,[21–23] and was shown to inhibit M^{Pro}.[20] Other compounds have been repurposed and tested on M^{Pro} as well.[24–31]
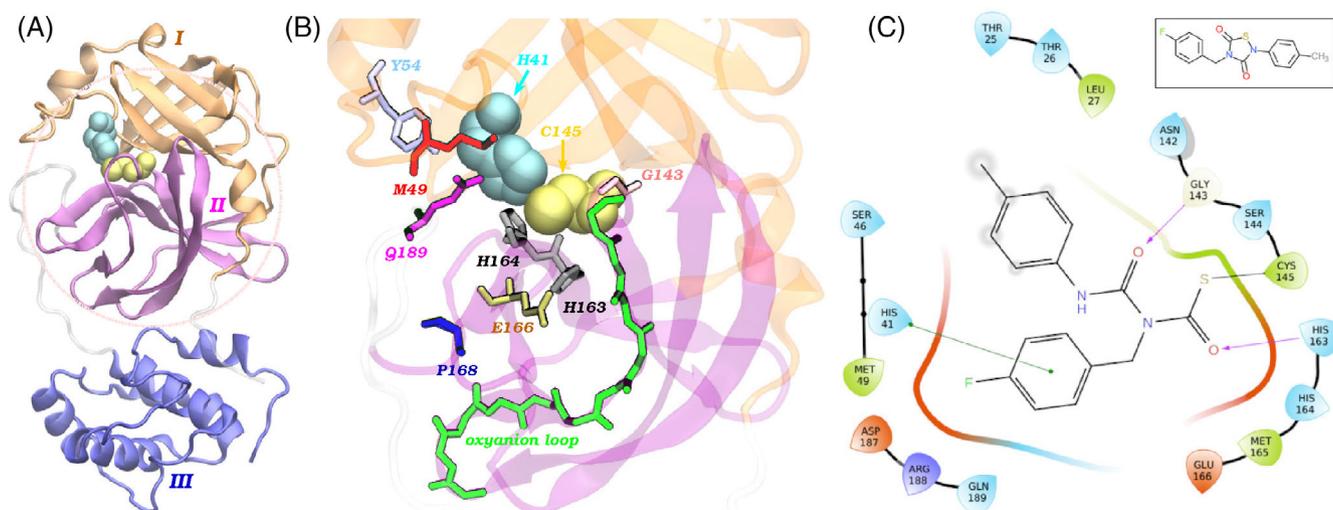
FIGURE 1 Structural details of M^Pro and key ligand interactions. (A) The protomer A of M^Pro with domain I (orange), domain II (purple), and domain III (blue) highlighted. Shown also are the catalytic residues H41 (cyan) and C145 (yellow). (B) A zoomed view of the active site with key residues colored, labeled, and shown in stick representations: M49 (red), G143 (pink), H163/H164 (gray), E166 (tan), P168 (blue), Q189 (purple), and the backbone atoms of residues that form the oxyanion loop (green). (C) Ligand interaction map of covalently bound CCG-50014. We show covalent bonds by black lines, hydrogen bonds by pink arrows, and π–π stacking interactions by green lines. Also shown are neighboring active site residues, in "guitar pick" representations where the narrow side shows the orientation of the sidechain of the residue and the wide side shows the orientation of the backbone of the residue, within 5 Å of CCG-50014. The van der Waals edges of negatively charged residues (red), hydrophobic residues (green), and polar residues (blue) are also shown. A chemical structure of CCG-50014 is also shown (inset box).

Given that the catalytic dyad of M^Pro contains a cysteine (C145) residue, which is the only reactive cysteine in the enzyme,[32] we explored the inhibitory properties of a potent thiadiazolidinone (TDZD) derivative, CCG-50014 (inset, Figure 1C), that is known to target regulators of G-protein signaling (RGS) proteins.[33,34] We have also studied the effect of various TDZD derivatives on the dynamics and inhibition of RGS proteins,[33,35–37] and showed that these compounds are efficient covalent modifiers of cysteine residues. Therefore, we studied the covalent modification of the catalytic cysteine residue, C145, to understand its effect on the conformational dynamics of M^Pro. We first performed covalent docking studies to derive the most energetically favorable protein/ligand conformation (Figure 1C) and then carried out molecular dynamics (MD) simulations of the covalently bound compound to probe its interactions in the binding pocket (Figures 2–4 and Figures S1–S5). In addition, we investigated the noncovalent binding and stability of CCG-50014 using docking and MD simulations (see Supporting Information: supplemental results, Tables S1 and S5, and Figures S7 and S8). Finally, we have expressed and purified recombinant M^Pro and performed IC_{50} assays to test the inhibitory effect of the ligand (Figure 6 and Figures S9–S11).

2 | MATERIALS AND METHODS

2.1 | Simulation methods

2.1.1 | Covalent docking and MD simulations

In docking and simulation studies, we used the atomic coordinates of protein atoms based on the protomer A from the crystal structure of

M^Pro (PDB ID: 6LU7).[38] We performed all docking studies in the Schrödinger Maestro Suite.[39] We prepared the protein structure for docking studies by adding hydrogen atoms to the protein, restrained minimization of the heavy atoms, and a short energy minimization step without restraints.[40] We conducted MD simulations using the NAMD software suite,[41] combined with the CHARMM forcefield.[42] We parameterized the ligand using CHARMM-GUI.[43] We solvated each system using TIP3P water molecules[44] and neutralized with NaCl. After an initial energy minimization of 1000 cycles, we optimized the volume of the simulation domain for each system in the NPT ensemble for 20 ps. We controlled the pressure at 1 atm using the Nosé–Hoover barostat, and we maintained the temperature at 310K using the Langevin thermostat.

Using the Schrödinger CovDock tool,[45] we performed covalent docking of CCG-50014 on the crystal structure of M^Pro. In these docking calculations, we identified the catalytic residue C145 as the target residue for the formation of a disulfide bond. We generated 100 covalently docked poses and performed molecular mechanics with generalized Born and surface area continuum solvation (MM-GBSA) calculations for each pose, and chose the docking pose with the most negative score (based on MM-GBSA values) for subsequent MD simulations. Briefly, the Schrödinger tool utilizes an implicit solvent model[46] to carry out MM-GBSA binding free energy calculations by deriving strain energies from the optimized structures of the receptor and ligand in the free and complexed states. The overall energies of the receptor and ligand are split into many components including the contributions from the Coulomb energy, Van der Waals energy, lipophilic energy, Born electrostatic solvation energy as well as the energies originating in correction terms for self-contact, hydrogen bonding, and π–π packing. Based on the chosen covalently docked
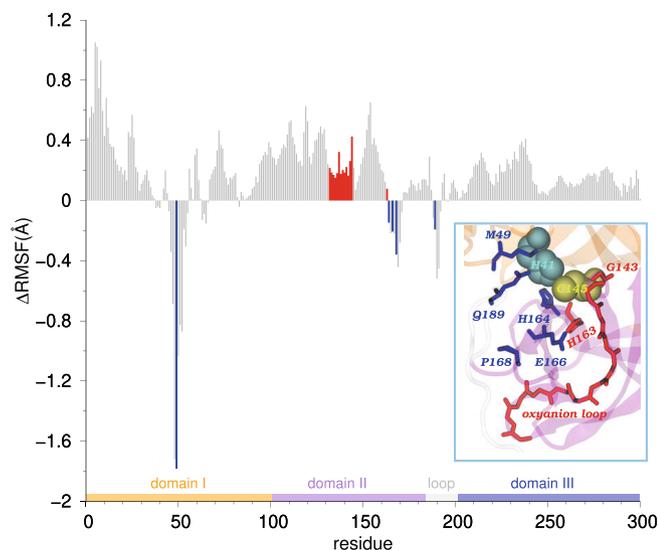
FIGURE 2 The ΔRMSF of all residues in M^Pro. We show the ΔRMSF values of residues 1 to 300 of M^Pro, where a negative value indicates the residue with a decreased flexibility in the liganded form of M^Pro. The boundaries of individual domains are shown near the x-axis: domain I (orange), domain II (purple), domain III (blue), and the loop (gray). A snapshot of the active site of M^Pro with the key active site residues (Figure 1B) colored by their ΔRMSF values, positive values (red) and negative values (blue). RMSF, root mean squared fluctuation.

pose, we conducted three independent MD simulations, each 500 ns long. We also performed three independent MD simulations (each 500 ns) of the *apo* form of the protomer A of M^Pro. The objective of these simulations was to investigate the dynamics of M^Pro in the *apo* and covalently bound CCG-50014 conformations.

2.1.2 | Noncovalent docking and MD simulations

We used the Schrödinger SiteMap tool[47] to identify three potential binding sites on the surface of the protein. Using the top potential binding site (determined based on the extent of solvent exposure) which was located around the active site, we specified a docking grid with a volume of $10 \times 10 \times 10$ Å^3 around the center of mass of the catalytic dyad (C145 and H41) to ensure that the coordinates of the docked ligand were in proximity of the catalytic dyad. We carried out standard precision docking[48] of CCG-50014 on 50 conformations of M^Pro, obtained from a short 1 ns long MD simulation, generating ~1000 bound poses per ligand per conformation, for a total of 50 000 poses. Based on docking scores, we then reduced the number of these poses to ~80 poses per ligand per conformation, for a new total of ~4000 poses. Furthermore, we performed MM-GBSA calculations to assess the binding affinity of each inhibitor to the active site. To conduct MD simulations of noncovalently docked CCG-50014, we
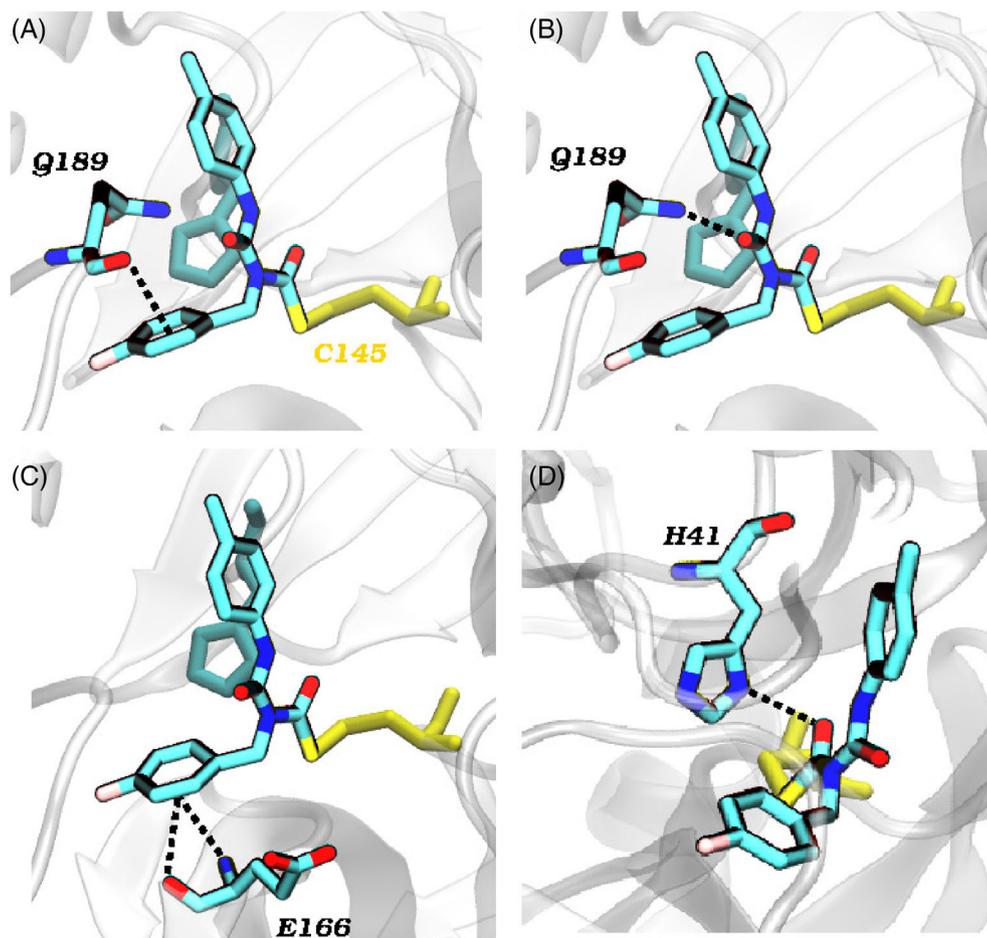

FIGURE 3 Key interactions of CCG-50014 in the binding pocket. (A) A snapshot highlighting the π–hydrogen interactions between the oxygen atom of the residue Q189 and the fluorinated aromatic ring of CCG-50014. (B) A snapshot showing the hydrogen bond between the nitrogen atom of the residue Q189 and the second oxygen atom of CCG-50014. (C) A snapshot showing the π–hydrogen interactions between the oxygen and nitrogen atom of the residue E166 and the 4-fluorophenyl ring of CCG-50014. (D) A snapshot highlighting the hydrogen bond between the nitrogen atom of the imidazole ring of the residue H41 and the first oxygen atom of CCG-50014.
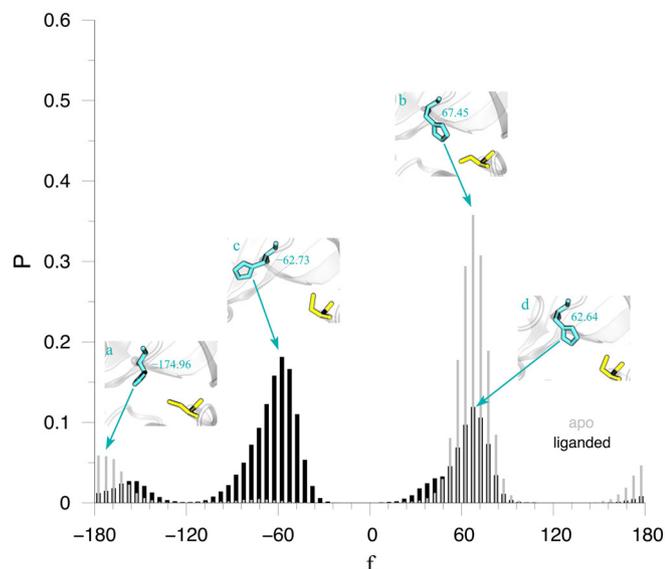
FIGURE 4   Conformational states of the sidechain of the residue H41, as characterized by the dihedral angle ϕ. We show the distributions of ϕ in the *apo* (gray) and liganded (black) states of M^{Pro}. Shown also are snapshots of the sidechain of H41 corresponding to the most populated states in apo (inset: a and b) and liganded (inset: c and d) states.

chose five poses (Figure S7) as the initial coordinates to conduct five independent MD simulations. We chose the poses by evaluating the proximity of the ligand to the catalytic C145 residue and then by the most negative MM-GBSA values. We conducted these five MD simulations, each with a maximum length of 500 ns, under the same simulation conditions as done for the covalently docked pose.

2.1.3 | Conformational metrics in MD simulations

The primary objective of our all-atom MD simulations is to investigate the dynamics of M^{Pro} in the *apo* and CCG-50014 bound conformations. We used various conformational metrics to characterize these dynamics at different scales in the protein structure including motions in residue sidechains and the overall interresidue communication network at the tertiary structure level. Specifically, the local dynamics are characterized through interresidue distances and interactions, fluctuations in residue sidechains, and measurements on backbone dihedral angles of key residues, while the interresidue communication is analyzed through a community network analysis approach. Some metrics are computed only for those residues that are directly used in the definition of the binding pocket due to their interactions and locations in the proximity of the inhibitor CCG-50014. For example, H41 and C145 are two key residues forming the catalytic dyad, the residues M49, Y54, and Q189 are in the vicinity of H41, while G143, H163, H164, and E166 are in the vicinity of C145 (Figure 1B). We note that several of these residues have direct interactions with CCG-50014 (Figure 1C) and are likely perturbed on inhibitor binding, thereby highlighting the significance of structural analyses of these residues. We provide additional details on all conformational metrics below.

Ligand interactions

In the ligand interaction maps (Figure 1C and Figure S7), we refer to the oxygen atoms in the ligand as 1 and 2. The oxygen atom 1 in the ligand is the oxygen bound to the carbon adjacent to the sulfur atom. The oxygen atom 2, therefore, refers to the other atom. We show the covalent bond by a black line, hydrogen bonds by pink arrows, π–π stacking interacting by green lines, and pi–cation interactions by red lines. We show negatively charged residues by a red border, hydrophobic residues by a green border, and polar residues by a blue border.

Root mean squared fluctuation

We calculated the root mean squared fluctuation (RMSF) values of protein residues 8–300 for *apo*, covalently bound, and noncovalently bound structures of M^{Pro}. The initial coordinates, excluding hydrogens, from docking were used as the reference frame. We also calculated ΔRMSF values to study the flexibility of each residue in the presence of covalently and noncovalently bound CCG-50014 in comparison to the *apo* state. A negative ΔRMSF value indicates decreased fluctuations in the presence of the ligand and a positive ΔRMSF value indicates increased fluctuations in the presence of the ligand.

Dihedral angle of H41 residue

We calculated the dihedral angle (ϕ) of H41 to investigate the rotameric conformational states of the sidechain of H41. We provide distributions of ϕ (Figure 4 and Figures S4 and S5) computed based on data from three independent MD simulations of apo and liganded systems.

Interresidue distances

We calculated the time-averaged Euclidean distances between key binding-pocket residues, and the residues in the catalytic dyad (H41 and C145) based on the center of mass of each residue (Table S2 and Figure S2). We also calculated the changes in distances (Δd) between *apo* and covalently bound runs. A negative Δd value indicates the distance between residues decreased in the presence of ligand.

Network analysis

To investigate allosteric pathways and correlated residues, we performed a dynamics-based network analysis on the *apo* and covalently bound CCG-50014 MD simulations (Figure 5). We used the method developed by Sethi et al.,[49] which has been applied previously to RGS[35] and other proteins.[50-55] For this analysis, we used data from three independent MD simulations of the *apo* and covalently bound CCG-50014 systems. We first performed a cross-correlation analysis using CARMA[56] by setting the C_α-atoms of residues as nodes. The correlations are calculated by the following equation:

$$C_{ij} = \frac{\langle \Delta \vec{r}_i(t) \cdot \Delta \vec{r}_j(t) \rangle}{\left(\langle \Delta \vec{r}_i(t)^2 \rangle \langle \Delta \vec{r}_j(t)^2 \rangle \right)^{1/2}}, \tag{1}$$

with $\Delta \vec{r}_i(t) = \vec{r}_i(t) - \langle \vec{r}_i(t) \rangle$, $\vec{r}_i(t)$ is the position of node i, and $\langle \vec{r}_i(t) \rangle$ is the mean position of node i. An edge is formed between two nodes
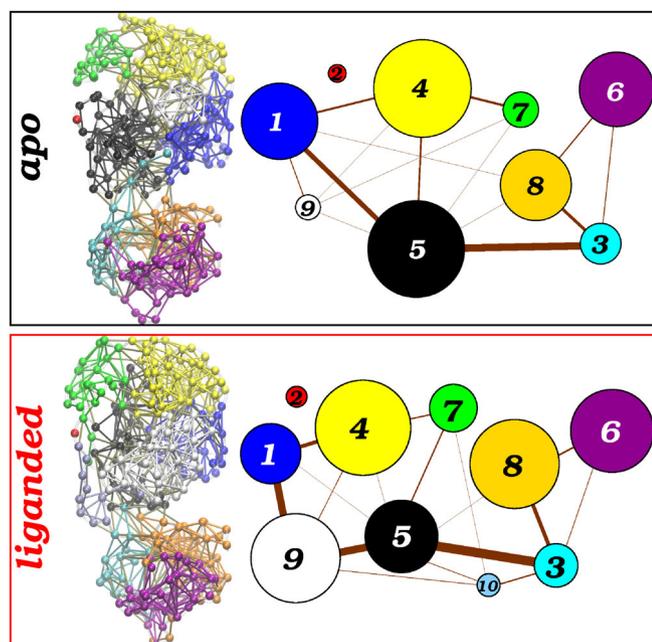
FIGURE 5 Residue communities and schematic community maps for the *apo* and liganded states of M^{Pro}. The C_α atoms of residues in each community are colored corresponding to the community color in the schematic community maps. Communities are labeled 1–9 in the *apo* form and 1–10 in the liganded form. The width of inter-community connections is representative of the combined edge betweenness. We provide additional details about residue memberships and critical nodes in Tables S3 and S4.

when the nodes are within a cutoff distance of 4.5 Å for at least 75% of an MD trajectory.[57] We also used the Girvan–Newman algorithm[57] to identify the community structures in each conformational state. The optimal community structure is chosen using the highest modularity value, which is a measure of intercommunity versus intracommunity edges, so that the maximum modularity value is 1. Typical modularity values are between 0.4 and 0.9.[57] In the *apo* system, the modularity value was 0.6519, and in the covalently bound CCG-50014 system, the modularity value was 0.6509. In a dynamic community network, several edges may exist between communities, referred to as critical edges, and the nodes forming these edges are labeled critical nodes.[35] In the results and discussion sections, C1 refers to community 1, C2 refers to community 2, and so on. We note that C145 refers to cysteine 145 and not to a community.

Ligand dissociation in noncovalently docked simulations
We measured the Euclidean distances between the center of mass of CCG-50014 and the center of mass of all key binding-pocket residues in the active site (H41, M49, G143, C145, H163, H164, P168, and Q189; Figure S7). We considered a distance of 15 Å as the maximum distance before the ligand was considered dissociated from the active site. This distance correlates well with the visual inspection of ligand dissociation.

RMSF per residue in noncovalently docked simulations
We calculated the RMSF of all protein residues while the ligand was still in the active site as defined by the cutoff distance of 15 Å (Figure S8). We also calculated the RMSF of all protein residues after the ligand had dissociated and until the simulation was terminated. The initial coordinates, excluding hydrogens, from docking were used as the reference frame.

2.2 | Experimental methods

2.2.1 | Protein expression and purification

The expression vector, M^{Pro}-3C pET21b(+), containing the full-length SARS-CoV-2 gene encoding M^{Pro} (NC_045512), was purchased from GenScript, with the gene sequence optimized for expression in *Escherichia coli*. The final protein product contains four additional amino acids (GPGG) before the C-terminal 6xHis-tag. Protein expression and purification procedures presented here were adopted and modified from Jin et al.[38] The vector was transformed into *E. coli* Rosetta-gami B(DE3) cells (EMD Millipore), which were grown for 5–6 h in Luria-Bertani (LB) broth with 100 μg/ml ampicillin at 37°C. Once the cells reached OD_{600} 0.7, the overexpression of M^{Pro} was induced by adding 0.5 mM IPTG to the medium. After 11–12 h of incubation at 16°C, the liquid cell culture was centrifuged at 9559g for 25 min (4°C), and the *E. coli* pellets were then stored at −80°C. Protein purification started by thawing out a pellet at 4°C for 30 min. Lysis buffer containing 20 mM Tris–HCl, 150 mM NaCl (pH 7.3), Halt™ Protease Inhibitor cocktail (Thermo Fisher Scientific), and Benzonase (EMD Millipore) was then added to the pellet. After 15–20 min of incubation with Lysis buffer, the cells were fully re-suspended and subjected to four rounds of mechanical cell lysis at 1500 psi using the G-M French Press (Glen Mills). The cell lysate was then centrifuged at 20,217g for 30 min (4°C). The supernatant was filtered through 0.45 and 0.22 μm filters before purification. Since M^{Pro} was expressed with a C-terminal 6xHis-tag, the protein was purified using fast protein liquid chromatography system (FPLC; GE Healthcare AKTA purifier 900) equipped with a HisTrap FF 1 ml column. The protein was eluted from the column using an imidazole concentration gradient going up to 500 mM. The selected protein fractions were pooled, and imidazole was removed using Zeba™ spin desalting columns (following the manufacturer's recommendation; Thermo Fisher Scientific). The sample was brought up to 5 ml before subjecting it to another round of Ni-affinity chromatography using a HisTrap HP 1 ml column. Upon collecting and pooling the selected fractions, the protein sample was exchanged into a buffer containing 50 mM Tris–HCl and 1 mM EDTA (pH 7.3) as described above. The protein purity was assessed using SDS-PAGE. The protein concentration was measured using JASCO V-650 UV–vis spectrophotometer. Upon determining the M^{Pro} purity and concentrations, the protein identity was confirmed using LC–MS/MS mass spectrometry.

2.2.2 | Activity and inhibition assays

Since it was previously shown that C-terminal 6xHis-tag does not alter M^{Pro} activity, the protease activity of uncleaved M^{Pro} containing the C-terminal 6xHis-tag was measured.[58] Both activity and inhibition assays were performed following a modified protocol originally established by Jin et al.[38,59] Fluorescence resonance energy transfer (FRET) peptide Mca-AVLQSGFRK(Dnp)K (GLBiochem) was used as the M^{Pro} substrate (excitation wavelength 325 nm, emission wavelength 392 nm). JASCO FP-8300 spectrofluorometer was used to measure M^{Pro} activity. 1 μM (final concentration) M^{Pro} was mixed with varying concentrations of the peptide substrate (1.87–40 μM). Initial rates were obtained by fitting the linear portion of curves to a straight line (typically around 35–50 s; Kinetic Analysis, JASCO). K_M and V_{max} were calculated from the Lineweaver–Burk plot (1/V vs. 1/[S]) using Kinetic Analysis, JASCO. These values were used to determine the K_{cat} (V_{max}/[M^{Pro}]) as well as the catalytic activity (K_{cat}/K_M). The degree of inhibition was measured for three known M^{Pro} inhibitors[20,38] tideglusib, carmofur, and ebselen (Figure S9); and a novel one, CCG-50014 (Figure 6). The inhibition assays were performed on a microplate reader SpectraMax Me (Molecular Devices). Similarly to the activity assays, Mca-AVLQSGFRK(Dnp)K peptide cleavage by M^{Pro} was measured by monitoring the fluoresence intensity. Each reaction contained 0.2 μM M^{Pro}, 20 μM FRET peptide, varied inhibitor concentration (0.025–100 μM), 5% DMSO in Assay buffer (50 mM Tris–HCl 1 mM EDTA pH 7.3). Data points were recorded every 4 s for 5 min, at 30 °C. SoftMax Pro 7 (Molecular Devices) was used to calculate the initial rates in units of relative fluorescence per second (RFU/s). To demonstrate that CCG-50014 is a specific inhibitor of SARS-CoV-2 M^{Pro}, and not achieving inhibition via potential protein aggregation, we performed additional inhibition assays in the presence of a mild detergent, 0.01% Triton-X 100 (EMD-Millipore). The GraphPad Prism
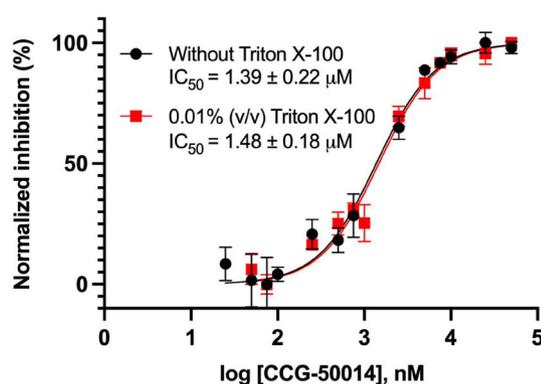
software was used to calculate the IC$_{50}$ values and plot the inhibition curves (Figure 6).

To further probe the binding interaction of CCG-50014, we mixed M^{Pro} and CCG-50014 to a final concentration of 0.9 and 5 μM, respectively, in 50 mM Tris-HCl, 1 mM EDTA, and 5% DMSO buffer of pH 7.3. This reaction mixture was incubated for 2 h at 25°C to ensure adequate time for CCG-50014 to bind to M^{Pro}. After 2 h, aliquots of the reaction were taken out and mixed with Mca-AVLQSGFRK(Dnp)K peptide ranging in concentration (0.95 to 20 μM) and a final concentration of 0.3 μM M^{Pro} in a 96-well plate. The activity assay was performed on the microplate reader SpectraMax Me (Molecular Devices) to determine the initial rate of each reaction. The rest of the reaction mixture was buffer exchanged using Zeba spin desalting columns (Thermo Fisher Scientific). After the buffer exchange, another set of aliquots of the reaction mixture were taken out and mixed with Mca-AVLQSGFRK(Dnp)K peptide using the same range of concentrations, and the initial rate of each reaction was recorded. The remaining reaction mixture was allowed to incubate for another 45 min at 25°C and was again buffer exchanged. Immediately following this second buffer exchange, the reaction mixture was mixed with the same concentrations of Mca-AVLQSGFRK(Dnp)K peptide, and once more, the initial rate of each reaction was measured. The initial rates of each set of assays were calculated in SoftMax Pro 7 (Molecular Devices) in units of relative fluorescence per second (RFU/s) and plotted in GraphPad Prism. All experiments were repeated at least three times. The results from these experiments are shown in Figure S10.

To test whether a common reducing agent, dithiothreitol (DTT), can diminish the inhibitory effect of CCG-50014 on M^{Pro} activity, we mixed the enzyme and the inhibitor CCG-50014 to a final concentration of 0.9 and 5 μM, respectively, in 50 mM Tris-HCl, 1 mM EDTA, 2 mM dithiothreitol (DTT), and 5% DMSO buffer of pH 7.3. Another identical control mixture was made using the same buffer without DTT and was treated in the same manner as the DTT containing sample. Both mixtures were incubated for 2 h at 25°C to ensure adequate time for CCG-50014 to bind. After 2 h, aliquots of the reaction were taken out and mixed with Mca-AVLQSGFRK(Dnp)K peptide ranging in concentration (0.95 to 20 μM) and a final concentration of 0.3 μM M^{Pro} in a 96-well plate. The activity assay was performed on the microplate reader SpectraMax Me (Molecular Devices) to determine the initial rate of each reaction. All experiments were repeated at least three times. The results from these experiments are shown in Figure S11.

FIGURE 6 Normalized inhibition of M^{Pro} by CCG-50014. We show the normalized inhibition percentage of M^{Pro} (y-axis) against the log of the concentration of CCG-50014 (nM, x-axis). Also shown is the IC$_{50}$ value with the error. The IC$_{50}$ value was also determined in the presence of 0.01% Triton X-100 to exclude the possibility of aggregation-based inhibition. See also Figure S9 for the IC$_{50}$ values of three other known inhibitors of M^{Pro}, Figure S10 for binding activity data after buffer exchanges, and Figure S11 for inhibitory effect in the presence of a reducing agent, dithiothreitol.

3 | RESULTS AND DISCUSSION

3.1 | Effect of CCG-50014 binding on the local dynamics of M^{Pro}

Since CCG-50014 is a covalent modifier of cysteine residues, we studied the effect of covalently docking (to C145, the catalytic cysteine residue) of the inhibitor on the shape of the active site, the

stability of key active site residues (M49, G143, H163, H164, E166, P168, and Q189),[60] and the dynamics of the catalytic dyad (C145 and H41). We generated several covalently-bound poses (see Section 2 for details) of CCG-50014 using Schrödinger's Maestro suite, and we show the most energetically favorable pose in Figure 1C. We used the initial coordinates of this covalently-bound pose for three independent MD simulations of M^Pro. We analyzed the root mean squared fluctuations (RMSF), a measure of the deviation of atomic coordinates of residues, and the change in RMSF values (ΔRMSF) between the *apo* and covalently-bound forms of M^Pro. We present the ΔRMSF values when CCG-50014 is covalently bound to M^Pro in Figure 2, and a representation of the binding pocket residues with increased fluctuations in the bound conformation (red) and with decreased fluctuations (blue). We present data from individual simulations and the averaged RMSF in Figure S1.

We observed that the key active site residues M49, H164, E166, P168, and Q189 are stabilized in the presence of CCG-50014 (blue in Figure 2), and several other residues (Y54, G143, H163, and those forming the oxyanion loop; red in Figure 2) exhibited increased fluctuations. The active site residues that are stabilized in the presence of covalently bound CCG-50014 are those residues located closest to the 4-fluorophenyl ring, suggesting that the fluorinated aromatic ring of CCG-50014 is stable in the binding pocket and stabilizes the neighboring residues through several stacking interactions and hydrogen bonds (Figure 3A–D). The active site residues that showed increased fluctuations (G143, H163, and oxyanion loop) in the presence of CCG-50014 are those residues that did not have direct interactions with the aromatic sidechains of CCG-50014. However, these residues are located in the vicinity of other residues that are directly interacting with the inhibitor (e.g., G143 is located near C145, and H163 is located near H164), and as a result, these residues (G143 and H163) show increased fluctuations.

In addition to the residue fluctuation analysis, we investigated the impact of CCG-50014 binding on the positioning of residues defining the active site. To quantify this, we calculated interresidue distances (*d*) between each of the two catalytic dyad residues, C145 and H41, and seven other key residues (M49, G143, H163, H164, E166, P168, and Q189). A negative Δ*d* value for a residue indicates that the residue remained closer to the reference catalytic residue in the inhibitor-bound form compared to the *apo* form. We defined Δ*d* values ≤ |1| Å as minor changes and Δ*d* values > |1| Å as significant changes. In Figure S2A, we present the interresidue Δ*d* values as a histogram for each residue. In Figure S2B,C, we show the snapshots of M^Pro with key residues colored and highlighted based on their Δ*d* values. In Table S2, we present the calculated Δ*d* values of the key active site residues of M^Pro. We report that the residues M49 and Q189 had significantly negative Δ*d* values, and the residues G143, H163, H164, E166, and P168 had minor negative Δ*d* values relative to the residue C145, and minor positive Δ*d* values relative to the residue H41. The Δ*d* values calculated indicate that the movement of M49 and Q189, which drift away from the catalytic dyad in the *apo* MD simulations, was significantly limited in the presence of CCG-50014. Further, since the residues G143, H163, H164, and P168 were closer to the residue

C145, but further away from the residue H41, we observed that the presence of CCG-50014 moved these residues closer in the binding pocket, and forced the catalytic residue H41 away from the key active site residues. Specifically, the residue G143 moved toward the two catalytic dyad residues H41 and C145, while the residue H163 moved toward H41 but away from C145. The movement of G143 toward C145, and of H163 away from C145 potentially contributes to the observed RMSF changes in these two residues (Figure 2).

We further investigated key interactions occurring in the active site to probe the effect of ligand binding. In Figure 1C, we provide the energetically favorable docked pose where we observed the hydrogen bonds between the oxygen atom 1 of CCG-50014 and the residue H163, as well as, between the residue G143 and the oxygen atom 2 of CCG-50014. Further, we observed a π–π stacking interaction with the aromatic ring of the residue H41 and the 4-fluorophenyl ring of CCG-50014. In addition, we observed the solvent exposure of the p-tolyl ring of CCG-50014. In all liganded simulations, we observed that the hydrogen bond formed between the oxygen atom 1 of CCG-50014 and the residue H163 was not stable, and that after the inhibitor CCG-50014 rearranged itself in the active site, the oxygen atom 1 formed a new hydrogen bond with the nitrogen atom in the imidazole ring of the residue H41. We also observed that the oxygen atom 2 of CCG-50014 did not maintain its initial hydrogen bond with the residue G143. The disruption of these interactions potentially further contributes to the observed RMSF changes in G143 and H163 (Figure 2). We find that the π–π stacking interaction between the residue H41 and the p-tolyl ring of CCG-50014 is stable only in one out of three independent MD simulations. In Figure S3, we present the distance between the center of mass of CCG-50014 and the center of mass of the residue H41 in each individual *apo* simulation. We observed in the first simulation that the ligand-residue distance was stable at ∼5 Å, but in the second and third simulations, the distance varied between 5 and 12 Å. We observed the most stable pose in the first MD simulation of the liganded form of M^Pro. In Figure 3, we highlight the key interactions that contributed to the stability of the ligand.

Notably, in the first MD simulation, we observed π–hydrogen bonding interactions between the fluorinated ring of CCG-50014 and the backbone of the residue E166 (Figure 3C). This interaction contributes to the stabilization of the residue E166, consistent with the fluctuation of this residue (Figure 2). In addition, we observed similar π–hydrogen bonding interactions between the backbone of the residue Q189 and the 4-fluorophenyl ring of CCG-50014 (Figure 3A). The hydrogen-bonding interactions between the oxygen atom of Q189 and the second oxygen atom of CCG-50014 (Figure 3B), as well as, the π–hydrogen bonding interaction contributed to the stabilization of the residue Q189 (Figure 2A). In addition, we observed hydrogen bonding between the nitrogen atom of the catalytic residue H41 and the first oxygen atom of CCG-50014 (Figure 3D). The active site residue H164 is stabilized by a hydrogen bond from its nitrogen atom to a backbone atom of the catalytic residue C145, which itself is stabilized due to its covalent bond to CCG-50014. The residue P168 does not directly form any hydrogen bonds with the ligand, but it is stabilized (ΔRMSF = −0.442 Å) by the π–hydrogen bonding interactions

between CCG-50014 and the neighboring residues E166 and Q189, which are located on the same loop as the residue P168.

In second and third MD simulations of the inhibitor-bound form of M^{Pro}, we observed interactions with the key residues M49, Y54, G143, H163, E166, and Q189. Notably, the residue M49, which is located on a flexible loop exhibited the most significant ΔRMSF (−1.78 Å) and Δd (−2.6 Å) when CCG-50014 was bound. We observed that the nitrogen atom of the residue M49 formed a hydrogen bond with the oxygen atom in the residue Y54 (ΔRMSF = −0.189 Å), contributing to the stabilization of the residue M49. In addition, in the *apo* simulations, the loop where M49 is located adopts an "open" conformation away from the active site when the residue H41 adopted a "flipped out" conformation (Figure 4A), but in the presence of CCG-50014 the frequency of this "flipped out" state decreased, reducing the fluctuations of M49 and its average distance to the active site catalytic residues.

In *apo* simulations of M^{Pro}, we observed two unique states of the sidechain of H41. We further investigated the effect of CCG-50014 binding on these two conformational states. In Figure 4, we present the distributions of the key dihedral angle (ϕ) for the residue H41, computed based on three independent simulations, of the *apo* and the inhibitor-bound form of M^{Pro}. In addition, we show key conformations of the residue H41 that are representative of the most populated states in the *apo* form, −174.96° (Figure 4A) and 67.45° (Figure 4B). Further, we also point to key representations that highlight the most populated states of ϕ of the sidechain of the residue H41 in the inhibitor-bound form, −62.73° (Figure 4C) and 62.64° (Figure 4D). In Figures S4 and S5, we show time traces and distributions of ϕ for each individual *apo* and inhibitor-bound MD simulation. In individual *apo* simulations, we observed that the ϕ states of the sidechain of the residue H41 had similar populations and the ϕ values. The two main ϕ states of H41 that we observed represent the "flipped out" state (Figure 4A) and the "flipped in" state (Figure 4B). For reference, the ϕ value in the crystal structure is −44.14°. It is known that the catalytic mechanism of M^{Pro} is initiated by a proton transfer from the residue C145 to the residue H41,[8,9] a process likely dependent on the orientation of the imidazole group of H41.

The presence of CCG-50014 results in another conformational state of $\phi = -62.73°$ (Figure 4C), which is the most populated state for the side-chain of the residue H41. The two other ϕ states observed in the inhibitor-bound simulations, −157.24° and 62.64° (Figure 4D), are similar in orientation to the two states observed in *apo* simulations. The third ϕ state, −62.73° (Figure 4C), is the most prevalent due to (i) the π–π stacking interaction between the 4-fluorophenyl ring of CCG-50014 and the imidazole ring of the residue H41, and (ii) the hydrogen bond between the nitrogen atom of the imidazole ring of the residue H41 and the first oxygen atom of CCG-50014 (Figure 3D), restricting the motion of H41 and potentially inhibiting the proton transfer between the two catalytic residues. To understand how the stabilizing effect of CCG-50014 on neighboring residues and its effect on the ϕ states of the residue H41 impacted the global protein dynamics of M^{Pro}, we performed a dynamic network analysis on the *apo* and liganded forms of M^{Pro}.

3.2 | Effect of CCG-50014 binding on the global dynamics of M^{Pro}

We applied a dynamic network analysis, specifically the Girvan-Newman algorithm[57] (see Section 2), to analyze partitioning of the dynamical conformational space of M^{Pro} into residue communities and an intercommunity network, both in the *apo* and inhibitor-bound forms. In Figure 5, we show the community partitioning of M^{Pro} with/ without the inhibitor and the schematic community maps highlighting the intercommunity connections which are defined by the metric "edge betweenness", where an edge is a connection between two nodes, and the edge betweenness is the number of shortest pathways between a pair of nodes. The larger the edge betweenness value, the more node pathways that pass through it, and the thicker the connecting line is in the schematic community maps (Figure 5). The higher values of edge betweenness indicate increased correlations between communities, and offer insight into how the perturbations from the active site of M^{Pro} may propagate throughout the enzyme. A critical node is a residue that forms an edge, termed a critical edge, with another residue in a different community. A full list of residues in each community and critical nodes is provided in Tables S3 and S4.

In the *apo* form, the active site contains residues from the communities C5, C7, and C9. Specifically, the residues H41 and M49 are assigned to C7, the residues G143 and C145 are assigned to C9, and the residues H163, H164, E166, P168, and Q189 are assigned to C5. In the inhibitor-bound form of M^{Pro}, the active site contains residues from the communities C2, C5, C7, and C9 and the same key active residues are partitioned as follows: H41 (C7), M49 (C7), G143 (C9), C145 (C9), H163 (C9), H164 (C5), E166 (C9), P168 (C9), and Q189 (C2). Notably, the catalytic residues H41 and C145 are both critical nodes in their respective communities with the critical edge between the two catalytic residues in the *apo* form (Table S3). In the inhibitor-bound form of M^{Pro}, the critical edge between the two catalytic residues does not exist (Table S4). The lack of this critical edge indicates that the effect of CCG-50014 on the local dynamics of the catalytic residue also affects global protein dynamics. In the inhibitor-bound form of M^{Pro}, compared to the *apo* form, we observed new connections between the communities C3 and C10, C5 and C10, C7 and C10, C9 and C10; the loss of connections between C1 and C8, C7 and C9; the weakening of connections between C1 and C5, C4 and C5, C4 and C7; and the strengthening of connections between C1 and C9, C5 and C9, C3 and C5, and C3 and C8.

The inhibitor CCG-50014 caused increased fluctuations in all residues located away from the active site, notably in residues 200–240, which are located in domain III and are members of the communities C3, C6, and C8. Domain III is also involved in the dimerization of the enzyme,[7,61] therefore fluctuations of residues in this domain are likely to disrupt the protein–protein interface, indicating an allosteric effect. We used the community schematic maps to analyze these increased fluctuations in domain III. In the *apo* form of M^{Pro}, the perturbations from the catalytic residue H41 (C7) propagate to domain III through correlations in C7 to C5 and C5 to C3, and the community C3 showed correlations with the communities C6 and C8.

In the inhibitor-bound form of M^{Pro}, the correlation between the communities 7 and 5, and communities 5 and 3, increased. In addition, in the liganded-state the community C9, where the catalytic residue C145 is located, increased in size, and had increased correlations to C5. As previously mentioned, the catalytic residue C145, as well as residues 140–170 (found in C7 and C9), showed increased fluctuations when CCG-50014 was bound. Therefore, the increased correlations observed in our dynamic network analysis are consistent with the increased fluctuations in the distant domain III, specifically residues 200–240. In addition, the residues 1–20 and 150–160, all of which are located in C1, showed significant increases in fluctuations, again explained by the increased correlations between the active site communities C1 and C9. A tenth community, C10, was observed in the inhibitor-bound state and contained residues found in domain III, suggesting that CCG-50014 not only causes allosteric perturbations in domain III, but also disrupts the interresidue communication in that domain. Based on these simulation results, we expressed and purified recombinant M^{Pro} (see Section 2) and tested the inhibitory effect of CCG-50014.

3.3 | Inhibitory effect of CCG-50014

We first tested that the SARS-CoV-2 M^{Pro} catalytic efficiency ($K_{cat}/K_M = 27\,900\,M^{-1}\,s^{-1}$) used in this study is comparable to the previously reported value by Jin et al.[38] Then, the IC_{50} values of the three known M^{Pro} inhibitors, carmofur, tideglusib, and ebselen,[62] were calculated (Figure S9). At the time of the initiation of this study, there were only a few known inhibitors of M^{Pro} of SARS-CoV-2 virus, including ebselen, tideglusib, and carmofur.[38] We have chosen ebselen and carmofur as these inhibitors were shown to have the IC_{50} values in the low micromolar range, and tideglusib was specifically chosen due to its structural similarity to CCG-50014. Since CCG-50014 is structurally similar to the known M^{Pro} inhibitor tideglusib, we hypothesized that it could fit inside the substrate-binding pocket. The IC_{50} values of tideglusib (Figure S9A; $1.39 \pm 0.2\,\mu M$) and ebselen (Figure S9B; $0.40 \pm 0.05\,\mu M$) are comparable to reported values,[38] while the calculated IC_{50} value of carmofur (Figure S9C; $4.45 \pm 0.52\,\mu M$) was approximately 2.5-fold higher than the reported value. Importantly, the inhibition experiments showed that CCG-50014 also inhibited M^{Pro} activity. The IC_{50} value of CCG-50014 was $1.39 \pm 0.22\,\mu M$ (Figure 6); thus CCG-50014 exhibited inhibition on a scale comparable to that of other known inhibitors. We have performed the M^{Pro} inhibition assay with CCG-50014 in the presence of a detergent using an assay previously described in the literature,[38,63] to exclude the possibility of CCG-50014 inhibiting the protease activity non-specifically via aggregation. The detergent (0.01% Triton X-100) did not affect these results (Figure 6), which indicates that CCG-50014 is not an aggregate-based inhibitor. We further characterized the M^{Pro} and CCG-50014 interaction in a series of buffer exchanges and in the presence of a disulfide reducing agent. A series of buffer exchanges could not restore M^{Pro} activity by removing the inhibitor CCG-50014 (Figure S10), thus suggesting potential covalent binding. In addition,

the disulfide reducing agent, dithiothreitol (DTT), diminished the inhibitory effect of CCG-50014 on M^{Pro} (Figure S11), further supporting a covalent interaction between CCG-50014 and M^{Pro}.

3.4 | Correlations with other studies

Previous studies[6,30] identified that other covalent and noncovalent inhibitors of M^{Pro} had interactions with the residues H41, G143, C145, H163, and E166. Our energetically favorable docked pose (Figure 1C) showed hydrogen bonds between the inhibitor and the residues G143 and H163. In subsequent simulations, we observed that these hydrogen bonds were not stable and that CCG-50014 instead formed new hydrogen bonds and stacking interactions with the residues H41, M49, E166, and Q189. While a limited work has gone into investigating covalent inhibitors that may form disulfide bonds with the catalytic cysteine residue,[64] other studies have focused on several potential covalent inhibitors[65–71] which may modify the catalytic cysteine residue of M^{Pro} via a C-S bond. The covalent inhibitors investigated in these studies formed interactions with the residues N142, G143, H164, E166, P168, and Q189, the residues that we also observed to interact with CCG-50014 in our work. In Figure S6, we show snapshots of the active site of M^{Pro} with six potential covalent inhibitors[71] to compare our chosen covalently bound conformation against these structures that were obtained by near physiological-temperature crystallography. Most notably, we report that the ligands investigated and CCG-50014 occupy the same region of the active site. Further, other studies[72–74] that investigated the RMSF of residues in the presence of different inhibitors reported typical RMSF values between 1 and 4 Å, similar to our observed results for CCG-50014. However, in our work, we also report on a residue community analysis of the enzyme, which notably suggested that CCG-50014 disrupted inter-community correlations stemming for the catalytic residues C145 and H41. These observations highlight the role of allostery in this enzyme, which could prove useful in designing potential therapeutics targeting the enzyme via an allosteric mechanism.[75–77]

3.5 | Noncovalent docking and MD simulations

Outside of covalent binding of CCG-50014 to the residue C145 in the active site, we also carried out comprehensive non-covalent docking and MD simulations of CCG-50014 (see Section 2 for details). Using the initial coordinates from the top five non-covalent protein/inhibitor complexes, we carried out five independent MD simulations. In Table S5, we present the calculated MM-GBSA values for each pose, with values ranging between -49.98 kcal/mol and -71.51 kcal/mol. In addition, in Figure S7, we provide the initial poses and interactions of the five protein/inhibitor complexes chosen for subsequent simulations. From these poses, we observed hydrogen bonding interactions between CCG-50014 and the residues N142, G143, S144, C145, E166, and Q189. In addition, we report π–π stacking and pi–cation

interactions between the residue H41 and the aromatic rings of CCG-50014. By comparing the initial non-covalent poses and the covalent pose, it is evident that the ligand interacts with the same key residues of the active site. We also calculated the RMSF values of all residues in M^{Pro} for each of the MD simulations (Figure S8).

In all five MD simulations, we observed that the residue M49 was stabilized, similar to the covalently-bound results, with the ΔRMSF values between −0.471Å and −3.19 Å. We also observed that the residue Q189 was stabilized in all MD simulations except in the simulation that used the fifth pose as its initial coordinates. Notably, we observed that the residues M49 and Q189 were both stabilized in the covalently-bound form of M^{Pro}, with the RMSF values of −1.78Å and −0.19 Å, respectively. The fluctuations of all other key active site residues (G143, H163, H164, E166, and P168) were variable between the five simulations conducted with the non-covalently docked poses . In the covalently-bound MD simulations, we observed that the residues H164, E166, and P168 were stabilized and the residues G143 and H163 showed increased fluctuations. A noteworthy observation from MD simulations of non-covalently docked CCG-50014 is that the inhibitor dissociates in all simulations at different time scales, suggesting that the covalent linking to the residue C145 is a mode of stability and inhibition.

For example, the five distinct noncovalent poses (Figure S7) had an average dissociation time of 56.66 ns; pose 1 had the maximum dissociation time of 101.94 ns and pose 2 had the minimum dissociation time of 13.06 ns. In the MD simulation from the first pose, which had the longest dissociation time, we report that the ΔRMSF values of residues M49, H164, E166, P168, and Q189 correlate with the results of the covalently bound simulations. This observation, and the key interactions highlighted in Figure 3A–D, indicate that the stabilization of the residues E166 and Q189 is important for the stability of the inhibitor. Overall, our results suggest that CCG-50014, when present in the binding pocket in a non-covalent conformation or covalently-bound to the residue C145, disrupts the dynamics of M^{Pro}. Our experimental results on the inhibitory effect of CCG-50014 further support a covalent binding mechanism of inhibition of M^{Pro} (Figure 6, Figures S10 and S11). Overall, our computational and experimental studies reveal molecular-level details of interactions underlying inhibition of M^{Pro} by CCG-50014.

AUTHOR CONTRIBUTIONS

Jacob Andrzejczyk: Modeling and computational studies, data analysis, and manuscript writing (Computational). Katarina Jovic, Logan M. Brown, and Valerie G. Pascetta: Experimental studies, data analysis, and manuscript writing (Experimental). Harish Vashisth and Krisztina Varga: Funding acquisition and manuscript editing.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available in the supplementary material of this article

ORCID

Harish Vashisth https://orcid.org/0000-0002-2087-2880

REFERENCES

1. Astuti I, Ysrafil. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2): an overview of viral structure and host response. *Diabetes Metab Syndrome*. 2020;14(4):407-412.
2. Satarker S, Nampoothiri M. Structural proteins in severe acute respiratory syndrome coronavirus-2. *Arch Med Res*. 2020;51(6):482-491.
3. Yadav R, Chaudhary JK, Jain N, et al. Role of structural and nonstructural proteins and therapeutic targets of SARS-CoV-2 for COVID-19. *Cell*. 2021;10(4):821.
4. Qiu Y, Xu K. Functional studies of the coronavirus nonstructural proteins. *STEMedicine*. 2020;1(2):e39.
5. Silva SJR, Silva CTA, Mendes RPG, Pena L, et al. Role of nonstructural proteins in the pathogenesis of SARS-CoV-2. *J Med Virol*. 2020;92(9): 1427-1429.
6. Shitrit A, Zaidman D, Kalid O, et al. Conserved interactions required for inhibition of the main protease of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). *Sci Rep*. 2020;10(1):1-11.
7. Padhi AK, Rath SL, Tripathi T. Accelerating COVID-19 research using molecular dynamics simulation. *J Phys Chem B*. 2021;125(32):9078-9091.
8. Keillor J, Brown R. Attack of zwitterionic ammonium thiolates on a distorted anilide as a model for the acylation of papain by amides. A simple demonstration of a bell-shaped pH/rate profile. *J Am Chem Soc*. 1992;114(21):7983-7989.
9. Świderek K, Moliner V. Revealing the molecular mechanisms of proteolysis of SARS-CoV-2 M pro by QM/MM computational methods. *Chem Sci*. 2020;11(39):10626-10630.
10. Zhang L, Lin D, Sun X, et al. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science*. 2020;368(6489):409-412.
11. Khan SA, Zia K, Ashraf S, Uddin R, Ul-Haq Z. Identification of chymotrypsin-like protease inhibitors of SARS-CoV-2 via integrated computational approach. *J Biomol Struct Dyn*. 2021;39:2607-2616.
12. DiMasi JA. The value of improving the productivity of the drug development process. *Pharmacoeconomics*. 2002;20(3):1-10.

13. Viceconte G, Petrosillo N. COVID-19 R0: magic number or conundrum? *Infect Dis Rep*. 2020;12(1):8516.

14. Strittmatter SM. Overcoming drug development bottlenecks with repurposing: old drugs learn new tricks. *Nature*. 2014;20(6):590-591.

15. Hodos RA, Kidd BA, Shameer K, Readhead BP, Dudley JT. In silico methods for drug repurposing and pharmacology. *Wiley Interdiscip Rev Syst Biol Med*. 2016;8(3):186-210.

16. Pushpakom S, Iorio F, Eyers PA, et al. Drug repurposing: progress, challenges and recommendations. *Nat Rev Drug Discov*. 2019;18(1):41-58.

17. Domínguez JM, Fuertes A, Orozco L, del Monte-Millán M, Delgado E, Medina M. Evidence for irreversible inhibition of glycogen synthase kinase-3β by tideglusib. *J Biol Chem*. 2012;287(2):893-904.

18. del Ser T, Steinwachs KC, Gertz HJ, et al. Treatment of Alzheimer's disease with the GSK-3 inhibitor tideglusib: a pilot study. *J Alzheimers Dis*. 2013;33(1):205-215.

19. Höglinger GU, Huppertz HJ, Wagenpfeil S, et al. Tideglusib reduces progression of brain atrophy in progressive supranuclear palsy in a randomized trial. *Mov Disord*. 2014;29(4):479-487.

20. Ma C, Hu Y, Townsend JA, et al. Ebselen, disulfiram, carmofur, PX-12, tideglusib, and shikonin are nonspecific promiscuous SARS-CoV-2 main protease inhibitors. *ACS Pharmacol Transl Sci*. 2020;3(6):1265-1277.

21. Andújar I, Ríos JL, Giner RM, Recio MC. Pharmacological properties of shikonin—a review of literature since 2002. *Planta Med*. 2013;79(18):1685-1697.

22. Chen X, Yang L, Oppenheim JJ, Howard OZ. Cellular pharmacology studies of shikonin derivatives. *Phytother Res*. 2002;16(3):199-209.

23. Sankawa U, Ebizuka Y, Miyazaki T, et al. Antitumor activity of shikonin and its derivatives. *Chem Pharm Bull*. 1977;25(9):2392-2395.

24. Rai H, Barik A, Singh YP, et al. Molecular docking, binding mode analysis, molecular dynamics, and prediction of ADMET/toxicity properties of selective potential antiviral agents against SARS-CoV-2 main protease: an effort toward drug repurposing to combat COVID-19. *Mol Divers*. 2021;25:1905-1927.

25. Das S, Sarmah S, Lyndem S, Singha RA. An investigation into the identification of potential inhibitors of SARS-CoV-2 main protease using molecular docking study. *J Biomol Struc Dyn*. 2020;39:3347-3357.

26. Abel R, Paredes Ramos M, Chen Q, et al. Computational prediction of potential inhibitors of the main protease of SARS-CoV-2. *Front Chem*. 2020;8:1162.

27. Keretsu S, Bhujbal SP, Cho SJ. Rational approach toward COVID-19 main protease inhibitors via molecular docking, molecular dynamics simulation and free energy calculation. *Sci Rep*. 2020;10(1):1-14.

28. Zhang L, Lin D, Kusov Y, et al. α-Ketoamides as broad-spectrum inhibitors of coronavirus and enterovirus replication: structure-based design, synthesis, and activity assessment. *J Med Chem*. 2020;63(9):4562-4578.

29. Jiménez-Alberto A, Ribas-Aparicio RM, Aparicio-Ozores G, Castelán-Vega JA. Virtual screening of approved drugs as potential SARS-CoV-2 main protease inhibitors. *Comput Biol Chem*. 2020;88:107325.

30. Mittal L, Kumari A, Srivastava M, Singh M, Asthana S. Identification of potential molecules against COVID-19 main protease through structure-guided virtual screening approach. *J Biomol Struc Dyn*. 2021;39(10):3662-3680.

31. Bolelli K, Ertan-Bolelli T, Unsalan O, Altunayar-Unsalan C. Fenoterol and dobutamine as SARS-CoV-2 main protease inhibitors: a virtual screening study. *J Mol Struct*. 2021;1228:129449.

32. Kneller DW, Phillips G, O'Neill HM, et al. Room-temperature X-ray crystallography reveals the oxidation and reactivity of cysteine residues in SARS-CoV-2 3CL Mpro: insights into enzyme mechanism and drug design. *IUCrJ*. 2020;7(6):1028-1035.

33. Shaw VS, Mohammadiarani H, Vashisth H, Neubig RR. Differential protein dynamics of regulators of G-protein signaling: role in specificity of small-molecule inhibitors. *J Am Chem Soc*. 2018;140(9):3454-3460.

34. Blazer LL, Zhang H, Casey EM, Husbands SM, Neubig RR. A nanomolar-potency small molecule inhibitor of regulator of G-protein signaling proteins. *Biochemistry*. 2011;50(15):3181-3192.

35. Liu Y, Vashisth H. Allosteric pathways originating at cysteine residues in regulators of G-protein signaling proteins. *Biophys J*. 2021;120(3):517-526.

36. Shaw VS, Mohammadi M, Quinn JA, Vashisth H, Neubig RR. An inter-helical salt bridge controls flexibility and inhibitor potency for regulators of G-protein signaling proteins 4, 8, and 19. *Mol Pharmacol*. 2019;96(6):683-691.

37. Mohammadi M, Mohammadiarani H, Shaw VS, Neubig RR, Vashisth H. Interplay of cysteine exposure and global protein dynamics in small-molecule recognition by a regulator of G-protein signaling protein. *Proteins*. 2019;87(2):146-156.

38. Jin Z, Du X, Xu Y, et al. Structure of M pro from SARS-CoV-2 and discovery of its inhibitors. *Nature*. 2020;582(7811):289-293.

39. Schrödinger, LLC, New York, NY, Maestro; 2020.

40. Sastry GM, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des*. 2013;27(3):221-234.

41. Phillips JC, Braun R, Wang W, et al. Scalable molecular dynamics with NAMD. *J Comput Chem*. 2005;26(16):1781-1802.

42. Vanommeslaeghe K, Hatcher E, Acharya C, et al. CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem*. 2010;31(4):671-690.

43. Lee J, Cheng X, Swails JM, et al. CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *J Chem Theory Comput*. 2016;12(1):405-413.

44. Mark P, Nilsson L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *J Phys Chem*. 2001;105(43):9954-9960.

45. Zhu K, Borrelli KW, Greenwood JR, et al. Docking covalent inhibitors: a parameter free approach to pose prediction and scoring. *J Chem Inf Model*. 2014;54(7):1932-1940.

46. Li J, Abel R, Zhu K, Cao Y, Zhao S, Friesner RA. The VSGB 2.0 model: a next generation energy model for high resolution protein structure modeling. *Proteins*. 2011;79(10):2794-2812.

47. Halgren T. New method for fast and accurate binding-site identification and analysis. *Chem Biol Drug Des*. 2007;69(2):146-148.

48. Friesner RA, Murphy RB, Repasky MP, et al. Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J Med Chem*. 2006;49(21):6177-6196.

49. Sethi A, Eargle J, Black AA, Luthey-Schulten Z. Dynamical networks in tRNA: protein complexes. *Proc Natl Acad Sci U S A*. 2009;106(16):6620-6625.

50. VanWart AT, Eargle J, Luthey-Schulten Z, Amaro RE. Exploring residue component contributions to dynamical network models of allostery. *J Chem Theory Comput*. 2012;8(8):2949-2961.

51. Chopra N, Wales TE, Joseph RE, et al. Dynamic allostery mediated by a conserved tryptophan in the Tec family kinases. *PLoS Comput Biol*. 2016;12(3):e1004826.

52. Kornev AP, Taylor SS. Dynamics-driven allostery in protein kinases. *Trends Biochem Sci*. 2015;40(11):628-647.

53. McClendon CL, Kornev AP, Gilson MK, Taylor SS. Dynamic architecture of a protein kinase. *Proc Natl Acad Sci U S A*. 2014;111(43):E4623-E4631.

54. Ahuja LG, Kornev AP, McClendon CL, Veglia G, Taylor SS. Mutation of a kinase allosteric node uncouples dynamics linked to phosphotransfer. *Proc Natl Acad Sci U S A*. 2017;114(6):E931-E940.

55. Srivastava AK, McDonald LR, Cembran A, et al. Synchronous opening and closing motions are essential for cAMP-dependent protein kinase A signaling. *Structure*. 2014;22(12):1735-1743.

56. Glykos NM. Software news and updates carma: a molecular dynamics analysis program. *J Comput Chem*. 2006;27(14):1765-1768.

57. Girvan M, Newman ME. Community structure in social and biological networks. *Proc Natl Acad Sci U S A*. 2002;99(12):7821-7826.

58. Sacco MD, Ma C, Lagarias P, et al. Structure and inhibition of the SARS-CoV-2 main protease reveal strategy for developing dual inhibitors against Mpro and cathepsin L. *Sci Adv*. 2020;6(50):eabe0751.

59. Anand K, Ziebuhr J, Wadhwani P, Mesters JR, Hilgenfeld R. Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs. *Science*. 2003;300(5626):1763-1767.

60. Padhi AK, Tripathi T. Targeted design of drug binding sites in the main protease of SARS-CoV-2 reveals potential signatures of adaptation. *Biochem Biophys Res Commun*. 2021;555:147-153.

61. Mengist HM, Dilnessa T, Jin T. Structural basis of potential inhibitors targeting SARS-CoV-2 main protease. *Front Chem*. 2021;9:622898.

62. Kneller DW, Phillips G, O'Neill HM, et al. Structural plasticity of SARS-CoV-2 3CL M pro active site cavity revealed by room temperature X-ray crystallography. *Nat Commun*. 2020;11(1):3202.

63. Feng BY, Shoichet BK. A detergent-based assay for the detection of promiscuous inhibitors. *Nat Protoc*. 2006;1(2):550-553.

64. Niesor EJ, Boivin G, Rhéaume E, et al. Inhibition of the 3CL protease and SARS-CoV-2 replication by dalcetrapib. *ACS Omega*. 2021;6(25): 16584-16591.

65. Paul AS, Islam R, Parves MR, et al. Cysteine focused covalent inhibitors against the main protease of SARS-CoV-2. *J Biomol Struct Dyn*. 2020;40:1639-1658.

66. Soulère L, Barbier T, Queneau Y. Docking-based virtual screening studies aiming at the covalent inhibition of SARS-CoV-2 MPro by targeting the cysteine 145. *Comput Biol Chem*. 2021;92:107463.

67. Hassab MAE, Fares M, Amin MK, et al. Toward the identification of potential α-Ketoamide covalent inhibitors for SARS-CoV-2 Main protease: fragment-based drug design and MM-PBSA calculations. *Processes*. 2021;9(6):1004.

68. Hoffman RL, Kania RS, Brothers MA, et al. Discovery of ketone-based covalent inhibitors of coronavirus 3CL proteases for the potential therapeutic treatment of COVID-19. *J Med Chem*. 2020;63(21): 12725-12747.

69. Banerjee R, Perera L, Tillekeratne LV. Potential SARS-CoV-2 main protease inhibitors. *Drug Discov Today*. 2021;26(3):804-806.

70. Chen CC, Yu X, Kuo CJ, et al. Overview of antiviral drug candidates targeting coronaviral 3C-like main proteases. *FEBS Lett*. 2021; 288(17):5089-5121.

71. Durdagi S, Dağ Ç, Dogan B, et al. Near-physiological-temperature serial crystallography reveals conformations of SARS-CoV-2 main protease active site for improved drug repurposing. *Structure*. 2021; 29:1382-1396.

72. Mishra D, Maurya RR, Kumar K, et al. Structurally modified compounds of hydroxychloroquine, remdesivir and tetrahydrocannabinol against main protease of SARS-CoV-2, a possible hope for COVID-19: docking and molecular dynamics simulation studies. *J Mol Liq*. 2021;335:116185.

73. Hosseini FS, Amanlou M. Anti-HCV and anti-malaria agent, potential candidates to repurpose for coronavirus infection: virtual screening, molecular docking, and molecular dynamics simulation study. *Life Sci*. 2020;258:118205.

74. Ancy I, Sivanandam M, Kumaradhas P. Possibility of HIV-1 protease inhibitors-clinical trial drugs as repurposed drugs for SARS-CoV-2 main protease: a molecular docking, molecular dynamics and binding free energy simulation study. *J Biomol Struct Dyn*. 2021;39:5368-5375.

75. Sun Z, Wang L, Li X, et al. An extended conformation of SARS-CoV-2 main protease reveals allosteric targets. *Proc Natl Acad Sci U S A*. 2022;119(15):e2120913119.

76. El-Baba TJ, Lutomski CA, Kantsadi AL, et al. Allosteric inhibition of the SARS-CoV-2 Main protease: insights from mass spectrometry based assays. *Angew Chem Int Ed*. 2020;59(52):23544-23548.

77. DasGupta D, Chan WK, Carlson HA. Computational identification of possible allosteric sites and modulators of the SARS-CoV-2 Main protease. *J Chem Inf Model*. 2022;62(3):618-626.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.