communications biology

ARTICLE

https://doi.org/10.1038/s42003-022-03084-7 OPEN

Modulation of the monomer-dimer equilibrium and catalytic activity of SARS-CoV-2 main protease by a transition-state analog inhibitor

Nashaat T. Nashed[1], Annie Aniana[1], Rodolfo Ghirlando[2], Sai Chaitanya Chiliveri[1] & John M. Louis[1✉]

The role of dimer formation for the onset of catalytic activity of SARS-CoV-2 main protease ($MPro^{WT}$) was assessed using a predominantly monomeric mutant ($MPro^{M}$). Rates of $MPro^{WT}$ and $MPro^{M}$ catalyzed hydrolyses display substrate saturation kinetics and second-order dependency on the protein concentration. The addition of the prodrug GC376, an inhibitor of $MPro^{WT}$, to $MPro^{M}$ leads to an increase in the dimer population and catalytic activity with increasing inhibitor concentration. The activity reaches a maximum corresponding to a dimer population in which one active site is occupied by the inhibitor and the other is available for catalytic activity. This phase is followed by a decrease in catalytic activity due to the inhibitor competing with the substrate. Detailed kinetics and equilibrium analyses are presented and a modified Michaelis-Menten equation accounts for the results. These observations provide conclusive evidence that dimer formation is coupled to catalytic activity represented by two equivalent active sites.

[1] Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD, USA. [2] Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD, USA. ✉email: johnl@niddk.nih.gov

In Severe Acute Respiratory Syndrome CoronaVirus 2 (SARS-CoV-2), which causes the COronaVIrus Disease 2019 (COVID-19)[1,2], and its closely related SARS-CoV, the function of the main protease (MPro) is indispensable for its replication and propagation[1,3,4]. In its genome, a single copy of MPro is encoded within the polyproteins (pp) 1a and 1ab[4,5]. The active MPro functions as a homodimer to mediate its own release at its termini and processing of the polyproteins at various sites to generate the non-structural proteins nsp4 through nsp16 required for the assembly of the viral replication/transcription complex[3,6,7]. Thus, in addition to effective vaccines targeting the spike protein[8], MPro is a potential target for the development of antiviral agents for the treatment of SARS-CoV-2 infection[3,9,10]. An active site inhibitor (PF-7321332) of MPro is currently in clinical trials for the treatment of COVID-19[11].

MPro is composed of 306 amino acids comprising three domains. Domains I (residues 8–101) and II (residues 102–184) together exhibit a chymotrypsin-like fold, and domain III (residues 201–306) comprises a cluster of five alpha-helices connected to domain II by a long loop (residues 185–200)[3,7]. Each subunit of MPro harbors an active site consisting of the catalytic dyad H41-C145. Cleavage at the N-terminus of MPro has been proposed to modulate the quaternary structure for catalytic activity through extensive inter- and intra-subunit contacts formed by the free N-terminal strand with domains II and III[3,7,12–15]. The released fully active wild-type mature MPro, henceforth referred to as $MPro^{WT}$, exhibits a dissociation constant in the low micromolar range of 0.1 to 15[3,16–20] capable of cleaving polyprotein and synthetic peptide substrates. Deletion of the N-terminal residues (termed the N-finger) and domain III lead to a shift in the monomer-dimer equilibrium towards the monomer form accompanied by a drastic decrease in catalytic activity[17,21–25]. Various mutational analysis of SARS-CoV MPro and structural requirements for its regulation are summarized in references[20,26]. Despite the monomer form adopting a native-like tertiary fold, as shown for various mutations or deletions in the sequence, monomeric variants of MPro are reported to exhibit very low or no catalytic activity[7,17,21,26–33]. This has been attributed to a collapsed active site which impairs the binding of Q-P1 of the substrate in the S1 subsite leading to loss of catalytic function[28,29,32]. Specifically, in the monomeric structure the loop comprising residues S139 to L141 was shown to transform into a 3_{10}-helix such that the rearranged N142 interacting with E166 blocks entry to the S1 subsite. This is consistent with the observation that mutation E166A also impairs substrate binding[17,20]. Single mutations of R4, M6, G11, S139, E290, and R298 lead to increased dimer dissociation[20,27–29,31,32]. Specifically, inter-subunit contacts mediated by a salt bridge between residues R4 and E290[27] and an aromatic-hydrophobic interaction between Y126 and M6, as well as an intra-subunit hydrogen bond between the side chain NH2 of R298 with backbone oxygen of M6, are shown to be essential for maintaining dimerization[20,28]. Thus, to closely examine the role of dimerization for the onset of catalytic activity, we took advantage of critical contacts mediated by residues E290 and R298 in domain III, mutations of which substituted to Ala lead to a significant increase in the dimer dissociation constant ($K_d = K_1$). The resulting predominantly monomeric construct allowed for a detailed examination of the relationship between dimer formation and the kinetics and inhibition of MPro-catalyzed hydrolysis using the transition-state analog inhibitor GC376 that modulates the relative composition of the monomer and dimer form of the enzyme. These studies provide conclusive evidence for the role of dimer formation and its associated thermodynamic stability being pivotal for the appearance of mature-like catalytic activity, with two equivalent active sites.

Results

Preparation and characterization of $MPro^M$. A mini precursor of MPro (termed $^{+25}MPro^M$-6His) containing 25 amino acids of flanking nsp4 sequence, substitution mutations E290A and R298A in domain III and a C-terminal 6His-Tag was constructed (Fig. 1 and Supplementary Fig. 1). Consistent with an earlier observation of an analogous model precursor construct of $MPro^{WT}$ or with the single R298E mutation in MPro[7], expression of $^{+25}MPro^M$-6His results in its maturation at the nsp4/nsp5 junction and accumulation of mature $MPro^M$-6His. The 6His-Tag which permits facile purification was subsequently removed using human rhinovirus (HRV-3C) protease as described[34] yielding $MPro^M$ (Supplementary Fig. 1).

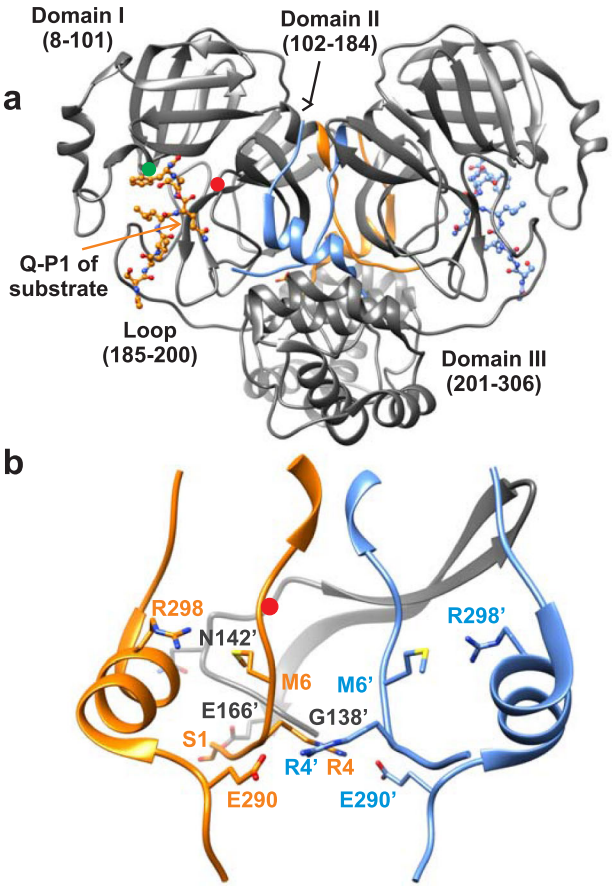

Fig. 1 Molecular representation of the SARS-CoV-2 main protease dimer and critical interactions which influence the monomer-dimer (M-D) equilibrium. a The two subunits of the mature dimer (PDB ID: 7N89[54]) with N- (1–11) and C-terminal (288–306) residues highlighted with subunit A in orange and subunit B in blue. Substrate SAVLQSGF bound to the active site of each subunit is shown with Q-P1 which forms the S1 subsite indicated by the orange arrow. The interface formed by the free N-finger residues 1–7, shown in the middle of the dimer, is critical for dimer stability. The dimer interface is formed by an extensive network of hydrogen bonds and hydrophobic interactions involving N-terminal residues 1′-16′, β-strand residues 118′-125′ and loop residues 137′-142′ as illustrated in reference[15]. Red and green circles denote position of active site C145 and H41 residues, respectively. **b** Enlargement of the region showing residue positions with the same coloring as in (A) critical for dimer interface stability. Mutations E290A and R298A increase the K_d by ~5000-fold based on our estimate shown in Table 1, and published reports[3,17,18]. The region (G138 to E166) encompassing the oxyanion loop (S139 to L141) for subunit B are shown in gray. S1 of subunit A interacts with E166 of subunit B[17]. Residues from subunit B are denoted with prime (′).

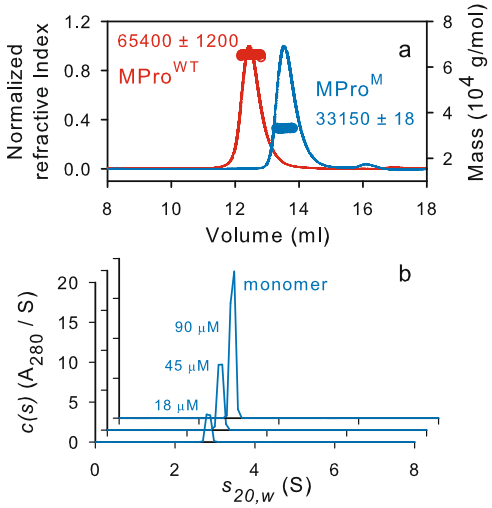

Fig. 2 Molecular mass estimation of MPro^{M}. a Mass estimation by SEC-MALS by injecting 125 µL of MPro^{M} and MPro^{WT} at ~58 µM and 30 µM, respectively. **b** SV-AUC absorbance *c(s)* distributions at concentrations ranging from 18 to 90 µM of MPro^{M}. SEC-MALS and SV-AUC were carried out in buffer A at 25 °C.

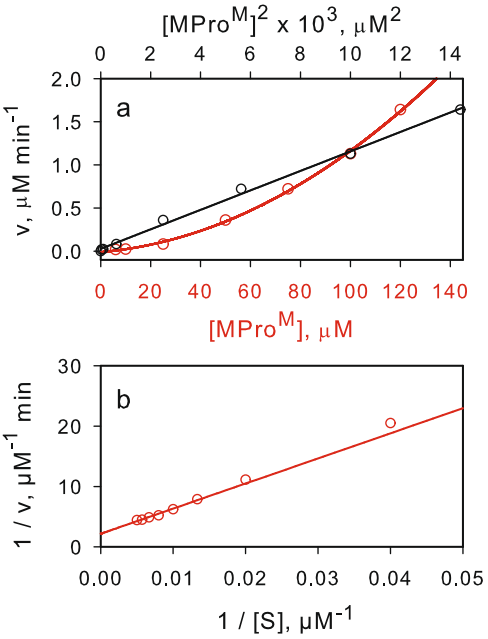

Fig. 3 Evaluation of the catalytic efficiency of MPro^{M}. a Non-linear relationship between the rate of catalyzed hydrolysis *vs* the protein concentration (red line), and the linear relationship between the rate of catalyzed hydrolysis *vs* the square of the protein concentration (black line). **b** Lineweaver-Burk plot for the catalyzed hydrolysis at a concentration of 40 µM MPro^{M}.

Processed (mature) MPro^{M} was purified using established methods (Supplementary Fig. 1) from cells induced for expression for 2–3 h. MPro^{M} showed an estimated mass of 33653 Da by ESI-MS (calculated = 33654 Da) and a single peak corresponding to a monomer mass of 33.2 kDa by size exclusion chromatography coupled with multi-angle light scattering (SEC-MALS, Fig. 2a). Sedimentation velocity analytical ultracentrifugation (SV-AUC), which precludes dilution during the experiment, at concentrations ranging from 18–90 µM clearly shows that MPro^{M} is mainly monomeric with no detectable dimer form. As shown in Fig. 2b, a single species of 2.84 S with an estimated mass of 33 kDa was observed up to 90 µM. This result is consistent with such mutations introduced in SARS-CoV main protease influencing the M-D equilibrium (Fig. 1b)[7,17,26–28]. In contrast, mature MPro^{WT} elutes as a dimer exhibiting a mass of 65.4 kDa by SEC-MALS as expected (Fig. 2a).

Kinetics of MPro^{WT} and MPro^{M}-catalyzed hydrolyses. MPro^{WT} and MPro^{M} catalyze the hydrolyses of a known fluorescence resonance energy transfer (FRET) peptide substrate[3,35–37] corresponding to the nsp4/nsp5 cleavage site sequence in pp1a polyprotein[3]. The rates of hydrolyses at a final substrate concentration of 200 µM display linear relationships with the square of the concentration of MPro^{WT} and MPro^{M} with intercepts at the origin [multiple correlation coefficient (R) = 0.9972, Supplementary Fig. 2a and R = 0.9984, Fig. 3a, respectively]. Second-order dependency on protein concentration provides clear evidence that the dimeric form of the enzyme is required for catalytic activity. These observations are in accordance with an earlier report of the monomeric R290A and R290L mutants of SARS-CoV main protease[17], which shares 95% sequence identity to MPro^{WT} [3]. The rates of MPro^{WT}- and MPro^{M}-catalyzed hydrolyses display good substrate saturation kinetics at constant protein concentrations (Fig. 3b and Supplementary Fig. 2b) and the resulting v_{max} and K_s are listed in Table 1 including the total initial monomer concentration (M_o).

Scheme in Fig. 4 is proposed to account for the observed results (see Supplementary Note 1).

It can be shown that:

$$v = \frac{k_3\{[M_o] - [M]\}[S]}{K_s + [S]} \quad (1)$$

where $[M_o]$ and [M] denote the total and free monomer concentrations, respectively. For MPro^{WT}, K_1 was determined by the kinetic method described previously (Supplementary Fig. 2c)[38,39]. The obtained K_1 value is 1.32 ± 0.2 µM, which is consistent with the earlier reported estimate by SV-AUC under similar buffer conditions for MPro^{WT} [3,17] and for SARS-CoV MPro^{WT} [3,17] was used to calculate the free monomer concentration. A plot of v *vs* $\{[M_o]-[M]\}$ is linear (R = 0.9979, Supplementary Fig. 2d) with an intercept at the origin indicating that enzymatic activity is observed only from the dimeric form of MPro^{WT}. The kinetic parameters and M-D equilibrium constant are listed in Table 1.

For MPro^{M}, the value of K_1 is much larger than 90 µM and thus could not be determined directly. Since $[M_o] \approx [M] >>> [D]$, $\{[M_o]-[M]\} = 2[D]$, and $[D] = [M]^2/K_1$, substituting $\{[M_o]-[M]\}$ with the value of 2[D] and $[M]^2$ with $[M_o]^2$ and rewriting Eq. 1 gives:

$$v = \frac{2k_3[S]\left\{\frac{[M_o]^2}{K_1}\right\}}{K_s + [S]} \quad (2)$$

As indicated above, a plot of v *vs* $[M_o]^2$ at substrate concentration of 200 µM is a straight line. From a plot of 1/v *vs* 1/[S] (Fig. 3b), the obtained K_s of 193 µM is about the substrate concentration used and thus, the slope of the line is:

$$\frac{k_3[S]}{K_1 K_s} = (1.0 \pm 0.03)10^{-4}\,\mu M^{-1} min^{-1} \quad (3)$$

Table 1 Kinetic, thermodynamic and inhibition parameters of MPro^{WT} and MPro^{M}-catalyzed hydrolysis of peptide substrate[a].

Protein	M_o μM	K_1 μM	V_{max} μM min^{-1}	k_{cat} min^{-1}	$K_m = K_s$ μM	$(k_{cat}/K_m)_{obs}$ μM^{-1} min^{-1}	k_3/K_s μM^{-1} min^{-1}	-ΔG kcal/mol	K_i μM
MPro^{WT}	0.2	1.3 ± 0.2	2.1 ± 0.1	54 ± 3	91 ± 2	0.6 ± 0.05	0.6 ± 0.05	0.306	0.15[b]
MPro^{M}	40	6600[b]	0.46 ± 0.04	0.97 ± 0.1	193 ± 29	(5 ± 0.6)10^{-3}	5 × 10^{-3}	3.2	
MPro^{M} + GC376	10					(1.3 ± 0.05)10^{-3}	3.4 × 10^{-3}[c]	3.4	6.2[d]

[a]Protease assays were carried out in buffer B (25 mM Tris-HCl, pH 7, 50 mM NaCl and 1 mM TCEP) at 28 °C. [b]Determined by ITC. [c]The value of k_{cat}/K_m is calculated from the observed k_{cat}/K_m at GC376 concentration of 10 μM and $[M_o]$ of 10 μM using the expression: $(k_{cat}/K_m) = k_3/2K_s(1 + [I]/K_i)$ (see Eq. 4). [d]The binding constant is the average of binding constants obtained from inhibition study, ITC, and SV-AUC. -ΔG values were calculated from the k_3/K_s values.

$$2M \underset{}{\overset{K_1}{\rightleftharpoons}} D + S \overset{K_s}{\rightleftharpoons} D.S \xrightarrow{k_3} \text{products}$$

Fig. 4 Mechanism of catalysis by MPro^{WT} and MPro^{M}. M, D and S denote monomer, dimer and substrate, respectively.

Modulation of the catalytic activity of MPro^{M} by feline coronavirus prodrug GC376. Several clinical drugs have been repurposed for developing rapid therapeutic intervention of COVID-19[35,36,40,41]. Prodrug GC376 elicits a broad-spectrum activity against human and animal coronaviruses including the recent SARS-CoV-2 MPro^{WT} and its replication[40–44]. In aqueous medium, the prodrug disproportionates to a sulfite ion and the aldehyde GC373 (Fig. 5a). GC373 inhibits MPro by reversibly binding and forming a covalent bond between the sulfur of C145 and the carbonyl carbon of GC373 to yield hemithioacetal, a transition state analog. Henceforth, GC376 and GC373 are used herein interchangeably. The 3D structures of MPro^{WT}-GC373 complexes were described recently[40,41]. The prodrug GC376 was chosen to examine its interaction with MPro^{M} because it is thermally more stable than other known inhibitor complexes by as much as 10.3 °C indicative of its higher affinity[40,45].

Surprisingly, the rate of hydrolysis catalyzed by 10 μM MPro^{M} increases with the increasing inhibitor concentration reaching a maximum at 10 μM followed by a decrease above 10 μM inhibitor (Fig. 5b). At a final concentration of 10 μM each of GC376 and MPro^{M}, K_m is much larger than the solubility limit of the substrate, and the observed first-order rate constant v_{max}/K_m is $(1.3 \pm 0.2) \times 10^{-3}$ min^{-1} in buffer B at 28 °C is obtained from the linear plot of rate *vs* [S] (R = 0.9984), shown in Fig. 5c.

Modulation of the M-D equilibrium of MPro^{M} and competitive inhibition by GC376. Since the rate of MPro^{M}-catalyzed hydrolysis displays a second-order dependency on the enzyme concentration indicating a protein dimer is required for catalytic activity, we suspected that GC376 influenced the M-D equilibrium. A series of SV analyses were carried out in the presence of varied inhibitor concentrations. Figure 6a shows SV absorbance c(s) distributions of 6–7 μM MPro^{M} at increasing concentrations of GC376. In addition to the monomeric form of the protein, a second species corresponding to the dimer form (4.47 S, 58 kDa) is observed. The amount of the dimer form increases with a corresponding decrease in the monomer form indicating a dynamic equilibrium. Figure 6b (black line) shows the decrease in the amount of the monomeric species with increasing GC376 concentration. The data can be fitted to the equation [M] + [I] ⇋ [MI], assuming a single binding site. A plot of {$[M_o]$-[M]} vs [M] is linear (Fig. 6b, red trace), where $[M_o]$ and [M] denote the total protein and monomer concentrations, respectively. This result indicates that both the active sites of the dimer are functional and equivalent with a calculated K_b of 6.7 ± 0.2 μM. CD spectra (Fig. 6c) at 10 μM MPro^{M} in the absence (monomer) and 10-fold molar excess GC376, where MPro^{M} is ~90% dimeric, are nearly identical indicating a folded monomer with a secondary structure that resembles the dimer, which is consistent with earlier observations reported for SARS-CoV main protease by CD and NMR[7,21,22,31].

The binding constant of GC376 to MPro^{M} was also determined by isothermal titration calorimetry (ITC). The isotherms are shown in Fig. 7 for MPro^{WT} and MPro^{M} and the binding constants and thermodynamic parameters are listed in Supplementary Table 1. The titration fits a single site binding model with nearly a 1:1 stoichiometry of MPro^{M} to inhibitor and an estimated K_b of 6.1 ± 0.3 μM in buffer C at 28 °C consistent with the estimated K_b value determined by SV-AUC analysis. The binding constant of GC376 to MPro^{M} is 41 times larger relative to MPro^{WT} which is indicative of the MPro^{WT}/GC376 complex being more stable by about 2.2 kcal/mol.

Scheme in Fig. 8 is proposed to account for the rise and fall in the catalytic activity of MPro^{M} upon increasing the inhibitor concentration (Fig. 5b).

The second-order dependency of the rate of hydrolysis on MPro^{M} concentration points to the presence of undetectable amounts of dimer by SV-AUC up to 90 μM (Fig. 2b). The concomitant increase of the dimer population and catalytic activity upon binding GC373 further supports the requirement of a dimer for catalytic activity. Importantly, the binding of the transition state analog GC373 requires a functional active site. X-ray crystallographic and NMR solution studies of MPro^{WT}-GC373 complex show that GC373 reversibly and covalently modifies the active site sulfur atom of the nucleophilic thiol of Cys145 by forming a hemithioacetal[40,41]. Therefore, the two active sites of a dimer must be equivalent and simultaneously functional contrary to the notion suggested for SARS-CoV main protease dimer that only one protomer is active at a time[26,46]. Thus, GC373 binds to a small population of the dimer leading to a shift in the equilibrium composition favoring the dimer. For MPro^{M} in the presence of GC373, the catalytically reactive species is a dimer containing one molecule of GC373 bound to one of the two active sites (DI) leaving the other available for catalytic function. It can be shown from the scheme that (see Supplementary Note 2):

$$v = \frac{K_3[S]\{[M_o] - [M]\}}{2K_s\left\{1 + \frac{[I]}{K_i}\right\} + [S]} \quad (4)$$

where M_o is the total protein and M is the amount of the monomeric form.

The values of [M] at various concentrations (M_o) are quantified by SV-AUC from the results shown in Fig. 6a, b. The K_i value is determined by evaluating inhibition at concentrations of GC376 at 10 μM and above. Because the maximal catalytic activity of MPro^{M} in the presence of GC376 is at 10 μM, the observed K_m/k_{cat} values were plotted versus ([I] – 10) (see Fig. 6d and Supplementary Table 2). The calculated K_i of 5.9 ± 3 μM is within the experimental error to the values determined by SV-AUC and ITC. Because GC373 is known to bind reversibly[40,41] and a single binding isotherm is observed in ITC both for MPro^{WT} and MPro^{M}, GC373 must bind to the same form of the protein at the active site of MPro^{M} and compete with the substrate. Since k_{cat}/K_m for MPro^{WT} is k_3/K_s, the value of $k_3/$

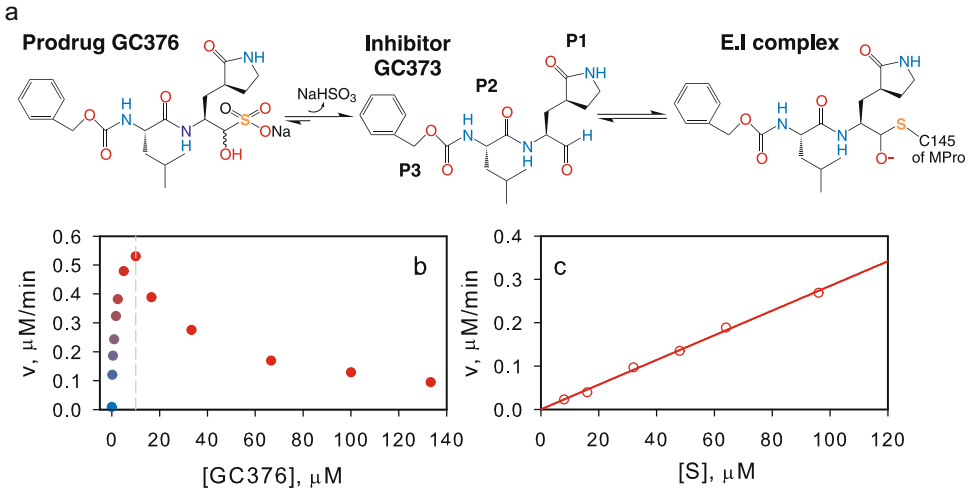

Fig. 5 Modulation of the catalytic activity of MPro^{M} by feline coronavirus prodrug GC376. a Chemical structure of GC376 and steps in its binding to the active site of 3C and 3CL proteases[40,41]. **b** Catalytic activity of 10 μM MPro^{M} as a function of increasing GC376 concentration. **c** A plot of the rate *vs* substrate concentration at a final concentration of 10 μM MPro^{M} and 10 μM GC376.

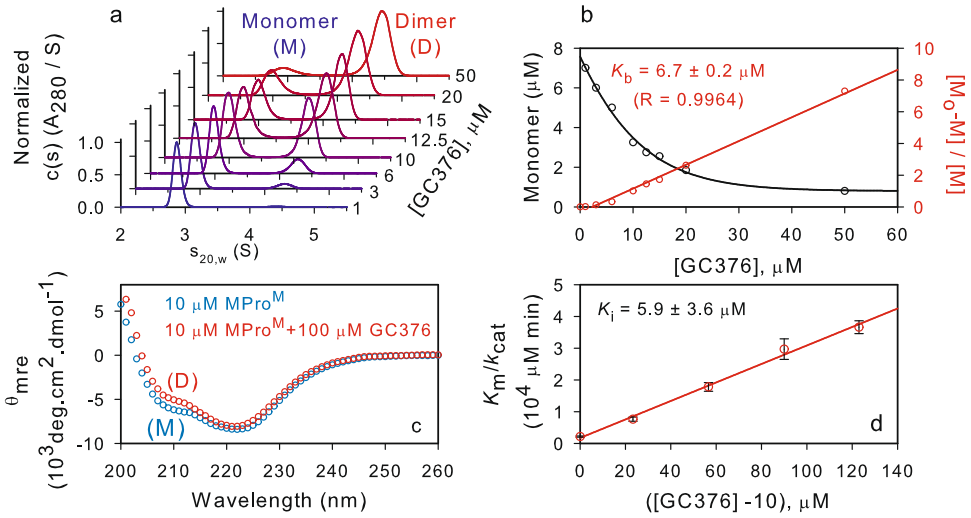

Fig. 6 Modulation of the M-D equilibrium of MPro^{M} and competitive inhibition by GC376. a SV-AUC absorbance c(s) distribution of MPro^{M} (6–7 μM) in the presence of increasing GC376 concentration ranging from 1-50 μM. **b** Plot of the monomer amount *vs* inhibitor concentration (black trace) from data derived from (A). Estimation of the K_b for the binding of GC376 to MPro^{M} (red trace). **c** CD spectra of 10 μM MPro^{M} in the absence (blue) and presence (red) of 100 μM inhibitor GC376. **d** A plot of the K_m/k_{cat} *vs* increasing concentration of GC376 at a final concentration of 10 μM MPro^{M}. Error values indicate a standard deviation of data points recorded 4 times in duplicate (Supplementary Table 2).

K_s, i.e., the intrinsic catalytic activity for MPro^{M}-catalyzed hydrolysis is calculated using Eq. 5.

$$\left(\frac{v_{\max}}{K_m}\right)obs = \frac{k_3\{[M_o]-[M]\}}{2K_s\left\{1+\frac{[I]}{K_i}\right\}} \quad (5)$$

From the observed V_{max}/K_m of 3.3×10^{-3} min^{-1}, the measured K_i of 6.2 μM, [M] value of 5 μM determined by SV-AUC at [M_o] and GC376 concentration at 10 μM each, k_3/K_s for MPro^{M}-catalyzed hydrolysis with one molecule of GC376 bound to one of the active sites is 3.4×10^{-3} μM^{-1} min^{-1} (Table 1). Since k_3/K_s is the intrinsic catalytic activity for MPro^{M}, it is possible to calculate the M-D equilibrium for MPro^{M} from Eq. 3 and the expression:

$$\frac{k_{cat}}{K_m} = k_3/K_1K_s \quad (6)$$

K_1 is calculated to be ~6600 μM. This allowed us to calculate the amount of the protein in the dimer form {[M_o]-[M]} at each protein concentration. A plot of v *vs* {[M_o]-[M]} is linear (R = 0.9992, Supplementary Fig. 3) as expected from Eq. 1. The calculated K_1 for MPro^{M} is 5077 times larger than that of MPro^{WT} corresponding to 5.1 kcal/mol.

Discussion

The results presented above indicate that MPro^{M} adopts a tertiary fold like MPro^{WT} (see refs. [40,41]) as shown by its catalytic activity and ability to bind GC373 and form a transition state analog at the active site, and by the CD spectrum. The observed second-order dependency of the rate of MPro^{M}-catalyzed hydrolysis, and the increase in catalytic activity upon inhibitor binding accompanied by dimer formation are *prima facie* evidence indicating that dimerization is required for mature-like catalytic activity. The kinetic parameters in Table 1 show that the catalytic activity of MPro^{WT} is higher than that of MPro^{M} in the absence and presence of GC373. For MPro^{WT}- and MPro^{M}-catalyzed

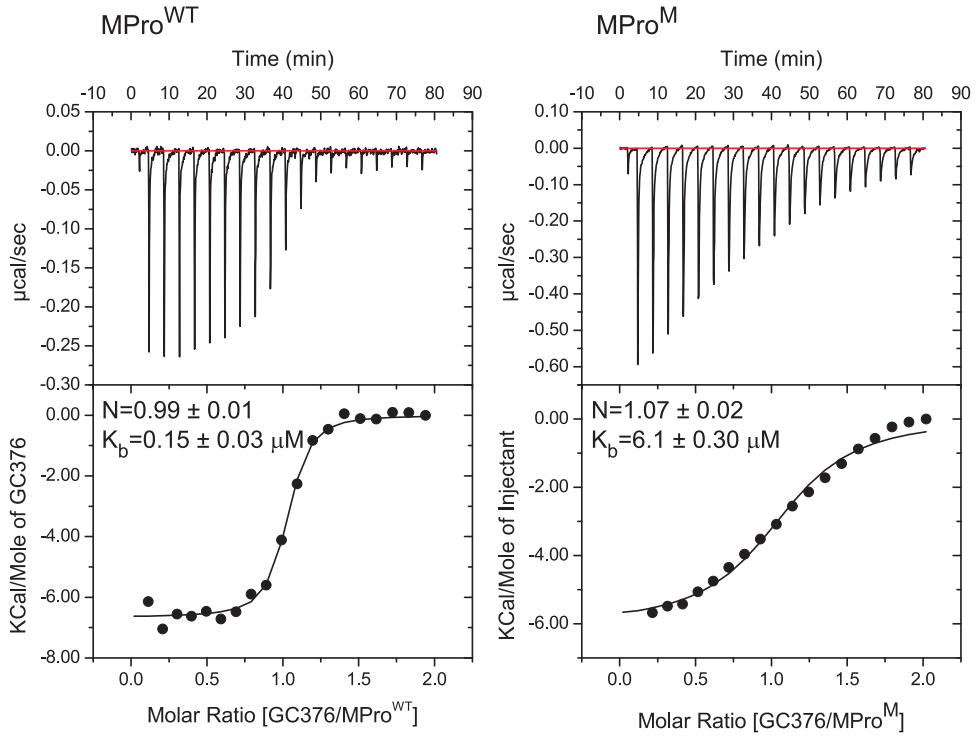

Fig. 7 Binding isotherm of GC376 to MPro^{WT} and MPro^{M}. Titrations were carried out with 30 µM MPro^{WT} and 98 µM MPro^{M} (in the cell) *vs* 300 µM and 1 mM GC376 (in the syringe), respectively, in buffer C at 28 °C. Thermodynamic parameters are listed in Supplementary Table 1.

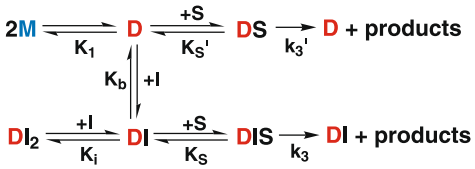

Fig. 8 Mechanism of activation and inhibition of MPro^{M} by GC376. M, D, S, I, DS, DI, DI_2, DIS denote monomer, dimer, substrate, inhibitor, dimer-substrate complex, dimer-inhibitor complex, dimer bound to 2 inhibitors, dimer bound to 1 inhibitor and 1 substrate, respectively.

hydrolysis, k_{cat} and k_{cat}/K_m are k_3 and k_3/K_s, respectively, and thus, they can be directly compared. MPro^{WT} k_{cat} is 56 times larger than that of MPro^{M} (see Table 1) corresponding to 2.4 kcal/mol of free energy which is the additional free energy provided by the MPro^{WT} to stabilize the transition state of the hydrolytic reaction. Similarly, MPro^{WT} k_3/K_s is 120 times larger than that of MPro^{M} corresponding to 2.9 kcal/mol. Since MPro^{M} is properly folded and has catalytic activity, it appears that the catalytic efficiency is also linked to the thermodynamic stability of the dimer form as seen from the comparison of the binding constants of GC373 to MPro^{WT} and MPro^{M} which reflect the difference in the thermodynamic stability of both complexes. The binding constant of GC373 (K_b) to MPro^{M} is 41 times larger than that for MPro^{WT} (see Table 1) indicating that MPro^{WT} complex is more stable than that of MPro^{M} by 2.2 kcal/mol. In contrast, in the absence of GC373, the E290A and R298A mutations destabilize the dimer form of the enzyme by 5.1 kcal/mol. Thus, these results indicate that enzymatic activity is highly dependent on thermodynamic stability of the dimeric form of the enzyme.

The result and conclusions presented above are consistent with published reports[31,47]. Lin et al. observed a clear correlation between k_{cat}/K_m and the rate constant for dimer dissociation (k_{off}), as well as the M-D equilibrium constant, using MPro constructs with N-terminal deletions and single or double substitution mutations which affect dimer formation of SARS-CoV[31]. Increasing both, the k_{off} and M-D dissociation constant (K_d), lead to decreased catalytic activities. The enhanced catalytic activity with increasing inhibitor concentration accompanied by a rise in the abundance of the dimer form (see Figs. 5b and 6A) indicates that binding of the inhibitor to the dimer form of a predominantly monomeric protein changes the equilibrium composition favoring the dimer and thereby establishing the proper geometry of the active site for catalytic activity. We note that dimerization of MPro^{M} does not restore the full catalytic activity (Table 1) probably due to a larger k_{off} for MPro^{M}, compared to that of MPro^{WT}. Silvestrini et al. have identified two groups of compounds that alter the monomer-dimer composition of MPro^{WT} by interacting with the dimer interface[47]. The first group shows the expected strong inhibition of enzymatic activity triggered by the dissociation of the dimer, whereas the second group displays the opposite behavior. The contrary effect of the second group is explained in terms of the presence of a fraction of the protein in the dimer form for which both the inhibitor and the substrate compete. The binding of the substrate to the active sites of the dimer form produces the observed catalytic activity[47]. This observation is consistent with our results that binding the transition state analog to the enzyme favors the dimer form of the enzyme leading to enhanced catalytic activity.

It is well established that the active site of an enzyme is complementary to the transition state of a reaction and not to the substrate, and that the resulting binding energy is utilized for lowering and stabilizing the transition state of a reaction[48]. The mutations E290A and R298A in MPro^{M} destabilize the dimeric form of the enzyme. Consequently, they distort the complementarity between the transition state of the hydrolytic reaction and the active site leading to a loss of 5.1 kcal/mol of stabilization energy to the transition state of the hydrolytic reaction. In

accordance, crystalline structures of the monomeric main protease constructs G11A, S139A and R298A of SARS-CoV show similar major reorganization of the active site including the P1 binding site of the substrate and the catalytic loop comprising the oxyanion hole constituting the weakly-functional catalytic machinery[20,27–29,32]. Thus, the collapsed loop conformation is likely to be an attribute of the monomeric MPro and independent of the position or type of mutation leading to dimer dissociation. It is worth noting that the crystal structure is a single static conformation, i.e., one of many conformations in solution, one of which, albeit at a much lower abundance, forms a dimer to provide the observed catalytic activity of $MPro^{M}$. As shown above, the rate of hydrolysis of the peptide substrate displays a second-order dependency on $MPro^{M}$ concentration indicating the presence of a finite amount of dimeric form of $MPro^{M}$ which is insufficient to be detected by optical methods. The largest difference between $MPro^{M}$ and $MPro^{WT}$ is in their M-D equilibrium constant which reflects the difference in overall catalytic activity. Thus, structure-based design of non- or un-competitive inhibitors that bind to a different form or site of the enzyme and interfere with dimer formation could prove to be an effective alternative strategy to impair catalytic activity[26]. In particular, the dimer interface may be targeted to identify a compound that directly interferes with dimer formation. Since $MPro^{M}$ is predominantly monomeric and displays measurable catalytic activity as well as activated and inhibited by a competitive inhibitor of $MPro^{WT}$ by modulating the monomer-dimer equilibrium, it may be a valuable tool in identifying such inhibitors.

Methods

Expression and purification. The coding sequence of the main protease (MPro) of SARS-CoV-2 (GenBank ID: MN908947.3) bearing the substitution mutations E290A and R298A ($MPro^{M}$, where M denotes monomer, Supplementary Fig. 1a) and the flanking 25 residues of nsp4 at the N-terminus and a 6His-Tag at the C-terminus was synthesized and cloned into pJ414 vector (ATUM, Newark, CA). The plasmid was transformed into BL21-DE3 cells (Agilent) and induced for expression at 0.7-0.8 optical density with 1 mM isopropyl β-D-1-thiogalactopyranoside for 3 h. The processed $MPro^{M}$ was purified from the cell lysate by nickel-affinity chromatography (NAC). The bound fraction was subjected to isocratic fractionation on Superose-12 column (step 1, Cytiva Life Sciences) and HRV-3C protease cleavage (step 2, purchased from Sigma-Aldrich) followed by repeating NAC and step 1 in a final buffer of 25 mM Tris-HCl, pH 7, 150 mM NaCl and 1 mM TCEP (buffer A). The full-length wild type ($MPro^{WT}$) was expressed and purified similar in strategy to that described previously[3] except for substituting the fusion partner GST with maltose binding protein (MBP) followed by a 36 amino acid spacer sequence corresponding to the immunoglobulin binding domain B1 of protein G[49]. Peak fractions were concentrated (5–6 mg/ml) and stored in aliquots at −20 °C and for long term storage at −80 °C. Purity was verified both by SDS-PAGE and electrospray ionization mass spectrometry.

Enzyme kinetics. Activity assays using the FRET substrate Dabsyl-KTSAVLQ/SGFRKM-E(Edans)-NH2[3,35–37], where (/) denotes the scissile peptide bond, were carried out in a total volume of 100 µl in 25 mM Tris-HCl, pH 7, 50 mM NaCl and 1 mM TCEP (buffer B) at 28 °C. Assays were initiated by adding the reaction mixture (95 µl) with or without the inhibitor GC376 to 5 µl of substrate in 100% DMSO kept in the microplate well (Reference 655809, Greiner bio-one). When involving inhibitor, the reaction mixture was incubated for a period of 10 min prior to initiating the reaction with substrate. The excitation and emission wavelengths were set to 336 nm and 490 nm, respectively, and the increase in emission fluorescence intensity was recorded 2–4 times per data point depending on the duration of the data collection as a function of time in a Tecan Infinite M plex microplate reader. After background correction of the average of no enzyme negative controls, concentration of substrate cleaved was determined from a EDANS standard plot and instrument specific inner filter correction values were applied as described[3,50] prior to calculating the kinetic parameters as described below. The substrate was custom synthesized (Biomatik, Ontario, Canada) and GC376 was purchased from Selleckchem, Houston, TX. ΔG was calculated according to the equation $\Delta G = -RTlnK$. The free monomer concentrations in the absence of the inhibitor are calculated from the equations: $K_1 = [M]^2/[D]$ and $[M_o] = [M] + 2[D]$.

Statistics and reproducibility. The FRET substrate is highly sensitive and widely used to assess MPro activity[3,35–37]. Stock solutions of the enzyme, FRET substrate (5 mM in 100% DMSO) and GC376 in (100 mM in 10% DMSO) stored in aliquots at −20 °C were freshly diluted prior to the experiment. The enzyme and inhibitor solutions were kept on ice and the substrate at room temperature while setting up the experiment. The solubility limit of the enzyme, substrate and the inhibitor were verified under the assay buffer conditions. The solubility of the substrate and its products were also monitored during the assay. The reproducibility of enzyme kinetics was tested at least 2-3 times with freshly prepared enzyme and stock solutions of the substrate and inhibitor. Once this was determined to provide consistent reaction rates within an error limit of 5%, the final experiment for the data displayed in the manuscript was carried out in duplicate and 2–4 reads per well for each time point. The mean of the data points was used for fitting. Measurements were processed using SigmaPlot (Systat) by fitting Michaelis-Menten (non-linear) or linear equations to data to calculate the kinetic parameters, standard deviations by the least squares method and multiple correlation coefficients (R). Also, the binding constant of the inhibitor to the enzyme was determined by SV-AUC and ITC using the same stock solutions of enzyme and inhibitor. The binding constants of the inhibitor to the enzyme obtained by the three independent methods were within the reported experimental errors.

Sedimentation velocity analytical ultracentrifugation (SV-AUC). Protein stock solutions maintained in buffer A were diluted to a final concentration ranging from 10-90 µM. Samples containing the inhibitor GC376 were prepared using a 1 mM stock solution of GC376 in buffer B to achieve the desired protein and GC376 ratios in buffer B and a final concentration of 0.1% DMSO.

Sedimentation velocity experiments were conducted at 50,000 rpm and 25 °C on a Beckman Coulter ProteomeLab XL-I analytical ultracentrifuge following standard protocols[51]. Samples were loaded in 2-channel centerpiece cells and scans were collected using both the absorbance (280 nm) and Rayleigh interference (655 nm) optical detection systems. Sedimentation data were time-corrected and analyzed in SEDFIT 16.1C[52] in terms of a continuous c(*s*) distribution of Lamm equation solutions. Solution densities ρ, solution viscosities η, and protein partial specific volumes were calculated in SEDNTERP[53].

Circular Dichroism. CD spectra were recorded in buffer B at 25 °C on a JASCO J-810 spectropolarimeter using Spectra Manager software version 2 (Jasco Analytical Instruments, Easton, MD) and a 0.1 cm pathlength cell. Spectra were processed using the same software.

Size exclusion chromatography with multi-angle light scattering (SEC-MALS). Molecular mass of $MPro^{M}$ and $MPro^{WT}$ was estimated by analytical SEC with in-line MALS (DAWN Heleos-II, Wyatt Technology Inc., Santa Barbara, CA), refractive index (Optilab T-rEX, Wyatt Technology Inc.) and UV (Waters 2487, Waters Corporation, Milford, MA) detectors. Sample was applied onto a pre-equilibrated Superose-12 column (1.0 × 30 cm) and eluted at a flow rate of 0.5 mL/min in buffer A at 25 °C. Molecular mass was calculated using the Astra software provided with the instrument.

Isothermal Titration Calorimetry (ITC). Purified proteins were diluted from a stock solution and dialyzed extensively against buffer C (25 mM Tris-HCl, pH 7.2, 20 mM NaCl and 1 mM TCEP). Concentrations were estimated after dialysis based on their 280 nm absorbance. A stock solution of GC376 in 10% DMSO was diluted in buffer C to the desired concentration. Titrations were performed at 28 °C on iTC200 microcalorimeter (Malvern Instruments Inc., Westborough, MA). A control titration of buffer with inhibitor showed negligible response. Data were processed using the Origin software provided with the instrument. For competitive inhibitors that bind at only one site, the binding constant $\{K_{binding}\ (K_b) = 1/K_a\}$ is equivalent to the inhibition constant measured by enzyme kinetics (K_i).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Source data files are provided in Supplementary Data 1 and 2.

Received: 14 October 2021; Accepted: 28 January 2022;
Published online: 01 March 2022

References

1. V'kovski, P., Kratzel, A., Steiner, S., Stalder, H. & Thiel, V. Coronavirus biology and replication: implications for SARS-CoV-2. *Nat. Rev. Microbiol* **19**, 155–170 (2020).
2. Wu, F. et al. A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265–269 (2020).
3. Zhang, L. et al. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved alpha-ketoamide inhibitors. *Science* **368**, 409–412 (2020).

4. Mariano, G., Farthing, R. J., Lale-Farjat, S. L. M. & Bergeron, J. R. C. Structural Characterization of SARS-CoV-2: Where We Are, and Where We Need to Be. *Front Mol. Biosci.* **7**, 605236 (2020).
5. Wang, C. et al. The establishment of reference sequence for SARS-CoV-2 and variation analysis. *J. Med Virol.* **92**, 667–674 (2020).
6. Groneberg, D. A., Hilgenfeld, R. & Zabel, P. Molecular mechanisms of severe acute respiratory syndrome (SARS). *Respir. Res* **6**, 8 (2005).
7. Chen, S., Jonas, F., Shen, C. & Hilgenfeld, R. Liberation of SARS-CoV main protease from the viral polyprotein: N-terminal autocleavage does not depend on the mature dimerization mode. *Protein Cell* **1**, 59–74 (2010).
8. Heinz, F. X. & Stiasny, K. Distinguishing features of current COVID-19 vaccines: knowns and unknowns of antigen presentation and modes of action. *NPJ Vaccines* **6**, 104 (2021).
9. Ghahremanpour M. M. et al. Identification of 14 Known Drugs as Inhibitors of the Main Protease of SARS-CoV-2. *bioRxiv*, (2020).
10. Baker, J. D., Uhrich, R. L., Kraemer, G. C., Love, J. E. & Kraemer, B. C. A drug repurposing screen identifies hepatitis C antivirals as inhibitors of the SARS-CoV2 main protease. *PLoS ONE* **16**, 1–13 (2021).
11. Owen D. R. et al. An oral SARS-CoV-2 M(pro) inhibitor clinical candidate for the treatment of COVID-19. *Science*, eabl4784 (2021).
12. Hsu, M. F. et al. Mechanism of the maturation process of SARS-CoV 3CL protease. *J. Biol. Chem.* **280**, 31257–31266 (2005).
13. Muramatsu, T. et al. Autoprocessing mechanism of severe acute respiratory syndrome coronavirus 3C-like protease (SARS-CoV 3CLpro) from its polyproteins. *FEBS J.* **280**, 2002–2013 (2013).
14. Li, C. et al. Maturation mechanism of severe acute respiratory syndrome (SARS) coronavirus 3C-like proteinase. *J. Biol. Chem.* **285**, 28134–28140 (2010).
15. Kneller, D. W. et al. Unusual zwitterionic catalytic site of SARS-CoV-2 main protease revealed by neutron crystallography. *J. Biol. Chem.* **295**, 17365–17373 (2020).
16. El-Baba, T. J. et al. Allosteric Inhibition of the SARS-CoV-2 Main Protease: Insights from Mass Spectrometry Based Assays. *Angew. Chem. Int Ed. Engl.* **59**, 23544–23548 (2020).
17. Cheng, S. C., Chang, G. G. & Chou, C. Y. Mutation of Glu-166 blocks the substrate-induced dimerization of SARS coronavirus main protease. *Biophys. J.* **98**, 1327–1336 (2010).
18. Graziano, V., McGrath, W. J., Yang, L. & Mangel, W. F. SARS CoV main proteinase: The monomer-dimer equilibrium dissociation constant. *Biochemistry* **45**, 14632–14641 (2006).
19. Wei, P. et al. The N-terminal octapeptide acts as a dimerization inhibitor of SARS coronavirus 3C-like proteinase. *Biochem Biophys. Res Commun.* **339**, 865–872 (2006).
20. Xia, B. & Kang, X. Activation and maturation of SARS-CoV main protease. *Protein Cell* **2**, 282–290 (2011).
21. Zhong, N. et al. Without its N-finger, the main protease of severe acute respiratory syndrome coronavirus can form a novel dimer through its C-terminal domain. *J. Virol.* **82**, 4227–4234 (2008).
22. Shi, J., Wei, Z. & Song, J. Dissection study on the severe acute respiratory syndrome 3C-like protease reveals the critical role of the extra domain in dimerization of the enzyme: defining the extra domain as a new target for design of highly specific protease inhibitors. *J. Biol. Chem.* **279**, 24765–24773 (2004).
23. Shi, J. & Song, J. The catalysis of the SARS 3C-like protease is under extensive regulation by its extra domain. *FEBS J.* **273**, 1035–1045 (2006).
24. Anand, K. et al. Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain. *EMBO J.* **21**, 3213–3224 (2002).
25. Chang, H. P., Chou, C. Y. & Chang, G. G. Reversible unfolding of the severe acute respiratory syndrome coronavirus main protease in guanidinium chloride. *Biophys. J.* **92**, 1374–1383 (2007).
26. Goyal, B. & Goyal, D. Targeting the Dimerization of the Main Protease of Coronaviruses: A Potential Broad-Spectrum Therapeutic Strategy. *ACS Comb. Sci.* **22**, 297–305 (2020).
27. Chou, C. Y. et al. Quaternary structure of the severe acute respiratory syndrome (SARS) coronavirus main protease. *Biochemistry* **43**, 14958–14970 (2004).
28. 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.* **82**, 4620–4629 (2008).
29. Chen, S. et al. 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 structure with molecular dynamics simulations. *J. Biol. Chem.* **283**, 554–564 (2008).
30. Hsu, W. C. et al. Critical assessment of important regions in the subunit association and catalytic action of the severe acute respiratory syndrome coronavirus main protease. *J. Biol. Chem.* **280**, 22741–22748 (2005).
31. Lin, P. Y., Chou, C. Y., Chang, H. C., Hsu, W. C. & Chang, G. G. Correlation between dissociation and catalysis of SARS-CoV main protease. *Arch. Biochem Biophys.* **472**, 34–42 (2008).
32. Hu, T. et al. Two adjacent mutations on the dimer interface of SARS coronavirus 3C-like protease cause different conformational changes in crystal structure. *Virology* **388**, 324–334 (2009).
33. Barrila, J., Bacha, U. & Freire, E. Long-range cooperative interactions modulate dimerization in SARS 3CLpro. *Biochemistry* **45**, 14908–14916 (2006).
34. Xue, X. et al. Production of authentic SARS-CoV M(pro) with enhanced activity: application as a novel tag-cleavage endopeptidase for protein overproduction. *J. Mol. Biol.* **366**, 965–975 (2007).
35. Ma, C. et al. Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease. *Cell Res.* **30**, 678–692 (2020).
36. Kneller, D. W. et al. Malleability of the SARS-CoV-2 3CL M(pro) Active-Site Cavity Facilitates Binding of Clinical Antivirals. *Structure* **28**, 1313–1320 (2020). e1313.
37. Noske, G. D. et al. A Crystallographic Snapshot of SARS-CoV-2 Main Protease Maturation Process. *J. Mol. Biol.* **433**, 167118 (2021).
38. Sayer, J. M., Agniswamy, J., Weber, I. T. & Louis, J. M. Autocatalytic maturation, physical/chemical properties, and crystal structure of group N HIV-1 protease: relevance to drug resistance. *Protein Sci.* **19**, 2055–2072 (2010).
39. Todd, M. J., Semo, N. & Freire, E. The structural stability of the HIV-1 protease. *J. Mol. Biol.* **283**, 475–488 (1998).
40. Wang, Y. C. et al. Structural basis of SARS-CoV-2 main protease inhibition by a broad-spectrum anti-coronaviral drug. *Am. J. Cancer Res.* **10**, 2535–2545 (2020).
41. Vuong, W. et al. Feline coronavirus drug inhibits the main protease of SARS-CoV-2 and blocks virus replication. *Nat. Commun.* **11**, 4282 (2020).
42. Kim Y. et al. Reversal of the progression of fatal coronavirus infection in cats by a broad-spectrum coronavirus protease inhibitor. *PLoS Pathog* **12**, (2016).
43. Kim, Y. et al. Broad-spectrum antivirals against 3C or 3C-like proteases of picornaviruses, noroviruses, and coronaviruses. *J. Virol.* **86**, 11754–11762 (2012).
44. Pedersen, N. C. et al. Efficacy of a 3C-like protease inhibitor in treating various forms of acquired feline infectious peritonitis. *J. Feline Med. Surg.* **20**, 378–392 (2018).
45. Bai Y. et al. Structural basis for the inhibition of the SARS-CoV-2 main protease by the anti-HCV drug narlaprevir. *Signal Transduct Tar.* **6**, (2021).
46. Chen, H. et al. Only one protomer is active in the dimer of SARS 3C-like proteinase. *J. Biol. Chem.* **281**, 13894–13898 (2006).
47. Silvestrini, L. et al. The dimer-monomer equilibrium of SARS-CoV-2 main protease is affected by small molecule inhibitors. *Sci. Rep.* **11**, 9283 (2021).
48. Fersht A. *Enzyme structure and Mechanism*. W.H. Freeman and company (1985).
49. Kneller D. W. et al. Structural, Electronic, and Electrostatic Determinants for Inhibitor Binding to Subsites S1 and S2 in SARS-CoV-2 Main Protease. *J. Med. Chem.* (2021).
50. Liu, Y. Y. et al. Use of a fluorescence plate reader for measuring kinetic parameters with inner filter effect correction. *Anal. Biochem* **267**, 331–335 (1999).
51. Zhao H., Brautigam C. A., Ghirlando R., Schuck P. Overview of current methods in sedimentation velocity and sedimentation equilibrium analytical ultracentrifugation. *Curr Protoc Protein Sci* **Chapter 20**, Unit20 12 (2013).
52. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.* **78**, 1606–1619 (2000).
53. Cole, J. L., Lary, J. W., T, P. M. & Laue, T. M. Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium. *Methods Cell Biol.* **84**, 143–179 (2008).
54. Kneller, D. W., Zhang, Q., Coates, L., Louis, J. M. & Kovalevsky, A. Michaelis-like complex of SARS-CoV-2 main protease visualized by room-temperature X-ray crystallography. *IUCrJ* **8**, 973–979 (2021).

Acknowledgements

We thank John Lloyd and the NIDDK Core Facility for mass spectrometry. We are grateful to Christopher Premanandan, Robert Best and Andrey Kovalevsky for many thoughtful discussions on SARS-CoV-2 main protease and related ongoing projects. This research was supported by the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health.

Author contributions

N.T.N. and J.M.L. designed the study and carried out enzyme kinetics; A.A. and J.M.L.: constructed, expressed, and purified proteins; R.G. performed and analyzed SV-AUC

data; J.M.L. and S.C.C. carried out ITC, CD and SEC-MALS analyses; N.T.N., R.G., S.C.C. and J.M.L.: analyzed data. N.T.N. and J.M.L.: wrote the paper with help from all co-authors.

Funding
Open Access funding provided by the National Institutes of Health (NIH).

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-022-03084-7.

Correspondence and requests for materials should be addressed to John M. Louis.

Peer review information *Communications Biology* thanks Stefan Siemann and the other, anonymous, reviewers for their contribution to the peer review of this work. Primary Handling Editor: Gene Chong.