

## CORONAVIRUS

## Molecular characterization of ebselen binding activity to SARS-CoV-2 main protease

Cintia A. Menéndez<sup>1\*</sup>, Fabian Byléhn<sup>1\*</sup>, Gustavo R. Perez-Lemus<sup>1\*</sup>,  
Walter Alvarado<sup>1\*</sup>, Juan J. de Pablo<sup>1,2†</sup>

There is an urgent need to repurpose drugs against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Recent computational-experimental screenings have identified several existing drugs that could serve as effective inhibitors of the virus' main protease, M<sup>Pro</sup>, which is involved in gene expression and replication. Among these, ebselen (2-phenyl-1,2-benzoselenazol-3-one) appears to be particularly promising. Here, we examine, at a molecular level, the potential of ebselen to decrease M<sup>Pro</sup> activity. We find that it exhibits a distinct affinity for the catalytic region. Our results reveal a higher-affinity, previously unknown binding site localized between the II and III domains of the protein. A detailed strain analysis indicates that, on such a site, ebselen exerts a pronounced allosteric effect that regulates catalytic site access through surface-loop interactions, thereby inducing a reconfiguration of water hotspots. Together, these findings highlight the promise of ebselen as a repurposed drug against SARS-CoV-2.

## INTRODUCTION

A coronavirus of zoonotic origin, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is the etiological agent responsible for the 2019–2020 viral pneumonia COVID-19 outbreak that commenced in Wuhan City, Hubei Province, China (1–4). Currently, targeted therapeutics are limited and, at this point in time, effective treatment options remain tentative and highly restricted.

Conventional drug design and drug development strategies require years of investigations and substantial investments. The repurposing of approved pharmaceutical drugs (and drug candidates already in clinical trials) therefore provides an essential, alternative approach to rapidly identify drugs with clinical potential to help manage emerging infectious diseases for which specific treatments and vaccines are not available. A recent high-throughput screening study considered more than 10,000 compounds against the main protease M<sup>Pro</sup> of SARS-CoV-2 (5). The main protease is one of the key enzymes in the viral life cycle, as it plays a central role in mediating viral replication and transcription. Its structure is shown in Fig. 1. Along with other, nonstructural proteins (papain-like protease, helicase, and RNA-dependent RNA polymerase) and the spike glycoprotein structural protein, M<sup>Pro</sup> is essential for interactions between the virus and host cell receptors during viral entry (6), making it an attractive drug target against SARS-CoV-2 (7, 8).

In a recent study by Jin *et al.* (5), a wide range of approved drugs, drug candidates, and natural products were screened by an enzymatic inhibition assay to identify a possible M<sup>Pro</sup> inhibitor. Ebselen, an organoselenium compound whose structure is shown in Fig. 1, emerged from that study as a promising drug lead to target this crucial enzyme. It was found to exhibit strong antiviral activity in cell-based assays (5). In previous, pre-COVID-19 work, ebselen was found to have anti-inflammatory, antioxidative, and cytoprotective properties. It has been investigated in the context of multiple diseases, such as bipolar disorders (9) and hearing loss (10, 11). In addition,

a recent report indicated that ebselen also showed potent bactericidal activity against multidrug-resistant (MDR) clinical isolates of *Staphylococcus aureus* (12). Likewise, the use of silver and ebselen in synergistic formulations has been shown to be effective against five, clinically difficult-to-treat MDR Gram-negative bacteria (13). Ebselen has low cytotoxicity (median lethal dose in rats >4600 mg/kg, per os) (14), and its safety in humans has been evaluated in several clinical trials (10, 11, 15). Together, these reports—and the underlying data—underscore the clinical potential of ebselen for CoV-2 treatment.

In an effort to understand the molecular mechanisms through which ebselen interacts with M<sup>Pro</sup>, in this work, we present results of atomistic molecular simulations that provide useful, previously unknown insights into the M<sup>Pro</sup>-ebselen complex, which might provide novel avenues to rationally enhance ebselen's activity. First, we present an analysis of the most probable interaction sites between M<sup>Pro</sup>-ebselen, as well as absolute binding free energy calculation for a complementary quantitative evaluation. Second, we evaluate the role of different binding sites on molecular stiffness and molecular strain. Last, we examine how ebselen binding modifies the structure and transport of water in M<sup>Pro</sup>'s catalytic site. We find that ebselen binds to two sites, as opposed to only one at the catalytic site, and we also identify an allosteric mechanism that influences the catalytic site when ebselen is bound at the distant site. Our findings are summarized in a concluding section, where we also provide several suggestions for future experimental work.

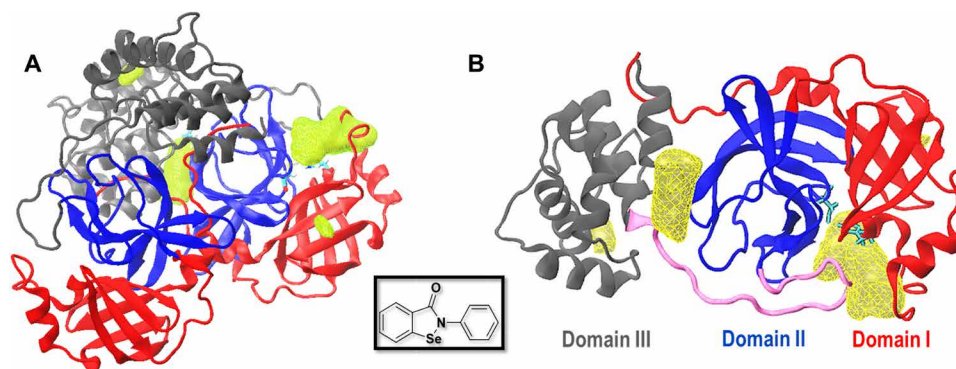
## RESULTS

Molecular dynamics (MD) simulations of M<sup>Pro</sup> and ebselen were carried out using the AMBER18 simulation package (see Materials and Methods for details). A total of 3  $\mu$ s of atomistic MD trajectories of M<sup>Pro</sup> using ebselen as a molecular probe were organized into 15 replicas of 200 ns each. These trajectories were analyzed to construct probability density maps for the preferred locations of ebselen around the protein. The results are shown in Fig. 1. Two distinct, highly probable binding sites emerged from this analysis. The first is located within the catalytic site, and the second is in a region that is essential for M<sup>Pro</sup>'s dimerization (16), between the II and III domains.

<sup>1</sup>Pritzker School of Molecular Engineering, University of Chicago, 5640, S. Ellis Avenue, Chicago, IL 60637, USA. <sup>2</sup>Argonne National Laboratory, Lemont, IL 60439, USA.

\*These authors contributed equally to this work.

†Corresponding author. Email: depablo@uchicago.edu



**Fig. 1. Structure of  $M^{\text{pro}}$  and density maps of ebselen binding.** (A)  $M^{\text{pro}}$  dimer and (B) protomer A (top view). In both panels,  $M^{\text{pro}}$  domains I, II, and III are shown in red, blue, and gray, respectively. Cys<sup>145</sup> and His<sup>41</sup> (the catalytic dyad) are shown in cyan. Yellow surfaces show the most probable interaction sites between  $M^{\text{pro}}$  and ebselen (highest probability density). The loop (residues 185 to 201) connecting domains I and II with domain III is shown in pink (B). The inset shows the chemical structure of the ebselen molecule.

Similar observations were reported in a previous study (17), where simple organic solvents or compounds such as acetonitrile, benzene, dimethyl sulfoxide, methanol, phenol, and urea were used as molecular probes. In addition, the authors considered the potential mutability of residues belonging to the  $M^{\text{pro}}$  catalytic site and explained that the development of drug resistance associated with the natural evolution of  $M^{\text{pro}}$  could wipe out efforts that target this protein for COVID-19 treatment. Instead, they emphasized the promise of alternative strategies aimed at targeting the region between the II and III domains, which is implicated in dimer formation. The results shown in Fig. 1 are therefore encouraging in that ebselen appears to target both the catalytic site and the dimerization domain.

Figure 2 provides a closer look at ebselen's binding motifs in both regions (the catalytic site and the intersection between domains II and III). In the catalytic site (Fig. 2A), hydrogen bonding interactions are formed between the carbonyl oxygen of ebselen and the Asn<sup>142</sup> and Gln<sup>189</sup> side chains. One can also appreciate the hydrophobic contacts between ebselen and Met<sup>165</sup>, Pro<sup>168</sup>, Met<sup>49</sup>, His<sup>164</sup>, and both residues belonging to the catalytic dyad, His<sup>41</sup> and Cys<sup>145</sup>.

Figure 2B shows a representative configuration of ebselen at the intersection between domains II and III. There are distinct hydrophobic contacts with Phe<sup>294</sup>, Pro<sup>108</sup>, Ile<sup>200</sup>, Val<sup>202</sup>, His<sup>246</sup>, Thr<sup>292</sup>, Ile<sup>249</sup>, Pro<sup>132</sup>, and Ile<sup>249</sup> (some residues are not shown for clarity). Highly dynamic hydrogen bonds were observed between ebselen and the Gln<sup>107</sup>, Gln<sup>110</sup>, and His<sup>246</sup> side chains.

The results shown in Fig. 1 were generated on the basis of direct MD simulations. To arrive at a quantitative estimate of the binding affinity of each site, we used thermodynamic integration (TI) to determine the absolute binding free energy for the catalytic site and the domains II and III site. More specifically, we used particle mesh Ewald molecular dynamics (PMEMD) as implemented in AMBER18 (18) with 11 windows per integration and 10 ns per window. In addition, multiple runs starting from the most probable binding cluster identified in the previous MD simulations were considered. In this way, three independent replicas for each site were used to calculate averages. The results are shown in Table 1.

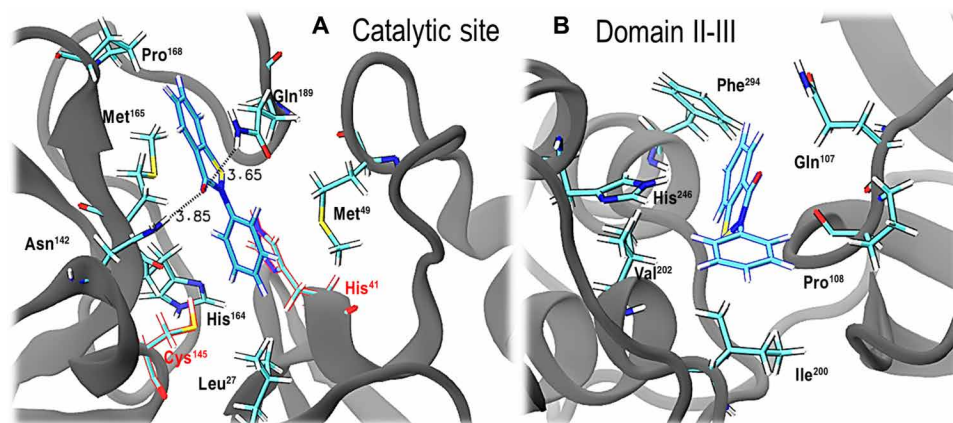
Consistent with the probability density maps shown in Fig. 1, the absolute binding free energies corresponding to both sites are negative, serving to underscore the thermodynamic stability of the ebselen- $M^{\text{pro}}$  complex at both sites. Here, we note the binding affinity of ebselen at the second binding site, located between the II

and III domains, is, in fact, greater than that at the catalytic site. A previous report suggested that this small molecule could also inhibit  $M^{\text{pro}}$  through noncovalent binding, particularly because it has been found to exhibit a stronger inhibition effect than other compounds that were also able to quantitatively modify Cys<sup>145</sup> in the catalytic dyad (5). The results presented here could explain ebselen's high enzymatic inhibition, even when this compound could only partially modify Cys<sup>145</sup>.

Having identified a distant binding site for ebselen between domains II and III, we examined the role of this binding site, if any, on the catalytic site of the protein. To do so, we determined the local strain induced by the binding throughout the protein. To examine potential synergies of drug binding to both sites, we also considered a third scenario in which both sites (catalytic and distant) are simultaneously occupied by ebselen. Note that local strain provides a measure of local deformation that filters out nontrivial, functional conformational changes from nonfunctional ones, and it is therefore ideally suited to highlight how a perturbation (i.e., ebselen binding) at one site induces conformational changes at other, potentially distant, sites. It is also important to highlight that different measures of strain have recently been used in studies of allostery in proteins (19–21). Here, we used a method originally introduced by our group in the context of local strain in polymeric glasses (22).

Figure 3 shows the shear strain throughout the protein upon binding of ebselen at different sites, alongside the  $\beta$  factor (estimated from the root mean squared fluctuation), which measures thermal fluctuations for each residue. The shear strains are measured relative to the average conformation from all frames of the apo protein trajectory; these strains, therefore, correspond to the deformation upon binding relative to the apo protein. The shear strain for the apo structure is then showing an average internal strain in the protein.

From the  $\beta$  factor analysis, it is evident that when ebselen is bound to the domains II and III interface (Fig. 3A, blue lines), the dynamics of the 44–52 loop, which flanks the catalytic site, is notably altered. On the other hand, when ebselen is bound at both sites simultaneously, there is a clear global reduction of the thermal fluctuations among the receptor except for one single region located around residues 137 to 140, where enhanced flexibility is apparent. Apart from these residues, ebselen bound at both sites shows the lowest  $\beta$  factor values; however, at this specific point, this system exhibits as high flexibility as  $M^{\text{pro}}$ -apo protein (Fig. 3A, green line).



**Fig. 2. Binding modes for ebselen- $M^{\text{PRO}}$  complexes.** (A) At the catalytic site and (B) at the intersection between domains II and III. In both panels, ebselen, as well as main residues displaying contacts, is shown in sticks.

**Table 1. Absolute binding free energy for  $M^{\text{PRO}}$ -ebselen complexes.** The average and uncertainty energy values were obtained from three independent replicas for each site. Details of the calculation are described in Materials and Methods.

Binding site	Absolute binding free energy (kcal/mol)
Catalytic site	$-5.55 \pm 2.28$
Domains II and III site	$-8.87 \pm 1.59$

The strain analysis helps disentangle the effects of functional and nonfunctional fluctuations and provides a more detailed view of the effects of ebselen binding. As seen in Fig. 3A for the  $\beta$  factor analysis, when ebselen is bound to the domain interface, it produces a large strain at the 44–52 loop (Fig. 3B, blue line). Likewise, the 185–201 loop that also flanks the catalytic site, as well as the region comprising residues 137 to 140, exhibits a high strain, which is not immediately apparent in the  $\beta$  factor. When both molecules are bound simultaneously to  $M^{\text{PRO}}$  (Fig. 3B, black line), a high strain signal is also exhibited around residues 137 to 140, as expected on the basis of the aforementioned  $\beta$  factor results. Last, when ebselen is located in the catalytic site, a much lower strain is showed around all three regions (Fig. 3B, red line). Figure 3C shows that the strain is primarily localized at the two loops flanking the catalytic site (44–52 and 185–201 loops), as well as a loop in catalytic site (residues 22 to 25). A highly strained region also appears around residues 137 to 140. To understand why the 137 to 140 residues show such a large strain when ebselen is bound to both sites, we turn our attention to the molecular images shown in Fig. 4.

In the close-up of the residues shown in Fig. 4A, a specific backbone hydrogen bond forms between Lys<sup>137</sup> and Phe<sup>140</sup> because of the binding of ebselen between domains II and III; this conformational change induced by the presence of ebselen causes the high strain shown in Fig. 3B. This conformational change takes place in the middle region between the catalytic site (Cys<sup>145</sup>) and the binding cleft between domains II and III (Pro<sup>132</sup>), which points to the relevance of these residues.

From these results, it is evident that when ebselen binds between the II and III domains, it exerts a pronounced allosteric effect that

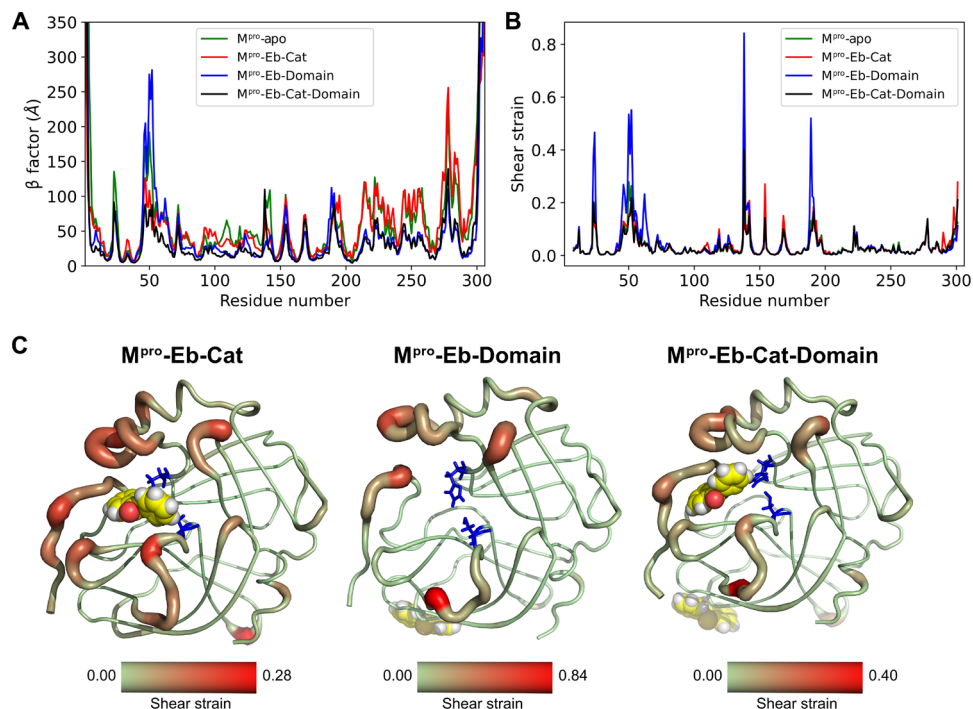
affects the loops that regulate access to the catalytic site (44–52 and 185–201 loops). In addition, it affects the residues 137 to 140, where a specific backbone hydrogen bond forms between Lys<sup>137</sup> and Phe<sup>140</sup>. The exact role of this conformational change will be the subject of future studies, but the results presented here show that it serves as a relay between domain III and the catalytic site.

Given that conformational changes in the catalytic site are observed when ebselen binds to  $M^{\text{PRO}}$  far away from this specific region, we turn our attention to the hydration characteristics of the catalytic site in the  $M^{\text{PRO}}$ -ebselen complex (when bound to the domains II and III interface) and in the  $M^{\text{PRO}}$ -apo structure [Protein Data Bank (PDB) code: 6M03] (23). AQUA-DUCT 1.0.5 (24) was used to analyze the water structure and water flux in the  $M^{\text{PRO}}$  protein, with a time window of 50 ns and sampling every 1 ps. The results are shown in Fig. 5.

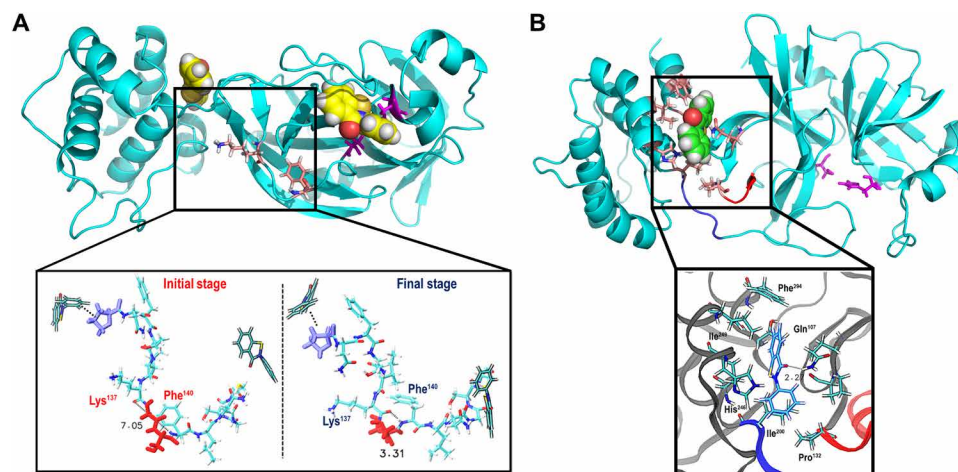
Figure 5 (top) shows that ebselen binding between domains II and III leads to fewer water inlets compared with the  $M^{\text{PRO}}$ -apo state; this is indicative of a water flux reduction in the catalytic site upon ebselen binding, although it binds far from the catalytic site. From the lower panel in Fig. 5, it is evident that a volume reduction of the catalytic site occurs when ebselen is located between domains II and III compared with the apo state; the maximum available volume (MAV) for water in the catalytic region is around 50% smaller in this case. Note that in the apo protein, there is a catalytic water molecule that forms a catalytic triad together with Cys<sup>145</sup> and His<sup>41</sup> (17, 25), and this catalytic water (red in Fig. 5A, bottom) is preserved and remains close to His<sup>41</sup> in the simulations. In contrast, it is clear from Fig. 5B (bottom) that the presence of ebselen induces a displacement and reconfiguration of water hotspots, including the catalytic water (red). These effects could prevent the normal enzymatic function of  $M^{\text{PRO}}$ , as this catalytic water displacement might damage the catalytic triad required for protein activity (26). Similarly, the observed pocket volume reduction might affect the accessibility of the polyprotein that  $M^{\text{PRO}}$  cleaves, thereby reducing enzymatic function.

## DISCUSSION

As mentioned before,  $M^{\text{PRO}}$  is an attractive drug target against the COVID-19 virus due to its central role in the viral life cycle. A previous structural and evolutionary investigation suggested that the



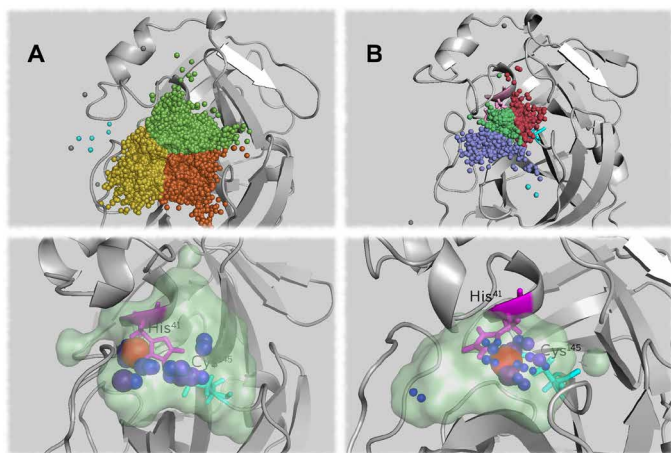
**Fig. 3. Strain analysis of ebnselen- $M^{\text{Pro}}$  complexes.** (A)  $\beta$  factor, estimated from the root mean squared fluctuation. (B) Shear strains. In both upper panels,  $M^{\text{Pro}}$ -apo results are shown in green ( $M^{\text{Pro}}$ -apo); for ebnselen, when bound to the catalytic site ( $M^{\text{Pro}}$ -Eb-Cat), results are in red. For the domain II–III interface ( $M^{\text{Pro}}$ -Eb-Domain), results are shown in blue. For two ebnselen molecules bound at both sites simultaneously ( $M^{\text{Pro}}$ -Eb-Cat-Domain), results are shown in black. For the shear strain calculation, only the  $C\alpha$  atoms are included. (C) Shear strains mapped onto the different protein complexes from (B). The redder the region and the larger the radius of the structure, the higher the shear strain. For clarity, only the catalytic site and domains I and II are shown, the drug is shown as spheres, and the catalytic dyad is shown as blue sticks.



**Fig. 4. Close-up of the highly strained regions identified from strain analysis.** (A) The  $M^{\text{Pro}}$  structure with ebnselen molecules bound to both catalytic site and domains II and III interface simultaneously. In the upper panel, Lys<sup>137</sup> and Phe<sup>140</sup> are shown as sticks. The lower panel shows the close-up of this region and shows the backbone H-bond between Lys<sup>137</sup> and Phe<sup>140</sup> at the beginning and end of the simulation. (B) The  $M^{\text{Pro}}$  structure with ebnselen bound to the domains II and III interface. The close-up shows that a hydrogen bonding interaction is formed with Gln<sup>107</sup> side chain, and hydrophobic contacts are formed with His<sup>246</sup>, Val<sup>202</sup>, Ile<sup>249</sup>, Phe<sup>294</sup>, Pro<sup>132</sup>, and Ile<sup>200</sup>. This last residue is located at the end of the 185–201 loop, which connects domain III with domains I and II (catalytic site). In both (A) and (B) upper panels, ebnselen is shown as spheres and the catalytic dyad as magenta sticks.

SARS-CoV-2  $M^{\text{Pro}}$  is not a suitable target for de novo development of inhibitors or the repurposing of drugs against the previous SARS coronavirus (17). However, only the flexibility and plasticity of the active sites in  $M^{\text{Pro}}$  for COVID-19 and the highly similar previous SARS-CoV  $M^{\text{Pro}}$  were compared. Nevertheless, major differences in

both shape and size of the catalytic site were observed, indicating that repurposing SARS drugs for COVID-19 may not be effective. On the basis of their evolutionary analysis, these authors also pointed out that the virus's mutability will pose further challenges to treatments against the COVID-19  $M^{\text{Pro}}$  protein. An alternative to this



**Fig. 5. Water analysis of  $M^{Pto}$  complexes.** Top: Water inlet clusters for (A) apo protein and (B) ebselen in domains II and III. The biggest cluster in both cases is divided into three parts using *K*-means (24) and colored for clarity. Bottom: MAV (as light green surface) and hotspots (blue-red color and sphere radius relative to occupation values) for (A) apo and (B) ebselen in domains II and III. The catalytic dyad is represented with sticks (magenta, His<sup>41</sup>; cyan, Cys<sup>145</sup>).

discouraging scenario, however, would be to target the region between the II and III domains, which is implicated in dimer formation.

Here, we find that there are two highly probable interaction sites between SARS-CoV-2  $M^{Pto}$  and ebselen. One is located within the catalytic cavity and, more importantly, in the region between the II and III domains, which is essential for  $M^{Pto}$  dimerization (16). Detailed calculations of the free energy reveal a higher binding affinity of ebselen to domains II and III than to the catalytic site.

The strain analysis reveals that ebselen bound between the II and III domains exert a pronounced allosteric effect that affects the loops regulating access to the catalytic site. In addition, it also affects residues 137 to 140, where a specific backbone hydrogen bond forms between Lys<sup>137</sup> and Phe<sup>140</sup>.

The catalytic site water analysis indicates that the proposed allosteric inhibition by ebselen could occur through a volume reduction of the catalytic pocket and a reconfiguration of water hotspots in that region. Given the catalytic role of water in this enzyme's activity, these effects could act to prevent the regular enzymatic function of the SARS-CoV-2  $M^{Pto}$  protein.

Together, the results from our comprehensive study of ebselen activity and the underlying data serve to underscore the potential of ebselen for COVID-19 treatment. The discovery of a distant binding site is also encouraging in that it offers potential for development of novel  $M^{Pto}$  inhibitors based on the ebselen scaffold; indeed, distant sites could be effective targets for other drugs that might have been deemed futile for catalytic sites. Note that current, massive virtual screening campaigns are focused primarily on targeting the SARS-CoV-2  $M^{Pto}$  catalytic site; our study, however, highlights the potential importance of a different, distant site (27). Further experimental structural characterization will be necessary to validate the predictions presented in this work, particularly those pertaining to the distant binding site. In the next stages of this project, we will seek to investigate the experimental crystal structure of the SARS-CoV-2  $M^{Pto}$ -ebselen complex, to validate our theoretical predictions and help elucidate the inhibitory mechanism. These efforts could also be combined with rational mutagenesis of interacting

residues to alanine and binding assays such as surface plasmon resonance, with which to quantify binding and provide a basis for comparisons to theoretical predictions.

A previous evolutionary study (17) showed that mutating a few residues belonging to the catalytic site is energetically unfavorable. Residues such as P39, R40, P52, G143, G146, or L167 could therefore be considered as key anchoring sites for future  $M^{Pto}$  inhibitor design. A recent study (5) identified ebselen as a promising repurposed drug against SARS-CoV-2. Our results build on that work by providing a detailed molecular picture of the interactions between ebselen and  $M^{Pto}$ . The insights provided here could help facilitate the rational molecular design of new analogs, based on the ebselen scaffold and its binding to a distant site, for future coronaviruses. This could be substantial, particularly given the fact that the SARS-CoV-2  $M^{Pto}$  protein shares a high sequence similarity to the SARS-CoV  $M^{Pto}$  (28), and is therefore a likely target for future coronaviruses. This alternative approach, uncovered through the extensive molecular simulations presented here, highlights the need to develop reliable, high-throughput methods to screen drug-protein interactions at the molecular level and that incorporate the role of explicit water molecules. We conclude by pointing out that our focus in this work has been on noncovalent complexes between ebselen- $M^{Pto}$ . In a subsequent stage, we also plan to investigate covalent complexes involving Cys<sup>145</sup>.

## MATERIALS AND METHODS

A total of more than 6  $\mu$ s of classical MD simulations of SARS-CoV-2  $M^{Pto}$ -apo state and  $M^{Pto}$ -ebselen complex were run using the AMBER18 (29) simulation package (3  $\mu$ s, ebselen as a molecular probe; 2.4  $\mu$ s, shear strain analysis; 100 ns, water structure and flux analysis; 990 ns, free energy analysis). The receptor initial configuration for the  $M^{Pto}$ -ebselen system was taken from a recently reported structure for the  $M^{Pto}$ -N3 inhibitor (5) (PDB ID: 6LU7); where the inhibitor and crystallographic water molecules were removed before starting simulations. Force field parameters for ebselen were determined using the Antechamber program and described by the general amber force field (30, 31). The partial atomic charges were determined by the restrained electrostatic potential fitting technique. Those electrostatic potential calculations were performed at the HF/6-31G level with Gaussian 09. For  $M^{Pto}$ -apo simulations, we used the recently reported structure (23) (PDB ID: 6M03). Approximately 20,000 TIP3P water model molecules and 4 Na<sup>+</sup> ions were added. All simulations were carried out using the ff14SB force field (32). The simulation protocol included a first minimization of 7000 steps, involving 3500 steepest descent steps followed by 3500 steps of conjugate gradient energy minimization, where constraints were applied on the protein heavy atoms (force constant, 500 kcal/mol-Å<sup>2</sup>) and a second minimization (7000 steps) with no constraints of conjugate gradient energy minimization. Next, during the first equilibration, the temperature was gradually increased from 0 to 300 K over 50 ps using a Langevin thermostat with a temperature coupling constant of 1.0 ps in the canonical ensemble. Density equilibration and production runs were carried out using a constant pressure ensemble (NPT). All simulations were performed using periodic boundary conditions and a 2-fs time step. Long-range electrostatic interactions were calculated using the particle mesh Ewald method with a nonbonded cutoff of 10 Å, and the SHAKE algorithm was used to implement rigid constraints.

## Ebselen global density map

To determine the most probable interaction sites between M<sup>PRO</sup>-ebselen (global hotspot), 15 replicas of 200 ns each were run (a total of 3  $\mu$ s), where configurations were saved every 20 ps. The CPPTRAJ (33) software was used to process the trajectories, where at the first stage all trajectories were aligned by means of minimizing the distance among protein backbone atoms (C, N, and CA). Grid command was used to track the ebselen molecule and to produce a number density map, where the grid resolution was selected to be 0.5  $\text{\AA}$ .

## Strain analysis

Initial M<sup>PRO</sup>-ebselen complex configurations were selected from previous global binding affinity study, with ebselen being bound to the catalytic site, to the intersection between domains II and III, and to both of these sites simultaneously. Three replicas of 200 ns for each of these initial setups were run. As a reference simulation, we used the M<sup>PRO</sup>-apo structure (PDB code: 6M03), and three replicas of 200 ns each were run. In this way, a total of 2400 ns of classical MD were run.

To apply the strain formalism from continuum theory to discrete, atomistic systems, differential operators replace the derivatives (19–21, 34, 35). A radius  $R$  around each central atom  $i$  containing  $n$  other atoms  $j$  defines the local neighborhood around the central atom. The instantaneous position of atom  $i$  at any timestep in the MD is  $\mathbf{x}_i$ , and the position of the same atom at any timestep of the reference simulation is  $\mathbf{x}_{0,i}$ . To first order, the distances between atom  $i$  and its neighbors  $j$  are related through the deformation matrix  $\mathbf{F}$  by  $\mathbf{x}_j - \mathbf{x}_i = \mathbf{F}(\mathbf{x}_{0,j} - \mathbf{x}_{0,i})$ , which forms an overdetermined system of linear equations, and an optimized  $\mathbf{F}^*$  is sought by minimizing the difference between the actual distances and the projected distances to an affine deformation:  $\min \sum_{j=1}^n [\mathbf{x}_j - \mathbf{x}_i - \mathbf{F}(\mathbf{x}_{0,j} - \mathbf{x}_{0,i})]^2$ . The atomic strain tensor is then found by  $\boldsymbol{\epsilon} = \frac{1}{2}(\mathbf{F}^T \mathbf{F} - \mathbf{I})$ , and the magnitude is defined as the L2 norm of the shear part of the strain tensor:  $\text{Tr}(\boldsymbol{\epsilon}^2) - \frac{1}{3}(\text{Tr}(\boldsymbol{\epsilon}))^2$ , since proteins are generally incompressible (19, 21). For this analysis, a radius of 10  $\text{\AA}$  around each atom is considered (21), and only the C $\alpha$  atoms are used in the calculations. The reference simulation is the apo protein trajectory, and the strain is then measured using the ebselen simulations to elucidate the effect of binding of ebselen at the different sites. In addition,  $\beta$  factors were estimated over the same trajectories using the atomic-fluct command of the CPPTRAJ module of AMBER18.

## AQUA-DUCT

To analyze the water structure and water flux in the M<sup>PRO</sup> protein of SARS-CoV-2, we used AQUA-DUCT 1.0.5 (24). We obtained inlets, MAV, and hotspots for water molecules in the M<sup>PRO</sup> catalytic region. For calculations, this region was defined as a 5  $\text{\AA}$  sphere around the center of geometry of the active site residues (H41, C145, H164, and D187) (17). The M<sup>PRO</sup> protein was studied in two different scenarios, with ebselen in the domains II and III site and the apo protein with no ligand. The time window used in both calculations was 50 ns, sampling every 1 ps. Images were created with open-source PyMOL (36).

## Free energy

The absolute binding free energy is defined as follows:  $\Delta G_{\text{binding}} = \Delta G^{\text{L}} - \Delta G^{\text{RL}}$ , where  $\Delta G^{\text{RL}}$  is the free energy change of ebselen annihilation in the M<sup>PRO</sup> complex and  $\Delta G^{\text{L}}$  is the free energy change of ebselen annihilation in water. To calculate these free energy changes, we use TI implemented in PMEMD for AMBER18. We use the one-step annihilation protocol with soft-core potentials (37). In addition,

we adopted a simple approach with multiple runs starting from the most probably binding cluster estimated from previous MD simulations. In this way, three independent replicas for each site were taken into account, as well as three replicas for ebselen solvated in pure water. Eleven equally spaced windows were used ( $\Delta\lambda = 0.1$ ) with 10 ns of simulation time per window. To keep the ligand from wandering in TI calculations, we used a soft restraint of 10 kcal/mol- $\text{\AA}^2$  (37).

## REFERENCES AND NOTES

- N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Wu, G. F. Gao, W. Tan; China Novel Coronavirus Investigating and Research Team, A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* **382**, 727–733 (2020).
- Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. M. Leung, E. H. Y. Lau, J. Y. Wong, X. Xing, N. Xiang, Y. Wu, C. Li, Q. Chen, D. Li, T. Liu, J. Zhao, M. Liu, W. Tu, C. Chen, L. Jin, R. Yang, Q. Wang, S. Zhou, R. Wang, H. Liu, Y. Luo, Y. Liu, G. Shao, H. Li, Z. Tao, Y. Yang, Z. Deng, B. Liu, Z. Ma, Y. Zhang, G. Shi, T. T. Y. Lam, J. T. Wu, G. F. Gao, B. J. Cowling, B. Yang, G. M. Leung, Z. Feng, Early transmission dynamics in wuhan, china, of novel coronavirus-infected pneumonia. *N. Engl. J. Med.* **382**, 1199–1207 (2020).
- P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, H.-R. Si, Y. Zhu, B. Li, C.-L. Huang, H.-D. Chen, J. Chen, Y. Luo, H. Guo, R.-D. Jiang, M.-Q. Liu, Y. Chen, X.-R. Shen, X. Wang, X.-S. Zheng, K. Zhao, Q.-J. Chen, F. Deng, L.-L. Liu, B. Yan, F.-X. Zhan, Y.-Y. Wang, G.-F. Xiao, Z.-L. Shi, A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020).
- F. Wu, S. Zhao, B. Yu, Y.-M. Chen, W. Wang, Z.-G. Song, Y. Hu, Z.-W. Tao, J.-H. Tian, Y.-Y. Pei, M.-L. Yuan, Y.-L. Zhang, F.-H. Dai, Y. Liu, Q.-M. Wang, J.-J. Zheng, L. Xu, E. C. Holmes, Y.-Z. Zhang, A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265–269 (2020).
- Z. Jin, X. Du, Y. Xu, Y. Deng, M. Liu, Y. Zhao, B. Zhang, X. Li, L. Zhang, C. Peng, Y. Duan, J. Yu, L. Wang, K. Yang, F. Liu, R. Jiang, X. Yang, T. You, X. Liu, X. Yang, F. Bai, H. Liu, X. Liu, L. W. Guddat, W. Xu, G. Xiao, C. Qin, Z. Shi, H. Jiang, Z. Rao, H. Yang, Structure of M<sup>PRO</sup> from SARS-CoV-2 and discovery of its inhibitors. *Nature* **582**, 289–293 (2020).
- A. Zumla, J. F. W. Chan, E. I. Azhar, D. S. C. Hui, K.-Y. Yuen, Coronaviruses — drug discovery and therapeutic options. *Nat. Rev. Drug Discov.* **15**, 327–347 (2016).
- K. Anand, G. J. Palm, J. R. Mesters, S. G. Siddell, J. Ziebuhr, R. Hilgenfeld, Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra  $\alpha$ -helical domain. *EMBO J.* **21**, 3213–3224 (2002).
- H. Yang, M. Yang, Y. Ding, Y. Liu, Z. Lou, Z. Zhou, L. Sun, L. Mo, S. Ye, H. Pang, G. F. Gao, K. Anand, M. Bartlam, R. Hilgenfeld, Z. Rao, The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 13190–13195 (2003).
- N. Singh, A. C. Halliday, J. M. Thomas, O. V. Kuznetsova, R. Baldwin, E. C. Y. Woon, P. K. Alely, I. Antoniadou, T. Sharp, S. R. Vasudevan, G. C. Churchill, A safe lithium mimetic for bipolar disorder. *Nat. Commun.* **4**, 1332 (2013).
- E. Lynch, J. Kil, Development of Ebselen, a glutathione peroxidase mimic, for the prevention and treatment of noise-induced hearing loss. *Semin. Hear.* **30**, 47–55 (2009).
- J. Kil, E. Lobarinas, C. Spankovich, S. K. Griffiths, P. J. Antonelli, E. D. Lynch, C. G. Le Prell, Safety and efficacy of ebselen for the prevention of noise-induced hearing loss: A randomised, double-blind, placebo-controlled, phase 2 trial. *Lancet* **390**, 969–979 (2017).
- S. Thangamani, W. Younis, M. N. Seleem, Repurposing ebselen for treatment of multidrug-resistant staphylococcal infections. *Sci. Rep.* **5**, 11596 (2015).
- L. Zou, J. Lu, J. Wang, X. Ren, L. Zhang, Y. Gao, M. E. Rottenberg, A. Holmgren, Synergistic antibacterial effect of silver and ebselen against multidrug-resistant Gram-negative bacterial infections. *EMBO Mol. Med.* **9**, 1165–1178 (2017).
- M. Renson, E. Etschenberg, J. Winkelmann, 2-Phenyl-1,2-benzisoxaselenazol-3(2H)-one containing pharmaceutical preparations and process for the treatment of rheumatic diseases (1982); <https://patents.google.com/patent/US4352799A/en>.
- C. Masaki, A. L. Sharpley, C. M. Cooper, B. R. Godlewska, N. Singh, S. R. Vasudevan, C. J. Harmer, G. C. Churchill, T. Sharp, R. D. Rogers, P. J. Cowen, Effects of the potential lithium-mimetic, ebselen, on impulsivity and emotional processing. *Psychopharmacology (Berl)* **233**, 2655–2661 (2016).
- C. Li, Y. Qi, X. Teng, Z. Yang, P. Wei, C. Zhang, L. Tan, L. Zhou, Y. Liu, L. Lai, Maturation mechanism of severe acute respiratory syndrome (SARS) coronavirus 3C-like proteinase. *J. Biol. Chem.* **285**, 28134–28140 (2010).
- M. Bzówka, K. Mitusińska, A. Raczynska, A. Samol, J. A. Tuszyński, A. Góra, Structural and evolutionary analysis indicate that the SARS-CoV-2 Mpro is an inconvenient target for small-molecule inhibitors design. *bioRxiv* 2020.02.27.968008 (2020).

18. J. W. Kaus, L. T. Pierce, R. C. Walker, J. A. McCammon, Improving the efficiency of free energy calculations in the amber molecular dynamics package. *J. Chem. Theory Comput.* **9**, 4131–4139 (2013).
19. M. R. Mitchell, T. Tlusty, S. Leibler, Strain analysis of protein structures and low dimensionality of mechanical allosteric couplings. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E5847–E5855 (2016).
20. P. M. Gullett, M. F. Horstemeyer, M. I. Baskes, H. Fang, A deformation gradient tensor and strain tensors for atomistic simulations. *Model. Simul. Mater. Sci. Eng.* **16**, 015001 (2007).
21. J.-P. Eckmann, J. Rougemont, T. Tlusty, *Colloquium: Proteins: The physics of amorphous evolving matter.* *Rev. Mod. Phys.* **91**, 031001 (2019).
22. G. J. Papakonstantopoulos, M. Doxastakis, P. F. Nealey, J.-L. Barrat, J. J. de Pablo, Calculation of local mechanical properties of filled polymers. *Phys. Rev. E* **75**, 031803 (2007).
23. R. P. D. Bank, RCSB PDB - 6M03: The crystal structure of COVID-19 main protease in apo form; [www.rcsb.org/structure/6M03](http://www.rcsb.org/structure/6M03).
24. T. Magdziarz, K. Mitusińska, M. Bzówka, A. Raczynska, A. Staińczak, M. Banas, W. Bagrowski, A. Góra, AQUA-DUCT 1.0: Structural and functional analysis of macromolecules from an intramolecular voids perspective. *Bioinformatics* **36**, 2599–2601 (2020).
25. K. Anand, J. Ziebuhr, P. Wadhvani, J. R. Mesters, R. Hilgenfeld, Coronavirus main proteinase (3CL<sup>pro</sup>) structure: Basis for design of anti-SARS drugs. *Science* **300**, 1763–1767 (2003).
26. J. D. R. Knight, D. Hamelberg, J. A. McCammon, R. Kothary, The role of conserved water molecules in the catalytic domain of protein kinases. *Proteins* **76**, 527–535 (2009).
27. M. Kandeel, M. Al-Nazawi, Virtual screening and repurposing of FDA approved drugs against COVID-19 main protease. *Life Sci.* **251**, 117627 (2020).
28. C. Liu, Q. Zhou, Y. Li, L. V. Garner, S. P. Watkins, L. J. Carter, J. Smoot, A. C. Gregg, A. D. Daniels, S. Jervey, D. Albaiu, Research and development on therapeutic agents and vaccines for COVID-19 and related human coronavirus diseases. *ACS Cent. Sci.* **6**, 315–331 (2020).
29. D. A. Case, I. Y. Ben-Shalom, S. R. Brozell, D. S. Cerutti, T. E. Cheatham III, V. W. D. Cruzeiro, T. A. Darden, R. E. Duke, D. Ghoreishi, M. K. Gilson, H. Gohlke, A. W. Goetz, D. Greene, R. Harris, N. Homeyer, Y. Huang, S. Izadi, A. Kovalenko, T. Kurtzman, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. J. Mermelstein, K. M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D. R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C. L. Simmerling, J. Smith, R. Salomon-Ferrer, J. Swails, R. C. Walker, J. Wang, H. Wei, R. M. Wolf, X. Wu, L. Xiao, D. M. York, P. A. Kollman, *AMBER 2018* (University of California, San Francisco, 2018).
30. J. Wang, W. Wang, P. A. Kollman, D. A. Case, Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* **25**, 247–260 (2006).
31. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174 (2004).
32. J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).
33. D. R. Roe, T. E. Cheatham III, PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* **9**, 3084–3095 (2013).
34. D. Dube, N. Ahalawat, H. Khandelia, J. Mondal, S. Sengupta, On identifying collective displacements in apo-proteins that reveal eventual binding pathways. *PLoS Comput. Biol.* **15**, e1006665 (2019).
35. M. L. Falk, J. S. Langer, Dynamics of viscoplastic deformation in amorphous solids. *Phys. Rev. E* **57**, 7192–7205 (1998).
36. W. L. DeLano, Pymol: An open-source molecular graphics tool. *CCP4 Newsletter on protein crystallography* **40**, 82–92 (2002).
37. T. Steinbrecher, I. Joung, D. A. Case, Soft-core potentials in thermodynamic integration: Comparing one- and two-step transformations. *J. Comput. Chem.* **32**, 3253–3263 (2011).

**Acknowledgments:** We are grateful to the Research Computing Center of the University of Chicago for computational resources. **Funding:** The authors are grateful to Andrzej Joachimiak and Rick Stevens for helpful discussions. The study of protein-DNA complexes in epigenetics is supported by the NSF, under grant EFRI CEE 1830969, and the development of new DNA models is supported by BIO/MCB 1818328. The study of allostery and its role in triggerable materials is supported by MURI: W911NF-15-1-0568. The development of advanced sampling simulation algorithms and efficient software is supported by the Department of Energy, Office of Science, Basic Energy Sciences, Division of Materials Science and Engineering. The simulations reported here were carried out on the GPU cluster supported by the NSF through grant DMR-1828629. **Author contributions:** C.A.M. and J.J.d.P. conceived and started the project. C.A.M., F.B., G.R.P.-L., and W.A. ran simulations, calculated properties, and analyzed results. J.J.d.P. guided the project and discussed analysis. All authors contributed to writing. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper. Additional data related to this paper may be requested from the authors.

Submitted 27 May 2020  
 Accepted 28 July 2020  
 Published First Release 14 August 2020  
 Published 11 September 2020  
 10.1126/sciadv.abd0345

**Citation:** C. A. Menéndez, F. Bylén, G. R. Perez-Lemus, W. Alvarado, J. J. de Pablo, Molecular characterization of ebselen binding activity to SARS-CoV-2 main protease. *Sci. Adv.* **6**, eabd0345 (2020).