

## CORONAVIRUS

# Toward wide-spectrum antivirals against coronaviruses: Molecular characterization of SARS-CoV-2 NSP13 helicase inhibitors

Gustavo R. Perez-Lemus<sup>1†</sup>, Cintia A. Menéndez<sup>1,2†</sup>, Walter Alvarado<sup>1,3†</sup>, Fabian Byléhn<sup>1†</sup>, Juan J. de Pablo<sup>1,4\*</sup>

To date, effective therapeutic treatments that confer strong attenuation against coronaviruses (CoVs) remain elusive. Among potential drug targets, the helicase of CoVs is attractive due to its sequence conservation and indispensability. We rely on atomistic molecular dynamics simulations to explore the structural coordination and dynamics associated with the SARS-CoV-2 Nsp13 apo enzyme, as well as their complexes with natural ligands. A complex communication network is revealed among the five domains of Nsp13, which is differentially activated because of the presence of the ligands, as shown by shear strain analysis, principal components analysis, dynamical cross-correlation matrix analysis, and water transport analysis. The binding free energy and the corresponding mechanism of action are presented for three small molecules that were shown to be efficient inhibitors of the previous SARS-CoV Nsp13 enzyme. Together, our findings provide critical fresh insights for rational design of broad-spectrum antivirals against CoVs.

## INTRODUCTION

Severe acute respiratory syndrome coronavirus (SARS-CoV), which emerged in 2003, was the first CoV identified as a severe human pathogen. CoVs are positive-strand RNA (+RNA) viruses belonging to the order Nidovirales (1). Ten years later, a similar CoV known as the Middle East respiratory syndrome CoV (MERS-CoV) arose and was found to have higher fatality rates than the first SARS-CoV. On December 2019, a new CoV of zoonotic origin, SARS-CoV-2, appeared in Wuhan City, Hubei Province, China (2–5). SARS-CoV-2, a  $\beta$ -CoV, is the etiological agent responsible for the 2019–2020 viral pneumonia coronavirus disease 2019 (COVID-19) outbreak. Targeted therapeutics for CoVs are not available, and at this point in time, effective treatment options for CoVs remain limited.

The nidovirus contains the largest known RNA genome (6). This genome encodes two large polyproteins, pp1a and pp1ab, which, after being proteolytically processed, produce 16 nonstructural proteins (Nsp), including primase (Nsp8), RNA-dependent RNA polymerase RdRp (Nsp12), and helicase (Nsp13). These three enzymes and other Nsps are components of a replication and transcription complex (RTC) that is essential for the virus life cycle. The helicase is essential for replication in the nidovirus equine arteritis virus (7) and the  $\beta$ -CoV murine hepatitis virus (8) and is presumed to be essential in all nidoviruses (9). Previous studies of SARS-CoV Nsp13 helicase have shown that this enzyme unwinds double-stranded nucleic acids (either DNA or RNA), with a 5' > 3' polarity, into two single strands by relying on the energy derived from nucleotide hydrolysis (10). Nsp13 can target all natural nucleotides and deoxynucleotides as substrates when performing its adenosine triphosphatase (ATPase) activity (11, 12). Nsp13 is a member of the SF1

helicase family, where the Walker A (phosphate loop) and Walker B ( $Mg^{2+}$  binding site) motifs play a fundamental role in the adenosine triphosphate (ATP) binding and catalysis (13). In addition, it has been shown that both SARS-CoV Nsp12 and MERS Nsp12 can enhance the helicase and ATPase activities of the SARS-CoV Nsp13 enzyme (10). The mechanistic role of the helicase in SARS-CoV-2 replication and transcription has yet to be established. However, a structural basis for the Nsp12-Nsp13 coupling in the SARS-CoV-2 RTC complex has recently become available. A stable Nsp13:holo-RdRp:RNA complex has been produced, and its structure has been reconstructed by cryo-electron microscopy (14). From this study, a possible role for Nsp13 has been proposed in the generation of backtracked replication transcription complexes for proofreading, template switching during subgenomic RNA transcription, or both (14). From the aforementioned structure, it was also shown that the primary interaction determinant of the helicase with the RTC occurs between the Nsp13 protein through its zinc binding domain (ZBD) and the Nsp8 extensions (14). The principal SARS-CoV-2 Nsp13 domains are highlighted in Fig. 1. The SARS-CoV-2 Nsp13 protein represents an ideal target for development of anti-CoV drugs because of its sequence conservation and indispensable role across all CoV species (15, 16). The primary amino acid sequence of SARS-CoV Nsp13 and SARS-CoV-2 Nsp13 shares a sequence identity of 99.8% and a sequence similarity of 100.0% (17). This implies that previously found therapeutics that target this enzyme in other CoVs species might be effective against the new COVID-19 disease.

The repurposing of approved pharmaceutical drugs and drug candidates already in clinical trials provides an essential approach to identify drugs with clinical potential against new infectious diseases. A particularly attractive strategy is based on targeting key enzymes such as proteases, polymerases, and helicases, which are all involved in the replication cycle of the virus. Good examples of antiviral drugs with a protease and polymerase mechanism of action are indinavir, saquinavir, lopinavir/ritonavir, and remdesivir, which have recently gained attention in clinical trials against COVID-19 infection (18). Unlike the structural proteins of CoVs such as spike

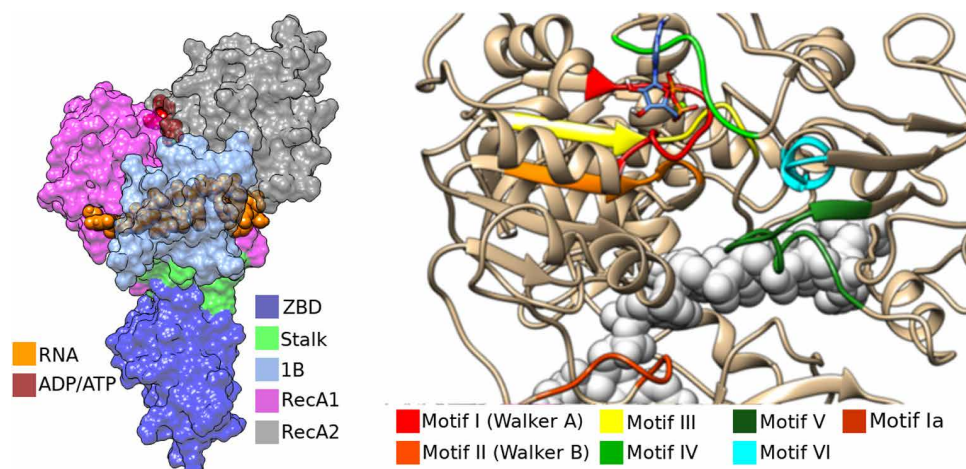
Copyright © 2022  
The Authors, some  
rights reserved;  
exclusive licensee  
American Association  
for the Advancement  
of Science. No claim to  
original U.S. Government  
Works. Distributed  
under a Creative  
Commons Attribution  
NonCommercial  
License 4.0 (CC BY-NC).

Downloaded from https://www.science.org on March 08, 2022

<sup>1</sup>Pritzker School of Molecular Engineering, University of Chicago, Chicago, IL 60637, USA. <sup>2</sup>INQUISUR, Departamento de Química, Universidad Nacional del Sur (UNS)-CONICET, Avenida Alem 1253, 8000 Bahía Blanca, Argentina. <sup>3</sup>Biophysical Sciences, University of Chicago, Chicago, IL 60637, USA. <sup>4</sup>Argonne National Laboratory, 9700 Cass Ave, Lemont, IL 60439 USA.

\*Corresponding author. Email: depablo@uchicago.edu

†These authors contributed equally to this work.



**Fig. 1. Nsp13 protein structure.** (Left) Surface representation of the Nsp13 model, with the principal domains highlighted with colors: ZBD (dark blue), Stalk (gray), RecA1 (pink), RecA2 (dark gray), and 1B (blue). (Right) Conserved motifs in Nsp13 among SF1 helicase family: motif I (red), motif II (orange), motif III (yellow), motif IV (green), motif V (dark green), motif VI (cyan), and motif Ia (dark orange), along with the ADP binding mode (light blue in sticks) and the RNA (white spheres). Protein images generated with Chimera (38).

(S) proteins, which exhibit a great deal of variability among CoV species, NSPs such as helicase (Nsp13) have been shown to be highly conserved. As noted earlier, the primary amino acid sequence of the Nsp13 protein from the previous SARS-CoV and SARS-CoV-2 Nsp13 shares a sequence identity of 99.8% and a sequence similarity of 100.0% (17). This means that previously reported drugs that target these enzymes might be effective against the new SARS-CoV-2 system. A recent study has shown that it is possible to target multiple CoVs through broad-spectrum inhibitors (19).

More specifically, Nsp13 inhibition could be achieved through ATP binding, direct nucleoside triphosphatase (NTPase) activity, nucleic acid binding to the helicase, blocking helicase translocation, etc. Several compounds have been reported as Nsp13 inhibitors; benzotriazole, imidazole, imidazodiazepine, phenothiazine, quinoline, triphenylmethane, pyrrole, acridone, small peptides, and bananin derivatives all constitute potential new agents for SARS-CoV-2 treatment (20). Here, we have selected three different ligands belonging to some of the aforementioned families: bananin (21), SSYA10-001 (19), and chromone-4c (22); the rationale behind our choices is provided below. In what follows, we explore the binding mechanism for each of them.

Bananins represent a family of antiviral compounds with a unique structural signature that incorporates a trioxa-adamantane moiety covalently bound to a pyridoxal derivative. Bananin deserves special attention because it was an effective inhibitor of the ATPase and helicase activities of the previous SARS-CoV helicase (21). In a cell culture system of SARS-CoV, bananin exhibited a median effective concentration ( $EC_{50}$ ) of less than 10  $\mu$ M (21). This adamantane derivative acts as a noncompetitive inhibitor of the ATPase activity of the SARS-CoV helicase with respect to both ATP and nucleic acids. This suggests that bananin displays Nsp13 inhibition by binding at a site different from the highly conserved ATP and nucleic acid binding sites (20). Note that the human pharmacology for several adamantane derivatives has already been studied, thereby providing a starting point for possible clinical trials. Recent reports have argued in favor of the potential protective effects of adamantanes against COVID-19 (23); if confirmed for larger sample sizes, adamantanes might prove helpful in limiting SARS-CoV-2 infection.

Previous studies have demonstrated through virological, biochemical, and molecular modeling that SSYA10-001, a 1,2,4-triazole, exhibits notable antiviral activity against multiple CoVs (19). SSYA10-001 is an efficient inhibitor of the helicase activity of the Nsp13 enzyme and is a strong replication inhibitor of at least three CoVs: the previous SARS-CoV, mouse hepatitis virus (MHV), and MERS-CoV (19). This evidence suggests that the binding pocket of SSYA10-001 to Nsp13 helicase is conserved among different CoVs and supports the notion of targeting multiple CoVs through broad-spectrum inhibitors. Here, we explore its potential mechanism of inhibition for the new SARS-CoV-2 Nsp13 protein.

Naturally occurring chemicals are also considered to be a promising source of potential medications against a wide range of diseases. In the case of SARS-CoV, 64 natural chemical compounds were examined for their effect over helicase activity (24). Relying on the aforementioned antecedent, the pharmacophore space of SARS-CoV NTPase/helicase by synthetic dihydroxychromone derivatives has been studied (22). Three classes of dihydroxychromone derivatives were synthesized, and a pharmacophore model of SARS-CoV NTPase/helicase was proposed. The model includes a diketone acid core, a hydrophobic arylmethyl substituent, and a free catechol unit (22). Among this series, chromone-4c exhibits the highest SARS-CoV helicase activity and marginal ATP/ase activity. Here, we explore the binding mechanism of this dihydroxychromone derivative, which has shown Nsp13 inhibition.

In this work, we present results of atomistic molecular dynamics (MD) simulations that provide useful, previously unknown insights into the structure and dynamics associated with the Nsp13 apo enzyme, as well as their complexes with the natural ligands. At the time of writing this report, the experimental structure of the SARS-CoV-2 Nsp13 protein is not available in the apo form or in complex with its natural ligands [nucleoside triphosphate (NTP) nucleotide and the oligonucleotides]. We have therefore constructed a homology model based on the experimental structure reported for SARS-CoV Nsp13 (10, 25). We first evaluate the impact of the binding of the different Nsp13 natural ligands on the protein molecular stiffness and molecular strain to unravel any communication patterns among

distal sites, if present. In addition, we examine how ligand binding modifies the structure and transport of water in the Nsp13 catalytic site. Second, we explore the potential interaction sites for three small organic molecules that are efficient inhibitors of the helicase or ATPase activity of the previous SARS-CoV Nsp13 protein: bananin (21), SSYA10-001 (19), and chromone-4c (22). We estimate their absolute binding free energy to arrive at a quantitative measure of their potential effectiveness. The protein molecular stiffness, molecular strain, and the structure and transport of water in the Nsp13 catalytic site are also examined for the different small organic molecules, in efforts to unravel the mechanism of Nsp13 inhibition at the molecular level. We conclude with suggestions for future work.

## RESULTS

### Structure and dynamics of the Nsp13 apo protein and their complexes with ADP and ssRNA natural ligands

MD simulations of seven different complexes were carried out using the Amber18 package (26). Three different atomistic trajectories were collected for each complex. The full list is compiled in Table 1, which includes the total simulation times. The first four complexes involve the Nsp13 monomer with natural ligands: adenosine diphosphate (ADP) and single-stranded RNA (ssRNA) (10, 25), while the last three complexes consider the conformational and energetic changes induced by drugs that were previously tested on SARS-CoV helicase (19, 21, 22). As indicated earlier, at the beginning of this project, an experimental structure of the SARS-CoV-2 helicase was not available, so a SwissModel (25) based on the SARS-CoV helicase was adopted [Protein Data Bank (PDB): 6JYT]. The left side of Fig. 1 shows the general shape of the Nsp13 monomer with its principal domains: ZBD (1 to 100), Stalk domain (101 to 150), 1B domain (151 to 261), Rec1A domain (262 to 442), and Rec2A (442 to 596), among with the ATP/ADP binding sites and the ssRNA binding site. As member of the SF1 family, Nsp13 has seven characteristic motifs that participate in both ATP and ssRNA binding (27). The right side of Fig. 1 emphasizes the conserved helicase motifs in the active site of the protein, which are motifs I (281 to 291) (also known as Walker A motif), Ia (340 to 347), (Walker B) II (373 to 379), III (398 to 406), IV [also known as IIIa (28)] (438 to 445), V (532 to 546), and VI (565 to 572). Figure S1 shows the NSP13 sequence of the highlighted motifs.

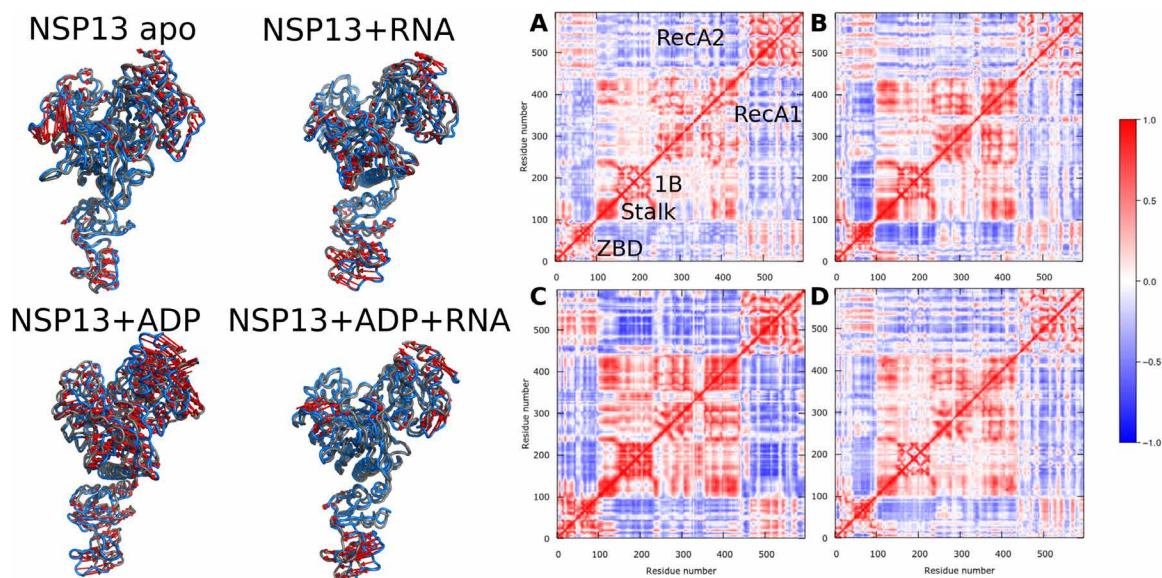
**Table 1. The Nsp13-ligand complexes that were studied by means of atomistic molecular dynamics simulations Nsp13 monomer with their natural ligands and small organic molecules and the corresponding simulation times.**

Complex	Simulation time (ns)
Nsp13 apo	1800
Nsp13+ADP	1800
Nsp13+ssRNA	1800
Nsp13+ssRNA+ADP	1800
Nsp13+ssRNA+ADP+Bananin	1500
Nsp13+ssRNA+ADP+SSYA10-001	1500
Nsp13+ssRNA+ADP+Chromone-4c	1500

Principal components analysis (PCA) was used to identify the dominant motions of the Nsp13 protein and relate them to the presence or absence of the ligands (29). Correlations between domains in the protein were determined by computing a dynamical cross-correlation matrix (DCCM) (29). Figure 2 shows the first mode of the PCA for the Nsp13 enzyme in the apo state, as well as Nsp13+ADP, Nsp13+ssRNA, and Nsp13+ADP+ssRNA, where both Nsp13's binding sites are occupied simultaneously. In the Nsp13 apo form, the first mode shows a limited movement of the protein, with the ZBD and the RecA1 and RecA2 domains being the more active parts of the protein. The DCCM map for the apo form exhibits the main domains of the protein clearly defined, with high intradomain positive correlation and general reduced interdomain correlation. The highest interdomain anticorrelation is located between the ZBD and the Stalk domains (Fig. 2A). A number of visible changes occur when ssRNA binds to the protein. In the first PCA mode, a clear coordinated movement appears between the ZBD and the 1B domains, where they move toward each other (Fig. 2, left). In DCCM (Fig. 2B), a definite anticorrelation emerges between the previously mentioned domains, and a weak one between ZBD and the 330 to 420 residues of the RecA1 domain; at the same time, the anticorrelation in the ZBD and the Stalk domains is increased with respect to the apo form. On the other hand, the correlations in the 320 to 350 regions that include motif Ia, which is an RNA binding region of the protein, are reduced with respect to the apo form. The addition of ADP to the Nsp13 protein induces the largest changes observed in DCCM (Fig. 2C). For this complex, the domains Stalk, 1B, and RecA1 become completely positively correlated, and they show an anticorrelation with ZBD and RecA2. The latter two domains correlate positively with a small band of anticorrelations in residues 440 to 460. The PCA first mode shows a coordinated movement between ZBD and other upper domains of the protein that resembles an opening of the upper part of the protein. Figure 2D is DCCM for the system that includes both natural ligands simultaneously, ssRNA and ADP. Therefore, the correlation map shows a behavior intermediate between those of Fig. 2 (B and C), with the domains Stalk and RecA1 correlated, similarly to the Nsp13+ADP system, whereas the 1B domain shows a clear reduction in the number of correlations with the RecA1 domain compared with the behavior observed in the Nsp13+ADP system. The anticorrelations between ZBD and the three domains listed above persist as in the Nsp13+ADP complex. The same is true for the correlations between RecA2 and Stalk.

In summary, the results from PCA and the DCCM matrix reveal a tightly coupled structural assembly characterized by a complex communication network among the five principal protein domains; moreover, this network presents a distinct configuration depending on whether the natural ligands are present or not. For example, the transition from Nsp13 in the apo form to the Nsp13+ADP system is accompanied by the emergence of a high degree of interdomain communication that is absent in the apo state of the enzyme. A similar response is also evident when both natural ligands, ADP and ssRNA, are incorporated into the assembly. This last complex was therefore selected as the reference state for all subsequent studies presented in the next section, which involve the incorporation of small drug molecules. This choice is motivated by the higher level of response of the global system compared with the apo state, which might be useful in exposing the effect of nonnatural ligands on the assembly.

We have also performed a detailed analysis of water behavior in the active site of the protein. This analysis has been shown to be



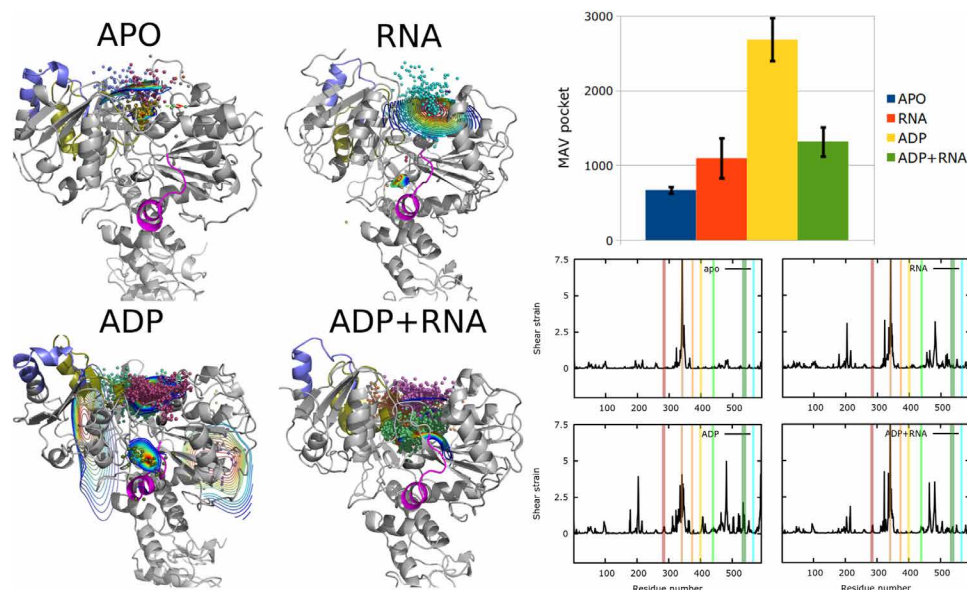
**Fig. 2. Principal components analysis (PCA) and the dynamical cross-correlation matrix (DCCM).** Left: Movement representation of the first mode of the PCA of the  $\alpha$  carbons of Nsp13 in the presence of nucleotides. Right: Dynamical cross-correlation network for  $\alpha$  carbons in (A) Nsp13 apo, (B) Nsp13+RNA, (C) Nsp13+ADP, and (D) Nsp13+ADP+RNA.

helpful in the study of different proteins, where the description of pocket sizes and tunnel formation provides insight into protein function (30). The left side of Fig. 3 shows the water inlets, the surfaces of water pores, and three highlighted regions of the Nsp13 protein. These highlighted regions correspond to the 427 to 437 residues for region 1 (magenta in Fig. 3), the 496 to 511 residues for region 2 (slate blue), and the 512 to 542 residues for region 3 (in olive). In a previous study, those regions were analyzed to study the dynamics of the helicase, where region 1 was related to the opening and closing of the RecA domains, and regions 2 and 3 were related to the ssRNA binding (10). Given their importance in SARS-CoV helicase, these regions were analyzed by surface-accessible solvent area (SASA), with the LCPO (linear combination of pairwise overlaps) algorithm as implemented in CPPTRAJ (31) to measure conformational changes upon ligand binding. The results are included in Table 2, where, in the case of the Nsp13+ADP complex, regions 1 and 2 experience minimal changes, while region 3 increases their surface with respect to the apo state: Region 3 experiences the largest variance. When the ssRNA is bound to the protein, a similar trend is observed in the sense that region 2 minimally reduces its values and region 3 exhibits an increase in surface area with the largest variance with respect to the apo form. However, region 1 exhibits a small reduction of surface exposure that was not evident before. Upon addition of ADP to the Nsp13+ssRNA complex, regions 2 and 3 exhibit an increase in surface exposure compared with the previous Nsp13+ssRNA assembly, and region 1 presents a reduction in the exposed area. These results show again that a complex cross-communication system exists between the ADP binding site and the oligonucleotide binding region. When ADP is bound to the Nsp13 enzyme in the active site, region 3, which binds ssRNA and is located in a distal position with respect to the NTP nucleotide binding site, experiences a marked increase in its solvent surface compared with the apo state, meaning that it separates from the bulk protein. Then, when ssRNA is attached to the protein, region 1 experiences a reduction

of its surface that is correlated with a higher proximity between the RecA1 and RecA2 domains. For regions 2 and 3, where the ssRNA binds, a lower exposure compared to the apo form is observed, as expected because of direct interactions with the oligonucleotide. When ADP is added to the complex, we observe a small reduction of the RecA1-RecA2 interface and a small detaching of the ssRNA in regions 2 and 3, as revealed by a higher solvent surface compared to the Nsp13+ssRNA system.

Returning to the water analysis, as shown in the left panels of Fig. 3, when ADP is present, a clear cluster of water inlets appears in the vicinity of region 1, which is related to an opening between the RecA1 and RecA2 domains. This water pore also appears when ssRNA is bound to the protein and disappears when both binding regions in Nsp13 are occupied at the same time by ADP and ssRNA. To elucidate the structural role of water, we calculate the maximum available volume (MAV) of the NTP nucleotide binding region. From the results included in the histogram of Fig. 3, we can see how the presence of ADP causes the RecA1 and RecA2 regions to separate, thereby generating a larger volume in the active site. The presence of ssRNA expands this volume slightly, compared with the apo system, whereas the subsequent addition of ADP to the Nsp13+ssRNA assembly creates another increase in the volume.

The NTP active site of Nsp13 is examined in more detail at the bottom right of Fig. 3, which shows the strain graphs for the four complexes shown on the left panel. Strain is considered a useful quantity for the analysis of local protein deformations, which serves to underscore the mechanical allosteric coupling between two regions on the same protein (32). This deformation is calculated as the shear strain measured relative to the average conformation of the Nsp13 apo form. The shear strain in the apo form points to a high strain in motif Ia, which intervenes in the helicase activity (10). In addition, a region in the 1B domain (190 to 210) exhibits a visible strain along with the region comprised by residues 480 to 490. These two regions mediate RNA binding. When ssRNA is present,



**Fig. 3. Characterization of the water behavior in the active site of the Nsp13 protein and shear strain analysis.** (Left) Water for the different Nsp13 systems. Cartoon representation of Nsp13 protein, with regions 1, 2, and 3 highlighted in colors: magenta (427 to 437), slate blue (496 to 511), and olive (512 to 542). The water inlet clusters are represented in colored spheres, with cluster area in rainbow contour maps. (Top right) Histogram of the maximum available volume (MAV) in  $\text{\AA}^3$  for the NTP active pocket for the four complexes on the left. (Bottom right) Strain analysis of the four complexes, with the seven principal motifs highlighted in the same colors as represented in Fig. 1 (right).

**Table 2. SASA estimated by the LCPO algorithm.** For all the systems, the SASA values for regions 1, 2, and 3 are reported in  $\text{\AA}^2$ . In all cases, it was estimated from over the heavy backbone atoms and the amide hydrogen belonging to each region.

Complex	Region 1: 427–437 residues ( $\text{\AA}^2$ )	Region 2: 496–511 residues ( $\text{\AA}^2$ )	Region 3: 511–542 residues ( $\text{\AA}^2$ )
NSP13 apo	208.58 $\pm$ 13.46	154.88 $\pm$ 20.72	284.15 $\pm$ 30.48
NSP13+ADP	220.22 $\pm$ 18.57	150.17 $\pm$ 21.64	364.24 $\pm$ 43.54
NSP13+ssRNA	201.48 $\pm$ 15.94	157.65 $\pm$ 24.42	299.90 $\pm$ 49.76
NSP13+ssRNA+ADP	188.56 $\pm$ 31.43	189.64 $\pm$ 43.54	358.06 $\pm$ 27.47

we see a spreading and increase of strain into the residues involved in the oligonucleotide binding. In the presence of ADP, the protein increases its strain in the 1B and RecA2 ssRNA binding regions and reduces the strain in the Ia motif. These results reveal the existence of a communication network among these distal sites of the Nsp13 protein and a corresponding allosteric behavior, as expected for the helicase activity. In addition, when ADP is present in the system, the Walker A motif, motif IV, motif V, and the contiguous regions from motifs II and III show a small strain signal, which was not evident for the previous Nsp13 conditions; in contrast, motif VI does not exhibit any shear strain, following the same trend shown by Nsp13 in the apo form and the Nsp13+ssRNA assembly. When both binding regions of Nsp13 are occupied at the same time by ADP and ssRNA, the shear at the RecA1 RNA binding site increases and spreads considerably, while in the 1B domain, the strain disappears around residue number 190 and increases in the region between 200 and 210. For the RecA2 site, the strain raises among residues 470 to 500, and the signal in motif V is reduced considerably compared to the Nsp13+ADP system. Together, the results outlined above present a general landscape of the structure, dynamics, and hydration properties of the SARS-CoV-2 Nsp13 protein in its native state and help

reveal key roles for selected regions that might be of use in our subsequent comparative analysis of potential inhibitors.

### Inhibitors of the NSP13 helicase activities and replication of SARS-CoV

To identify potential binding sites at the SARS-CoV-2 helicase, the ligand candidates were docked using the AutoDock4.2 software (33). Two different binding sites were identified; the bananin binding site is located at the Rec1A domain, whereas SSYA10-001 and chromone-4c bind at the same pocket located in the RecA2 domain. SSYA10-001 and chromone-4c, which are efficient inhibitors of the helicase activity of the previous SARS-CoV Nsp13 enzyme, both bind to the same region of the protein.

The binding positions identified by means of docking simulations were subject to further refinement through atomistic MD simulations with explicit water. In all cases, three replicas with a total of 1500-ns simulations of classical MD were carried out. Figure 4 shows Nsp13-ligand interactions for the different complexes after MD refinement.

To arrive at a quantitative estimate of the binding affinity of each ligand, we used thermodynamic integration (TI) and determined the absolute binding free energy for the different Nsp13 regions. In

all cases, 11 windows per integration and 20 ns per window were taken into account. 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 and their associated SDs. The results are shown in Table 3.

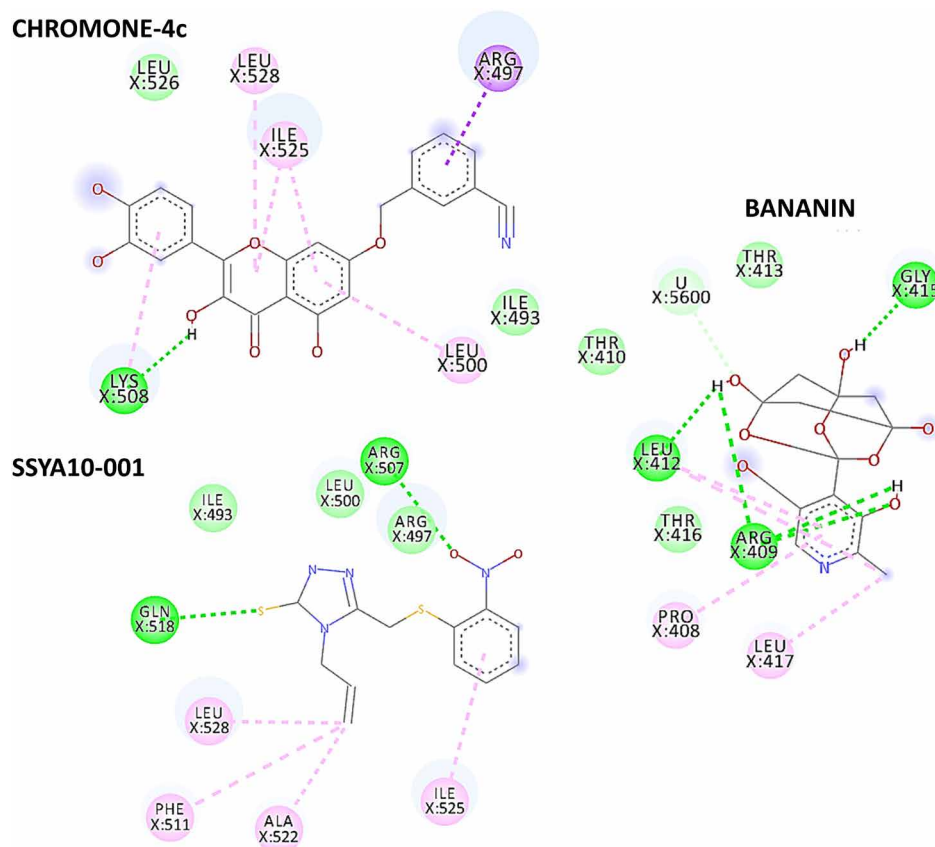
The absolute binding free energies corresponding to the different nonnatural ligands are negative, serving to underscore the thermodynamic stability of the nsp13-ligand complexes. SSYA10-001 and chromone-4c, which bind at the RecA2 domain, exhibit the highest affinities. Moreover, the calculated free energies for SSYA10-001 and chromone-4c correspond to a micromolar order for the constant inhibition, which are the same order of magnitude than the experimental results (19, 21, 22, 24), while the calculated energy for bananin underestimates its experimental inhibition value.

All the Nsp13+ADP+ssRNA+drug assemblies considered here are thermodynamically stable; their mechanism of inhibition of each drug, however, is different. A detailed analysis of the water behavior in the active site of the protein was conducted using the AQUADUCT 1.0.5 software (30). In particular, we studied water structure and water flux in the Nsp13+ligand complexes. The time window was selected as 100 ns, and sampling was performed every 10 ps in all cases. The results are shown in Fig. 5. The left side of Fig. 5A represents the water averaged paths, the surface of water pores, and the MAV of the ATP/ADP binding site (included in the inset histogram).

In all cases, regions 1, 2, and 3, as described in the previous sections, are highlighted using the same color code as in Fig. 3. In all images, the binding site of the ligands is colored in brown.

From Fig. 5A, the presence of water pores below the RecA1 and RecA2 domains when SSYA10-001 and chromone-4c are included is evident. This contrasts with the reference state—when both binding regions in Nsp13 protein are occupied simultaneously by the natural ligands (ADP and ssRNA). In the absence of a small molecule, these water channels, which expand from the ATP/ADP binding site to the ssRNA/ssDNA binding region, are closed. The binding of SSYA10-001 and chromone-4c in the upper part of RecA therefore generates an important conformational change. This change occurs far away from the binding site, showing that both small molecules can trigger allosteric changes that open the water channels. Our results for the MAV of the ADP binding region are shown in the histogram of Fig. 5 and reported as the relative value to the reference system (Nsp13+ADP+ssRNA assembly with no ligand). We can see how the presence of SSYA10-001 and chromone-4c causes RecA1 and RecA2 to separate and generate a larger volume in the Nsp13 catalytic site. The presence of bananin does not change quantitatively the reference volume.

Figure 5 (B and C) shows the relative  $\beta$ -factor values alongside the relative shear strain throughout the protein upon binding of different small molecules. The reference state corresponds to the Nsp13 protein, with both binding regions occupied at the same time



**Fig. 4. Interaction diagrams for NSP13-ligand complexes: Bananin, SSYA10-001, and chromone-4c.** In all cases, light green represents van der Waals interactions, dark green represents hydrogen bonding interactions, light violet shows  $\pi$ -alkyl contacts, dark violet shows  $\pi$ -sigma interactions, and orange shows cation- $\pi$  interactions. The interaction diagrams were generated using BIOVIA software (39).

**Table 3. Absolute binding free energies for Nsp13-ligand complexes: The average energy and uncertainty values were obtained from three independent replicas.** Details of the calculation are provided in Materials and Methods.

TI (kcal/mol)	SSYA10-001	Bananin	Chromone-4c
Replica 1	−7.73	−3.83	−8.13
Replica 2	−8.53	−3.33	−9.23
Replica 3	−7.03	−4.33	−5.83
Average	−7.83 ± 0.9	−3.83 ± 0.5	−7.73 ± 1.8

by ADP and ssRNA in the absence of a small molecule. All the results given in Fig. 5 (B and C) are expressed as relative values with respect to that reference state to simplify the data interpretation. The  $\beta$ -factor values throughout the Nsp13 enzyme were estimated from the root mean square fluctuation (RMSF) of the heavy atoms belonging to the backbone of the protein to quantify thermal fluctuations for each residue.

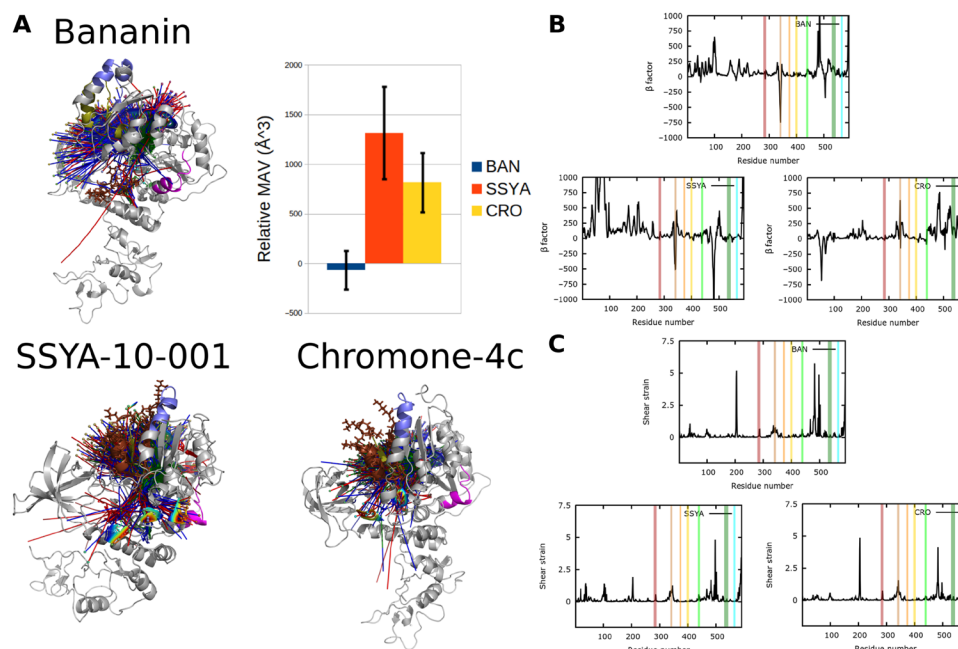
From Fig. 5B, it is apparent that the ligand bananin induces a small reduction in mobility at the Walker A motif located at the ATP/ADP binding site (region highlighted in red, motif I), while the opposite trend is observed in SSYA10-001 and chromone-4c. Note that all ligands bind the Nsp13 enzyme in a distal region with respect to the catalytic site; however, all of them trigger a mobility constraint at the active site of the protein. In addition, a larger reduction in mobility is observed for the conserved motif Ia (highlighted in dark orange), which binds the ssRNA, in bananin and SSYA10-001, where a mixed behavior is observed for chromone-4c. Bananin and chromone-4c induce an enhanced mobility in motif IV (highlighted in light green). Those ligands also show an important increase of the flexibility in a region around motif V (highlighted in dark green). With regard to the region from residues 150 to 261, the 1B domain, we observe an appreciable increase of global flexibility for all the ligands. Previous studies of the SARS-CoV Nsp13 helicase have shown that the residues 176 to 186 and 209 to 214, belonging to the 1B domain, constitute a probable nucleic acid binding region (10). We therefore expect the stiffening (and the enhanced flexibility) around those punctual sequences, displaying direct interactions with the ssRNA, to affect the normal Nsp13 function. However, time scale limitations preclude us from unraveling the full impact of the observations outlined above over the entire dynamics of the nucleic acid translocation processes.

Strain analysis helps disentangle the effects of functional and nonfunctional fluctuations and provides a more detailed view of the effects of the ligands. Figure 5C shows that all ligands induce strain in the Walker A motif (motif I, highlighted in red color) and motif Ia, consistent with the  $\beta$ -factor data. In addition, all ligands induce considerable strain in the region around residues 496 to 511 and the region around residue 200, which are a nucleic acid binding region on the 1B and RecA2 domains. This signal, which arises at a nucleic acid binding site on the 1B domain (residues 209 to 214), is reduced for SSYA10-001. SSYA10-001 also exhibits a high strain in the ZBD, but in this particular case, the signal is spread more broadly inside the ZBD, with a sharp spike around residue 100 at the border of the ZBD and Stalk domains. This pronounced spike is correlated with a conformational change of the ZBD, which can be seen in Fig. 5 (left), where the ZBD is located now much closer to the RecA1 domain.

The movement of the ZBD domain should be contrasted with its experimentally observed rigidity (10) and suggests a possible compromise of the helicase activity and its interaction with the RdRp complex (14). Note that studies of the previous SARS-CoV Nsp13 helicase have shown that both SARS-CoV Nsp12 and MERS Nsp12 are able to enhance the helicase and ATPase activity of the Nsp13 enzyme. Moreover, the interactions among the Nsp12 and Nsp13 proteins take place on the ZF3 motif belonging to the ZBD domain and the 1A domain. Therefore, it is plausible that the perturbations of the structure and dynamics of the ZBD also disturb the normal interaction patterns among these two Nsps and influence their function.

SASA also presents small changes upon ligand binding. The Nsp13+ssRNA+ADP system shows a value of  $188.56 \pm 31.43 \text{ \AA}^2$  for region 1,  $189.6 \pm 43.5 \text{ \AA}^2$  for region 2, and  $358.1 \pm 27.5 \text{ \AA}^2$  for region 3. Compared to that, SSYA10-001 exhibits noticeable changes only in regions 1 and 3, with values of  $171.87 \pm 28.6 \text{ \AA}^2$  and  $382.0 \pm 38.4 \text{ \AA}^2$ , respectively. Chromone-4c shows its biggest difference in region 3, with  $373.6 \pm 40.4 \text{ \AA}^2$ . For bananin, the reduction in the surface area occurs only in region 2, with the respective value of  $155.1 \pm 19.5 \text{ \AA}^2$ . These results point out that all these drugs may affect RNA binding to the RecA2 domain. To characterize the interactions between Nsp13 with ssRNA in the presence of the different ligands, we have calculated the relative contact map of the protein to the ssRNA and the relative linear interaction energy (LIE) (31). The reference state is again the Nsp13 protein with ADP and ssRNA but without any small molecules. The results are shown in Fig. 6. As shown in the upper left panel of Fig. 6, in general terms, the contacts among the RecA2 domain and the ssRNA are reduced with respect to the reference state. This is also evident from the relative LIE analysis depicted in the lower right panel of Fig. 6, where we observe a positive delta value for all ligands, with the largest changes occurring for bananin and chromone-4c. All ligands induce a loss of interactions between the ssRNA substrate and the RecA2 domain. In contrast, the RecA1 domain have a mixed behavior among the studied ligands, with motif Ia as one of the key interacting parts in RecA1, as mentioned above. LIE for this domain exhibits a stabilization in the ssRNA binding for bananin, while SSYA10-001 and chromone-4c present a small destabilization but with a huge variance. For the last two interacting domains, Stalk and 1B, the contact map shows a changing region around residue 180 for all ligands, with mixed gains and losses. These changes produce an overall positive LIE (in relative terms) for bananin and SSYA10-001, which implies a loss of interactions between the Nsp13 protein and the ssRNA substrate in the Stalk and 1B domains. The ligands with the strongest helicase inhibition, bananin and chromone-4c, exhibit a clear positive change in the ssRNA binding energy or, in other words, a pronounced loss of interactions between the Nsp13 protein and the oligonucleotide, which corresponds to the ssRNA substrate becoming detached from the Nsp13 enzyme; however, chromone-4c also has a bigger variance in the Stalk, 1B, and RecA1 domains, where only the RecA2 domain has fewer variable results. SSYA10-001 also shows a positive change, but the magnitude is smaller, and like chromone-4c, there is a high variance on those LIE results.

Figure 7 shows the results of our PCA and DCCM analyses of the four complexes. In all cases, the first two PCA modes account for more than 40% of the variance. The correlation maps focus on the residues in the RecA1 and RecA2 to capture the effects of drug binding in the helicase motifs and the ADP binding site. The left panel shows the first mode of the PCA. Bananin exhibits higher reduction in



**Fig. 5. Water trajectories in Nsp13 complexes (blue out and red in) and pore surface as rainbow contour maps. (A)** Cartoon representation of Nsp13 protein, with regions 1, 2, and 3 highlighted in magenta (427 to 437), slate (496 to 511), and olive (512 to 542), respectively. The ligand binding sites are shown with stick representation in brown color. Upper right: Relative MAV (in  $\text{\AA}^3$ ) of the water pocket in the active site of nsp13 due to the ligands' presence. **(B)** Relative  $\beta$ -factor values. **(C)** Shear strain analysis of the four different protein-ligand complexes. **(B)** and **(C)** highlight the regions that correspond to the conserved motifs in Nsp13 among SF1 helicase family: motif I (red), motif II (orange), motif III (yellow), motif IV (green), motif V (dark green), motif VI (cyan), and motif Ia (dark orange).

correlation values throughout the entire RecA1 and RecA2 domains, with a clear band of anticorrelation that expands from residues 440 to 460 that is absent in the reference. The first PCA mode exhibits a higher mobility than the reference system for the 1B and RecA2 domains but a reduced mobility in the ZBD domain. In the case of SSYA10-001, the correlation boundaries for RecA1 and RecA2 are sharper than for the two previous systems; the band around motif Ia is still present, but is barely distinguishable from its surroundings, as it is in the reference for bananin. The PCA mode on the left side panel of Fig. 7 shows a different behavior than that observed in the reference and the other ligands, largely as a result of the conformational change of the ZBD, where we now see a displacement of this domain, which gets closer to the RecA1, while 1B and RecA2 move in the opposite direction relative to RecA1. In the case of chromone-4c, the Nsp13 protein exhibits a reduced correlation band that spreads from the 400 to the 440 residues. In the three complexes, where this last band is still there, are the ones that exhibit helicase inhibition. Last, the first PCA mode of chromone-4c shows a mobility similar to that of the reference system, with the ZBD domain moving toward the main body of the protein.

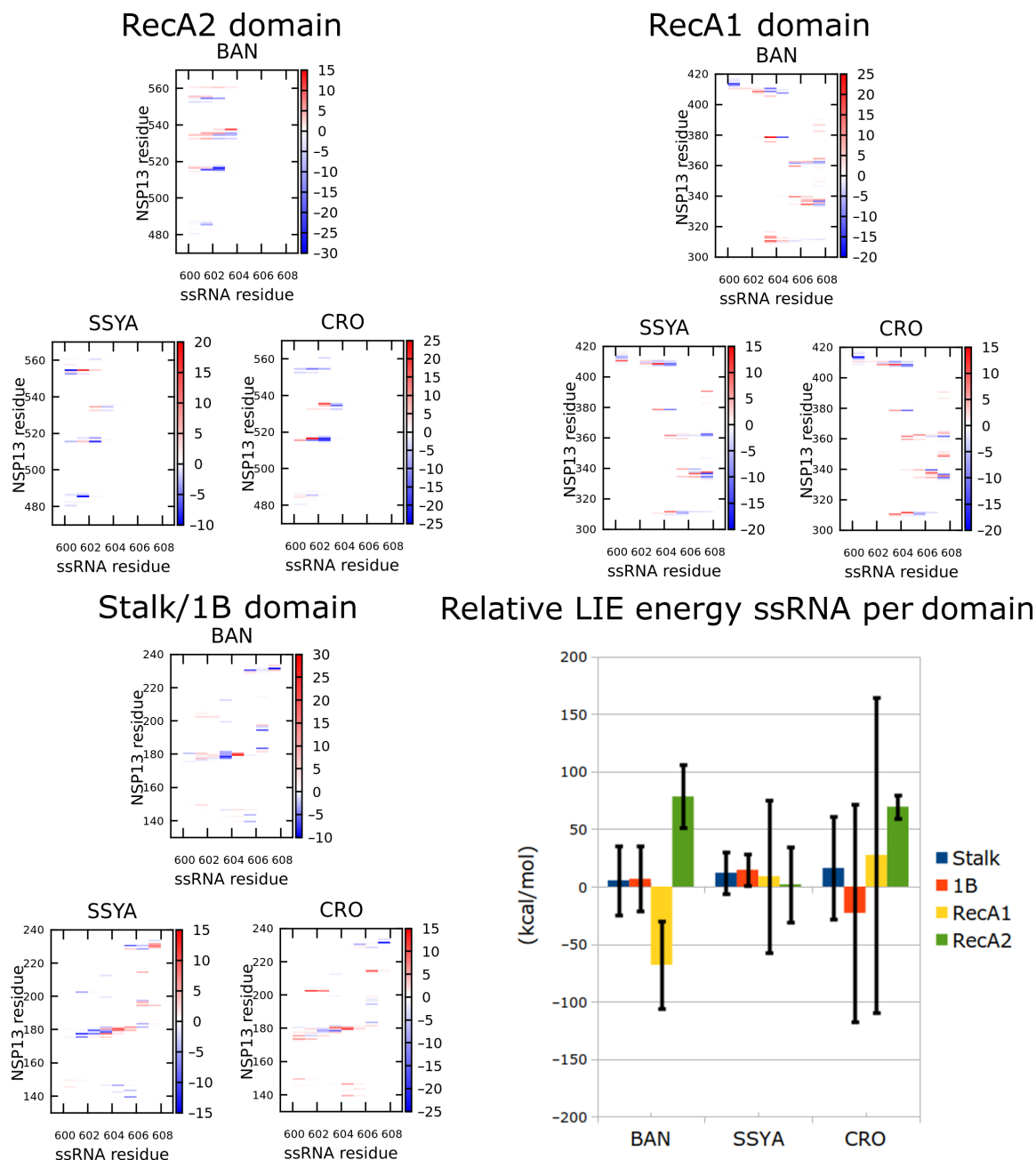
It is of interest to point out that past studies of the previous SARS-CoV helicase Nsp13 have shown that all five domains of the enzyme are directly or indirectly involved in the helicase activity and are finely coordinated with respect to each other in the assembly (10). Most of these observations (from  $\beta$ -factor, shear strain, PCA, and LIE analysis) suggest a high degree of communication throughout the entire Nsp13 helicase, where the binding of a small molecule at one punctual site can trigger some disruption of the normal Nsp13 structure and dynamics.

## DISCUSSION

Identifying compounds that exhibit a broad antiviral spectrum (34) by targeting highly conserved structures has proven to be effective against a wide array of viruses. In this sense, the SARS-CoV-2 Nsp13 helicase protein emerges as an attractive target given its sequence conservation and importance across all CoV species (15, 16).

The first part of this study focused on deciphering the structural coordination that exists between the Nsp13 domains and the behavior of the enzyme's complex with its natural ligands. The results from PCA and the DCCM matrix reveal a structural assembly that is characterized by a complex communication network among the five principal protein domains that is selectively activated by the presence of different natural ligands. As an example, the transition from Nsp13 apo form to Nsp13+ADP, the final product of the NTP nucleotide hydrolysis in the active site, gives rise to a high degree of interdomain communication that is absent in the apo state of the enzyme. A similar response was observed when both natural ligands, ADP and ssRNA, are incorporated into the assembly. A SASA analysis supports the existence of a communication mechanism between the ADP binding site and the ssRNA binding site in the RecA2 domain; the solvent surface exposure of regions 1, 2, and 3 is altered, and the increase of the active pocket volume (Fig. 3) serves to point out how the binding of ADP induces the opening of the RecA1 and RecA2 domains. Experimental studies on the previous SARS-CoV Nsp13 helicase have reported a similar behavior and response of the Nsp13 assembly to the natural ligand, ADP (10), suggesting that similar Nsp13 dynamics exist among CoVs. The opening of the RecA1 and RecA2 domains is also evident from the PCA, where the 1B and RecA2 domains (Fig. 2) separate from RecA1, either when



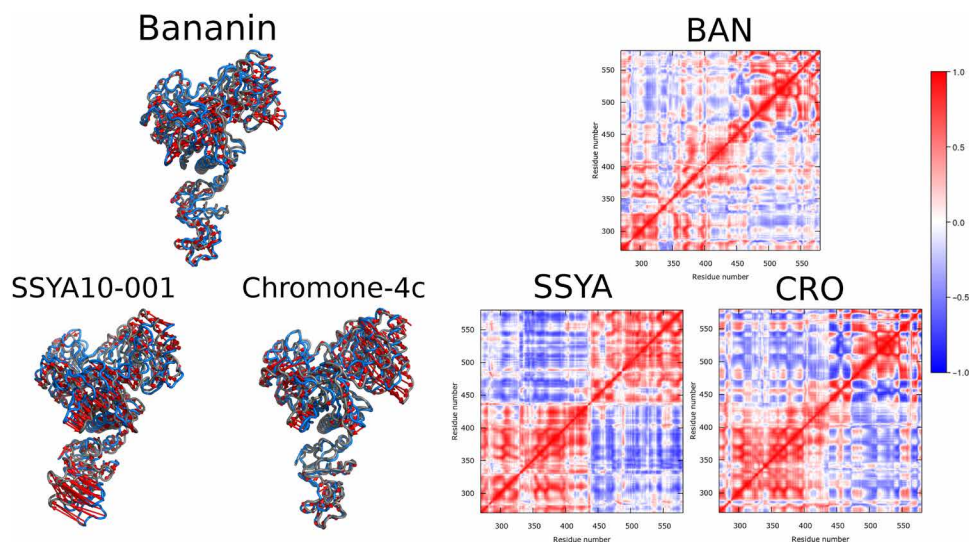


**Fig. 6. Characterization of the total relative contact maps and LIE of ssRNA and NSP13 in the presence of ligands.** The contact maps and LIE are divided by domains for each ligand. Bananin (BAN), SSYA10-001 (SSYA), and chromone-4c (CRO).

only ssRNA or when both ligands are attached to the protein. PCA also highlights the rigidity of the ZBD and how the presence of the natural ligands affects its mobility. The cross-correlation maps reveal pronounced anticorrelations with the other domains upon binding to the natural ligands. The allosteric communication in the Nsp13 protein is also evident in the strain analysis, which shows conformational changes of the helicase conserved motifs located in the ADP binding site and the RNA binding region, with motif I (Walker A) and motif Ia showing the largest effects.

Having characterized the structure and dynamics of the SARS-CoV-2 Nsp13 protein, the second part of this work was focused on

the characterization of three compounds that offer potential to hinder its activity. Specifically, we found that SSYA10-001, bananin, and chromone-4c all form thermodynamically stable complexes with the Nsp13+ADP+ssRNA assembly, with SSYA10-001 and chromone-4c exhibiting the highest affinity. Both ligands bind to Nsp13 at the RecA2 domain, in regions 2 and 3, underscoring these as important regions for the helicase's activity (10). In addition, the numerical value of the binding free energies is in the same order of magnitude than the experimental values, giving confidence on the model and results obtained. On the other hand, the computed binding free energy for bananin underestimates its inhibition power, where the



**Fig. 7. PCA and DCCM in the presence of ligands.** (Left) Movement representation of the first mode of the PCA of the  $\alpha$  carbons of Nsp13+ADP+ssRNA in the presence of ligands. (Right) Dynamical cross-correlation network for  $\alpha$  carbons in the presence of bananin (BAN), SSYA10-001 (SSYA), and chromone-4c (CRO).

causes of that could be related to the limitations of the selected force field to capture the complex interaction in the adamantane motif present in bananin. On the other hand, it could be possible that bananin exerts its antiviral activity by inhibition of other molecular target apart from Nsp13. Shear strain and  $\beta$ -factor analyses revealed that the largest perturbations of the assembly take place in motif I (Walker A) and motif Ia; the reduced flexibility in motif Ia may be responsible for the reduced Nsp13 helicase activity, as reported for these ligands and as predicted for SARS-CoV-2 by our models. In addition, the largest strain responses occur in the three binding regions for ssRNA (190 to 210), (330 to 350), and (490 to 530). This result, in conjunction with the contact maps and LIE analysis (which shows a reduction in the interaction between the ssRNA oligonucleotide and the Nsp13 protein), could explain the helicase inhibition mechanism. SSYA10-001 shows a different behavior than that of the other molecules: This small drug leads to a large strain in the ZBD-Stalk region, which, in turn, induces a conformational change of the protein. Given the importance of ZBD in the helicase activity, as well as its leading role on the interaction among Nsp13 and the RdRp, it is reasonable to expect these perturbations over the enzyme structure and dynamics to impair its function. Following the important role of NSP13 in the RTC complex, additional atomistic studies on the protein-protein interactions are now being pursued. The possible mechanism of disruption of the NSP13-RdRp complex by small molecules will be the focus of that work.

## MATERIALS AND METHODS

### Model of Nsp13 and natural ligands

The model of the Nsp13 helicase protein of SARS-CoV-2 adopted here is the SwissModel generated with the SARS-CoV helicase (PDB: 6JYT). To determine the binding models for ADP and ssRNA, we used the Dali server (35) to align the protein to PDB 2XZO and determine the initial configuration for those natural ligands. The generated PDB model of Nsp13 with ADP and ssRNA is included in the Supplementary Materials (data S1).

### Molecular dynamics

The parameters for each of the complexes considered here were determined using the ff14SB for the protein, RNA.OL3 for the oligonucleotide, and GAFF for the small molecules. The charges for the ligands were determined using the AM1-BCC method. An octahedral box with  $\sim 40,000$  TIP3P water molecules and  $\sim 23$   $\text{Cl}^-$  ions was used. Energy minimization included 3000 steps, which involved 1500 steepest descent steps and constraints to heavy atoms, followed by a second minimization of 30,000 steps involving 15,000 steepest descent steps. Equilibration was performed through a gradual temperature increase from 0 to 300 K over 400 ps using a Langevin thermostat with a temperature coupling constant of 1.0 ps in a constant volume ensemble. Density equilibration and production runs were conducted in a constant pressure ensemble (NPT). All simulations were performed using periodic boundary conditions and a 2-fs time step. Long-range electrostatic interactions were modeled using the Particle Mesh Ewald method with a nonbonded cutoff of 12 Å. The SHAKE algorithm was used to implement rigid constraints to covalent bonds involving hydrogen atoms.

### Molecular docking

AutoDock4.2 was used for the molecular docking between the target protein and ligands using the Lamarckian genetic algorithm and pseudo-Solis and Wets local search method (33). The initial configuration of NSP13+ADP+ssRNA was taken after an NPT relaxation for 600 ns. A total of 200 docking runs were performed for each replica, and each run was set to terminate after 25,000,000 energy calculations. The best result of each ligand was selected for further analysis in MD simulations.

### TI calculations

The absolute binding free energy for the nonnatural ligands is defined as  $\Delta G_{\text{Absolute}} = \Delta G^{\text{L}} - \Delta G^{\text{RL}}$ , where  $\Delta G^{\text{RL}}$  is the free energy change of the ligand annihilation in the Nsp13+ADP+ssRNA+drug assembly and  $\Delta G^{\text{L}}$  is the free energy change of ligand annihilation in water. To calculate these free energy changes, we used TI implemented in PMEMD (Particle Mesh Ewald Molecular Dynamics) for

Amber 18. A one-step annihilation protocol with soft core potentials was implemented. We used multiple runs starting from the equilibrated ligand position from previous 350-ns MD simulations. In this way, three independent replicas for each ligand, as well as three replicas for solvation in pure water, were considered. Eleven equally spaced windows were selected with 20 ns of simulation time per window. To keep the ligand from wandering in TI calculations, we used a soft positional restraint of 1 kcal/mol Å<sup>2</sup>, whose contribution to the total free energy calculations can be obtained analytically using the following equation (36)

$$\Delta G_{\text{restraint}} = -kT \ln \left( \frac{V^0}{(2\pi kT/K_{\text{harm}})^{3/2}} \right)$$

where  $V^0$  is a reference volume and  $K_{\text{harm}}$  is the harmonic constant applied to the positional restraint.

### SASA and LIE calculations

To calculate SASA for determined regions of the NSP13 protein, we used the LCPO method implemented in CPPTRAJ (31). For each of the studied regions, the calculations were averaged over the last 100 ns of simulation using conformations taken every 20 ps and using the backbone atoms (N, CA, C, and O) of each residue plus the amide hydrogen (H).

LIE was estimated using CPPTRAJ (31), where the nonbonded energy (van der Waals and electrostatic interactions) was calculated between the oligonucleotide and the protein divided by domains. The distance cutoff selected was 12 Å, and minimum image distance is applied. The results are averaged over 10,000 frames from the last 200 ns for each system in each replica.

### Water analysis

To analyze the water flux and active site changes in the NSP13 protein of SARS-CoV-2 for each complex, we used AQUA-DUCT 1.0.5. We obtained MAV, water paths, inlet clusters, and their surface. For calculations, the region of interest was defined as a 5-Å sphere around the center of geometry of the NTP active site residues (K288, S289, D374, E375, Q404, and R567). The time window used in all calculations was 100 ns, sampling every 10 ps. The results reported were block-averaged for the total simulation time used. Images were created with open-source PyMOL (37).

### Contact maps and RMSF calculations

Contacts maps by residue and RMSF calculations were generated using the native contacts and RMSF functions in CPPTRAJ (31). The contacts between the ssRNA and the Nsp13 protein were defined relative to the averaged PDB of Nsp13+ADP+ssRNA with a cutoff distance of 4 Å. The maps were averaged over 2500 snapshots taken every 100 ps for each ligand. The relative RMSF calculations were calculated with respect to a 200-ns averaged structure using the mass-weighted average over the CA, N, and C atoms and reported by residue. A total of 5000 snapshots were used, taken every 20 ps.

### Strain analysis

The effects of functional and nonfunctional fluctuations after ligand binding can be distinguished with the aid of the strain formalism from continuum theory. This type of analysis has been proven useful in proteins before (32). For this study, a 10-Å radius around each Cα atom for our analysis was considered. In the first part of this work,

a 300-ns simulation of Nsp13 was used as reference for strain calculations; a 300-ns simulation of Nsp13+ADP+ssRNA without ligand was used as reference for strain calculations in the second part of the article.

### PCA and DCCM

PCA and DCCM calculations were performed using the Bio3D package for R (29) and CPPTRAJ (31), respectively. PCA is a dimensionality reduction technique that is effective at identifying correlated motions in atomic simulations of proteins from experimental structures or MD trajectories. The first several PCs are often considered “essential dynamics,” and the rest are neglected without important loss of information (29).

DCCM represents all atom-wise cross-correlations whose elements,  $D_{ij}$ , may be displayed in a graphical representation

$$D_{ij} = \frac{\langle (r_i - \langle r_i \rangle) (r_j - \langle r_j \rangle) \rangle}{\sqrt{(\langle r_i^2 \rangle - \langle r_i \rangle^2) (\langle r_j^2 \rangle - \langle r_j \rangle^2)}}$$

where  $r_1, \dots, r_{3N}$  are the atomic coordinates of the protein, averaged over all the sampled snapshots from the simulation. If  $D_{ij} = 1$ , the fluctuations of atoms  $i$  and  $j$  are completely correlated (same period and same phase); if  $D_{ij} = -1$ , the fluctuations of atoms  $i$  and  $j$  are completely anticorrelated (same period and opposite phase); and if  $D_{ij} = 0$ , the fluctuations of  $i$  and  $j$  are not correlated. Off-diagonal positive and negative correlations may indicate potentially interesting correlations between domains of noncontiguous residues (29). For PCA and DCCM, 5000 equally spaced snapshots taken from a 600-ns simulation of each complex were used, and the analysis was performed on the Cα atoms of the protein.

### SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <https://science.org/doi/10.1126/sciadv.abj4526>

[View/request a protocol for this paper from Bio-protocol.](#)

### REFERENCES AND NOTES

1. R. J. de Groot, S. C. Baker, R. Baric, L. Enjuanes, A. E. Gorbalenya, K. V. Holmes, S. Perlman, L. Poon, P. J. M. Rottier, P. J. Talbot, Virus taxonomy, in *Family Coronaviridae*, A. M. Q. King, M. J. Adams, E. B. Carstens, E. J. Lefkowitz, Eds. (Elsevier, 2012), pp. 806–828.
2. 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).
3. 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).
4. 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).
5. 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).
6. A. Saberi, A. A. Gulyaeva, J. L. Brubacher, P. A. Newmark, A. E. Gorbalenya, A planarian nidovirus expands the limits of RNA genome size. *PLoS Pathog.* **14**, e1007314 (2018).
7. L. C. Van Dinten, H. van Tol, A. E. Gorbalenya, E. J. Snijder, The predicted metal-binding region of the arterivirus helicase protein is involved in subgenomic mRNA synthesis, genome replication, and virion biogenesis. *J. Virol.* **74**, 5213–5223 (2000).

8. R. Zhang, Y. Li, T. J. Cowley, A. D. Steinbrenner, J. M. Phillips, B. L. Yount, R. S. Baric, S. R. Weiss, The nsp1, nsp13, and M proteins contribute to the hepatotropism of murine coronavirus JHM.WU. *J. Virol.* **89**, 3598–3609 (2015).
9. K. C. Lehmann, E. J. Snijder, C. C. Posthuma, A. E. Gorbalenya, What we know but do not understand about nidovirus helicases. *Virus Res.* **202**, 12–32 (2015).
10. Z. Jia, L. Yan, Z. Ren, L. Wu, J. Wang, J. Guo, L. Zheng, Z. Ming, L. Zhang, Z. Lou, Z. Rao, Delicate structural coordination of the Severe Acute Respiratory Syndrome coronavirus Nsp13 upon ATP hydrolysis. *Nucleic Acids Res.* **47**, 6538–6550 (2019).
11. K. A. Ivanov, V. Thiel, J. C. Dobbe, Y. van der Meer, E. J. Snijder, J. Ziebuhr, Multiple enzymatic activities associated with severe acute respiratory syndrome coronavirus helicase. *J. Virol.* **78**, 5619–5632 (2004).
12. J. A. Tanner, R. M. Watt, Y. B. Chai, L. Y. Lu, M. C. Lin, J. S. M. Peiris, L. L. M. Poon, H. F. Kung, J. D. Huang, The severe acute respiratory syndrome (SARS) coronavirus NTPase/helicase belongs to a distinct class of 5' to 3' viral helicases. *J. Biol. Chem.* **278**, 39578–39582 (2003).
13. J. M. Caruthers, D. B. McKay, Helicase structure and mechanism. *Curr. Opin. Struct. Biol.* **12**, 123–133 (2002).
14. C. James, B. Malone, E. Llewellyn, M. Grasso, P. M. M. Shelton, P. D. B. Olinares, K. Maruthi, E. T. Eng, H. Vatandaslar, B. T. Chait, T. M. Kapoor, S. A. Darst, E. A. Campbell, Structural basis for helicase-polymerase coupling in the SARS-CoV-2 replication-transcription complex. *Cell* **182**, 1560–1573.e13 (2020).
15. A. O. Adedeji, B. Marchand, A. J. T. Velthuis, E. J. Snijder, S. Weiss, R. L. Eoff, K. Singh, S. G. Sarafianos, Mechanism of nucleic acid unwinding by SARS-CoV helicase. *PLOS ONE* **7**, e36521 (2012).
16. K. J. Jang, S. Jeong, D. Y. Kang, N. Sp, Y. M. Yang, D. E. Kim, A high ATP concentration enhances the cooperative translocation of the SARS coronavirus helicase nsP13 in the unwinding of duplex RNA. *Sci. Rep.* **10**, 4481 (2020).
17. F. K. Yoshimoto, The proteins of severe acute respiratory syndrome coronavirus-2 (SARS CoV-2 or n-COV19), the cause of COVID-19. *Protein J.* **39**, 198–216 (2020).
18. N. Katsiki, M. Banach, D. Mikhailidis, Lipid-lowering therapy and renin-angiotensin-aldosterone system inhibitors in the era of the COVID-19 pandemic. *Arch. Med. Sci.* **16**, 485–489 (2020).
19. A. O. Adedeji, K. Singh, A. Kassim, C. M. Coleman, R. Elliott, S. R. Weiss, M. B. Frieman, S. G. Sarafianos, Evaluation of SSYA10-001 as a replication inhibitor of severe acute respiratory syndrome, mouse hepatitis, and Middle East respiratory syndrome coronaviruses. *Antimicrob. Agents Chemother.* **58**, 4894–4898 (2014).
20. I. Briguglio, S. Piras, P. Corona, A. Carta, Inhibition of RNA helicases of ssRNA<sup>+</sup> virus belonging to Flaviviridae, Coronaviridae and Picornaviridae families. *Int. J. Med. Chem.* **2011**, 213135 (2011).
21. J. A. Tanner, B. J. Zheng, J. Zhou, R. M. Watt, J. Q. Jiang, K. L. Wong, Y. P. Lin, L. Y. Lu, M. L. He, H. F. Kung, A. J. Kesel, J. D. Huang, The adamantane-derived bananins are potent inhibitors of the helicase activities and replication of SARS coronavirus. *Chem. Biol.* **12**, 303–311 (2005).
22. C. Lee, J. M. Lee, N.-R. Lee, D.-E. Kim, Y.-J. Jeong, Y. Chong, Investigation of the pharmacophore space of severe acute respiratory syndrome coronavirus (SARS-CoV) NTPase/helicase by dihydroxochromone derivatives. *Bioorg. Med. Chem. Lett.* **19**, 4538–4541 (2009).
23. K. Rejdak, P. Grieb, Adamantanes might be protective from COVID-19 in patients with neurological diseases: Multiple sclerosis, parkinsonism and cognitive impairment. *Mult. Scler. Relat. Disord.* **42**, 102163 (2020).
24. M. S. Yu, J. Lee, J. M. Lee, Y. Kim, Y. W. Chin, J. G. Jee, Y. S. Keum, Y. J. Jeong, Identification of myricetin and scutellarein as novel chemical inhibitors of the SARS coronavirus helicase, nsP13. *Bioorg. Med. Chem. Lett.* **22**, 4049–4054 (2012).
25. S. Bienert, A. Waterhouse, T. A. P. de Beer, G. Tauriello, G. Studer, L. Bordoli, T. Schwede, The SWISS-MODEL Repository—New features and functionality. *Nucleic Acids Res.* **45**, D313–D319 (2017).
26. 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).
27. M. C. Hall, S. W. Matson, Helicase motifs: The engine that powers DNA unwinding. *Mol. Microbiol.* **34**, 867–877 (1999).
28. M. E. Fairman-Williams, U.-P. Guenther, E. Jankowsky, SF1 and SF2 helicases: Family matters. *Curr. Opin. Struct. Biol.* **20**, 313–324 (2010).
29. B. J. Grant, A. P. C. Rodrigues, K. M. ElSawy, J. A. McCammon, L. S. D. Cavas, Bio3D: An R package for the comparative analysis of protein structures. *Bioinformatics* **22**, 2695–2696 (2006).
30. T. Magdziarz, K. Mitusińska, M. Bzówka, A. Raczyska, A. Stańczak, M. Banas, W. Bagrowska, A. Góra, AQUA-DUCT 1.0: Structural and functional analysis of macromolecules from an intramolecular voids perspective. *Bioinformatics* **36**, 2599–2601 (2020).
31. 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).
32. M. R. Mitchell, T. Tlustý, 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).
33. G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, A. J. Olson, Autodock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–2791 (2009).
34. A. Adalja, T. Inglesby, Broad-spectrum antiviral agents: A crucial pandemic tool. *Expert Rev. Anti Infect. Ther.* **17**, 467–470 (2019).
35. L. Holm, DALI and the persistence of protein shape. *Protein Sci.* **29**, 128–140 (2020).
36. S. Boresch, F. Tettinger, M. Leitgeb, M. Karplus, Absolute binding free energies: A quantitative approach for their calculation. *J. Phys. Chem. B.* **107**, 9535–9551 (2003).
37. W. L. DeLano, Pymol: An open-source molecular graphics tool. *CCP4 Newsl. Protein Crystallogr.* **40**, 82–92 (2002).
38. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
39. Dassault Systemes, BIOVIA Discovery Studio Visualizer, 849 v20.1.0.19295 (Dassault Systemes, 2020).

**Acknowledgments:** We are grateful to the Research Computing Center of the University of Chicago for computational resources. **Funding:** This work was supported by NSF grants EFR1 CEE 1830969, BIO/MCB 1818328, and Multidisciplinary University Research Initiatives (MURI) Program grant W911NF-15-1-0568. The simulations reported here were carried out on the GPU cluster supported by the NSF through grant DMR-1828629. **Author contributions:** Conceptualization: G.R.P.-L., C.A.M., and J.d.P. Methodology: G.R.P.-L., C.A.M., W.A., and F.B. Investigation: G.R.P.-L., C.A.M., W.A., and F.B. Supervision: J.d.P. Writing—original draft: G.R.P.-L. and C.A.M. Writing—review and editing: G.R.P.-L., C.A.M., W.A., F.B., and J.d.P. **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 and/or the Supplementary Materials.

Submitted 13 May 2021  
 Accepted 17 November 2021  
 Published 7 January 2022  
 10.1126/sciadv.abj4526

## Toward wide-spectrum antivirals against coronaviruses: Molecular characterization of SARS-CoV-2 NSP13 helicase inhibitors

Gustavo R. Perez-LemusCintia A. MenéndezWalter AlvaradoFabian ByléhnJuan J. de Pablo

*Sci. Adv.*, 8 (1), eabj4526. • DOI: 10.1126/sciadv.abj4526

### View the article online

<https://www.science.org/doi/10.1126/sciadv.abj4526>

### Permissions

<https://www.science.org/help/reprints-and-permissions>

Use of think article is subject to the [Terms of service](#)

---

*Science Advances* (ISSN ) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS. Copyright © 2022 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).