Proposition of a new allosteric binding site for potential SARS-CoV-2 3CL protease inhibitors by utilizing molecular dynamics simulations and ensemble docking

Jurica Novak
Laboratory of Computational Modeling of Drugs, Higher Medical and Biological School, South Ural State University, Chelyabinsk, 454008, Russia, Chaikovskogo 20A https://orcid.org/0000-0001-5744-6677

Hrvoje Rimac (hrvoje.rimac@pharma.unizg.hr)
University of Zagreb Faculty of Pharmacy and Biochemistry https://orcid.org/0000-0001-7232-6489

Shivananda Kandagalla (kandagallas@susu.ru)
Laboratory of Computational Modeling of Drugs, Higher Medical and Biological School, South Ural State University, Chelyabinsk, 454008, Russia, Chaikovskogo 20A

Prateek Pathak
Laboratory of Computational Modeling of Drugs, Higher Medical and Biological School, South Ural State University, Chelyabinsk, 454008, Russia, Chaikovskogo 20A https://orcid.org/0000-0002-6160-9755

Maria Grishina
Laboratory of Computational Modeling of Drugs, Higher Medical and Biological School, South Ural State University, Chelyabinsk, 454008, Russia, Chaikovskogo 20A https://orcid.org/0000-0002-2573-4831

Vladimir Potemkin
Laboratory of Computational Modeling of Drugs, Higher Medical and Biological School, South Ural State University, Chelyabinsk, 454008, Russia, Chaikovskogo 20A https://orcid.org/0000-0002-5244-8718

Keywords: 3CL protease, COVID-19, drug repurposing, molecular dynamics, SARS-CoV-2, virtual screening

Posted Date: June 8th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-34002/v1

Version of Record: A version of this preprint was published at Journal of Biomolecular Structure and Dynamics on May 21st, 2021. See the published version at https://doi.org/10.1080/07391102.2021.1927845.

Proposition of a new allosteric binding site for potential SARS-CoV-2 3CL protease inhibitors by utilizing molecular dynamics simulations and ensemble docking

Jurica Novak[1], Hrvoje Rimac[1,2,*], Shivananda Kandagalla[1,*], Prateek Pathak[1], Vladislav Naumovich[1], Maria Grishina[1], Vladimir Potemkin[1]

1 Laboratory of Computational Modeling of Drugs, Higher Medical and Biological School, South Ural State University, Chelyabinsk, 454008, Russia, Chaikovskogo 20A

2 Department of Medicinal Chemistry, University of Zagreb, Faculty of Pharmacy and Biochemistry, 10000 Zagreb, Croatia, Ante Kovacica 1

Corresponding authors:
E-mail: hrvoje.rimac@pharma.unizg.hr, kandagallas@susu.ru

ORCID numbers:

Jurica Novak:	0000-0001-5744-6677
Hrvoje Rimac:	0000-0001-7232-6489
Prateek Pathak:	0000-0002-6160-9755
Maria Grishina:	0000-0002-2573-4831
Vladimir Potemkin:	0000-0002-5244-8718

Abstract: The SARS-CoV-2 3CL protease shows a high similarity with 3CL proteases of other beta-coronaviruses, such as SARS and MERS. It is the main enzyme involved in generating various non-structural proteins that are important for viral replication and is one of the most important proteins responsible for SARS-CoV-2 virulence. In this study, we have conducted ensemble docking of molecules from the DrugBank database using both crystallographic structure of the SARS-CoV-2 3CLpro, as well as five conformations obtained after performing a cluster analysis of a 300 ns molecular dynamics simulation. This procedure elucidated the inappropriateness of the active site for non-covalent inhibitors, but it has also shown that there exists an additional, more favorable, allosteric binding site, which could be a better target for non-covalent inhibitors, as it could prevent dimerization and activation of SARS-CoV-2 3CLpro.

Keywords: 3CL protease, COVID-19, drug repurposing, molecular dynamics, SARS-CoV-2, virtual screening.

Declarations

Funding: The authors acknowledge the financial support by the Croatian Science Foundation (research project UIP-2017-05-5160), the Government of the Russian Federation (Act 211, contract 02.A03.21.0011) and by the Ministry of Science and Higher Education of the Russian Federation (Grant FENU-2020-0019). The authors also acknowledge University of Zagreb, University Computing Centre (SRCE) for granting computational time on the Isabella cluster.

Conflicts of interest/Competing interests: The authors declare no conflict of interest.

Availability of data and material: Data is available in Supporting material

Code availability: NA

Introduction

Coronaviruses are a class of single-stranded positive-sense RNA viruses with a large viral RNA genome [1]. The SARS-CoV-2 is classified as a beta-coronavirus and has a highly similar genomic organization as other beta-coronaviruses (< 80% nucleotide identity and 89.10% nucleotide similarity with SARS-CoV genes) [2, 3]. A recent genome annotation of SARS-CoV-2 identified 14 orfs (open reading frames, i.e. continuous stretches of codons that have the ability of being translated) encoding for 27 proteins. Typically, beta-coronaviruses produce a polypeptide (~800 kDa) upon transcription of the genome. This polypeptide is proteolytically cleaved by a papain-like protease (PLpro) and a 3-chymotrypsin-like protease (3CLpro). 3CLpro cleaves the polyprotein at 11 distinct sites to generate various non-structural proteins that are important for viral replication [4]. Thus, this main protease is required for the maturation of coronaviruses and is vital for the viral life cycle, making it an attractive target for anti-SARS-CoV-2 inhibitors.

Further sequence comparison of the 3CLpro protein with its closest homologs shows that SARS-CoV-2 3CLpro shares sequence identity of 99.02% with bat SARS-like coronaviruses. It also shows sequence identity with SARS-CoV (96.08%), MERS-CoV (87.00%), human-CoV (90.00%) and bovine-CoV (90.00%) homologs. All these reports mean that 3CLpro is a highly conserved enzyme and a good target for anti-viral drugs [2, 5, 6].

Recent availability of SARS-CoV-2 3CLpro crystal structure (PDB ID: 6LU7) confirms its high structural similarity with SARS-CoV (PDB ID: 1UJ1) [7]. The 3CLpro protomers of both SARS-CoV-2 and SARS-CoV contain three distinct domains, i.e. domain I, domain II, and domain III, which are together comprised of nine α-helices and 13 β-strands (Fig. 1). Similar to other corona proteases, domain I (residues 8–101) and domain II (residues 102–184) include an antiparallel β-sheet structure with 13 β-strands, which resemble trypsin-like serine proteases structures. Domain III (C-terminal domain, residues 201–306) consists of five α-helices and is connected to domain II by a long loop (residues 185–200). The main substrate binding site is formed by a cleft between domains I and II, and has a catalytic dyad composed of conserved residues His 41 and Cys 145. Domains I and II are collectively referred as the N-terminal domain. Amino acids 1–7 in the N-terminus form an N-finger which plays a significant role in dimerization and formation of the active 3CLpro (Sang P, Tian S, Meng Z, Yang L. Insight derived from molecular docking and molecular dynamics simulations into the binding interactions between HIV-1 protease inhibitors and SARS-CoV-2 3CLpro, unpublished data).

Through enzyme activity measurements and molecular dynamics simulations, Chen et al.[8] determined that SARS 3CLpro monomer is unable to establish normal enzymatic activity and that only one protomer in the homodimer is active. Additionally, Shi and Song [9] identified four regions associated with 3CLpro dimerization: residues 1-5 from N-terminus forming N-finger (1), the residue Asn 214 (2), the region around residues Glu 288 – Asp 289 – Glu 290 in close contact with the N-finger (3), and the C-terminus last helix region around residues Arg 298 – Gln 299 (4). Due to their high similarity, this behavior is also expected for SARS-CoV-2.

Drug repurposing (or drug repositioning) strategy includes various data-driven and experimental procedures for identification of new uses for approved or investigational drugs that are outside the scope of the original medicinal indication [10–12]. The emergence of the SARS-CoV-2 virus motivated both science and pharmaceutical communities to speed up drug discovery against SARS-CoV-2 virus applying drug repurposing approaches [13]. According to clinicaltrials.gov, on March 10, 2020, 75 clinical trials were underway across the world, many of them trying to repurpose currently available antiviral drugs. One of these clinical trials (registration number: ChiCTR2000029603) is currently testing the suitability of HIV-1 protease inhibitors,

namely ASC09/ritonavir and lopinavir/ritonavir cocktails [14]. In the initial epicenter of the epidemic, in Wuhan, China, another clinical trial (registration number: ChiCTR2000029541) is running, testing the appropriateness of darunavir/cobicistat and lopinavir/ritonavir combined with thymosin α1. Based on pre-clinical studies in SARS-CoV and MERS-CoV infections, two clinical trials (Identifiers: NCT04252664 and NCT04257656) [15] are testing potential of remdesivir drug as a potential antiviral therapy for COVID-2019.

Nukoolkarn el al. [16] performed molecular dynamics simulations for the SARS-CoV 3CLpro free enzyme and its complexes with lopinavir and ritonavir. Complex intermolecular interactions when the inhibitors are bound to the proteinase active site result with enzyme's flap closing. Li et al. (Therapeutic Drugs Targeting 2019-nCoV Main Protease by High-Throughput Screening, unpublished data) aside from identifying drugs with high binding capacity to SARS-CoV main protease (prulifloxacin, bictegravir, nelfinavir, and tegobuvir), identified three binding sites: binding site of a natural substrate, a site in the groove between the two monomers and a small pocket near the C terminus.

The recent availability of the SARS-CoV-2 3CLpro 3D structure provides deep insights into the viral life cycle and aids the screening of anti-COVID-19 drugs. In this concern, the present work focused on identification of potential candidate drugs against 3CLpro protease through virtual screening, as well as identification of additional druggable sites on the enzyme.

Materials and Methods

Molecular dynamics (MD) simulations

The SARS-CoV-2 3CLpro structure was obtained from RCSB (code 6LU7). After removing the ligand and water molecules and visual inspection of amino acid residues, the protein was prepared for MD simulations to obtain its different conformations. AMBER ff14SB force field was used and the protein was solvated in a truncated octahedral box of TIP3P water molecules spanning a 12 Å thick buffer. The 3CLpro was neutralized by Na_+ ions and submitted to geometry optimization in AMBER16 program, employing periodic boundary conditions in all directions [21]. For the first 1500 cycles the complex was restrained and only water molecules were optimized, after which another 2500 cycles of optimization followed, where both water molecules and complex were unrestrained. Optimized systems were gradually heated from 0 K to 300 K and equilibrated during 30 ps using NVT conditions, followed by productive and unconstrained MD simulations of 300 ns employing a time step of 2 fs at a constant pressure (1 atm) and temperature (300 K). The latter was held constant using Langevin thermostat with a collision frequency of 1 ps^{-1}. Bonds involving hydrogen atoms were constrained using the SHAKE algorithm [22], while the long-range electrostatic interactions were calculated employing the Particle Mesh Ewald method [23]. The non-bonded interactions were truncated at 11.0 Å. After the simulation, k-means cluster analysis was performed using CPPTRAJ [24]. Enzyme structures were clustered into five groups based on RMSD in backbone atoms of amino acids located within 5.5 Å of the Cys 145 residue (Thr 24, Thr 25, Thr 26, Leu 27, His 41, Met 49, Tyr 54, Phe 140, Leu 141, Asn 142, Gly 143, Ser 144, Cys 145, His 163, His 164, Met 165, Glu 166, Leu 167, Pro 168, His 172, Asp 187, Arg 188, Gln 189, Thr 190, Ala 191, and Gln 192), with maximal number of iterations set to 500, randomized initial set of points used and sieving set to 10. After clustering, frames closest to the centroids of each cluster were identified as five different conformations and were used as receptors in the second round of docking.

Molecular docking

First, a database of drugs was obtained from DrugBank (Release Version 5.1.5) [25]. After cross-checking the collected drug molecules for redundancies an sdf file containing 8750 potential ligands was procured. All structures were then converted into pdb format using Open Babel 3.0.0. program [26] and were prepared for docking using AutoDockTools 4 [27] program's script prepare_ligand4.py and saved in pdbqt format. Before docking, the crystallographic inhibitor molecule was removed, hydrogens were added where necessary, all Lys, Arg, His, and Cys sidechains were protonated, all Asp and Glu sidechains were deprotonated, and the amino and carboxy termini were charged. Such molecule was then minimized and saved in the pdbqt format. All dockings were performed using AutoDock Vina [28] locally on 6 personal computers with 8 Intel® CoreTM i7-6700K CPU @ 4.00 GHz, 32 GB RAM, and 64-bit Windows 10 Pro operating system. Docking of the crystallographic ligand (code name N3) was performed to assess the suitability of the docking procedure. First round of ligand docking was performed on the protein in the crystallographic conformation and docking was centered at the Cys 145 residue, with coordinates -11.4, 12.8, 70.1 and the size of the box was 20 × 30 × 20 Å. Number of runs was set to 100 and exhaustiveness to 20. All ligands with binding energy lower than -7.0 kcal/mol proceeded to the second round of docking. These ligands were then docked to five different conformations of the SARS-CoV-2 3CLpro, which were obtained using MD simulations (the method is described in the previous paragraph), using the same parameters as for the first round of docking. For these ligands, a weighted binding constant across all five receptor conformations was calculated using proportion of time the receptor was in each conformation (eq. 1):

$$K_w = \sum_{i=1}^{n} K_i \times p_i \quad (1)$$

K_w is the ligand's weighted binding constant, Ki is ligand's binding constant for a given protein conformation, pi is the proportion of time the protein is in that conformation, and n is the number of different conformations (in this case 5). For the best 10 ligands obtained in this way a final, third, round of docking was performed on all 5 protein conformations, but this time exhaustiveness was set to 100, in order to confirm if the best binding poses were truly found. Additionally, blind docking was also performed for those best 10 ligands, with exhaustiveness set to 100, to possible locate an alternative binding site outside the active pocket.

Results and Discussion

Molecular dynamics (MD) simulations

In Fig. 2a changes in RMSD of the entire backbone compared to the first frame are displayed. It is immediately noticeable that in the last 25 ns, a significant conformational change occurs, depicted in Fig. 3a. Highly flexible loop, connecting domain III and an anti-parallel cluster of five α-helices of domain II, is responsible for this conformational change. According to our simulation, the secondary structure of all three domains is conserved in both conformations, with slight changes in the unstructured loops, connecting secondary structure motifs. Clustering of the SARS-CoV-2 3CLpro trajectory based on RMSD of the active pocket (amino acid residues within 5.5 Å of Cys 145) is shown in Fig. 2b, while three representative conformations of the three most populated clusters are depicted on Fig. 3b. Comparing the RMSD's obtained in these two ways (Fig. 2a and 2b) and by visual inspection of generated conformations, it was concluded that the greatest conformational change occurs outside of the active site.

Even though the clustering was based on the amino acid residues in the active site (Table 1.), the main conformational differences still occurred in the domain III. Obviously, structural changes in the domain III are

reflected on the structure of the active site itself. The only noticeable conformational differences in the active site are found between residues Cys 44 – Leu 67 (a region of the domain I with 3 small α-helices) and Val 186 – Gln 192 (in the loop region) (Fig. 3b). The same is also true for clusters 4 and 5 (data not shown).

Table 1. Cluster occupancy of SARS-CoV-2 3CLpro through time (300 ns).

Cluster #	Cluster population (%)	Average distance from the cluster centroid	Cluster standard deviation
1	32.6	1.119	0.256
2	31.9	1.474	0.300
3	18.6	1.375	0.299
4	8.5	1.441	0.298
5	8.3	1.504	0.306

Docking studies

Since ligands in the DrugBank database are compounds which are approved or undergoing clinical trials, they are usually small molecules. The mean and the highest molecular weights of the tested compounds were 331.3 Da and 1268.9 Da, respectively. The mean binding energy of potential inhibitors to the crystal structure of the SARS-CoV-2 3CLpro according to the present study is -6.4 ± 1.3 kcal/mol (Fig. 4). The arbitrary cut-off binding energy was -7.0 kcal/mol and 3056 out of 8750 ligands (35%) passed the first screening and proceeded to the second round.

These ligands were then docked to five different protein conformations and their weighted binding constant was calculated according to eq. 1. Best 10 performing ligands additionally underwent a third round of docking with increased exhaustiveness in order to confirm the best binding pose. The results for the most favorable binding poses of the best 10 ligands to all five protein conformations showed that the binding constants of the same ligand to different protein conformations do not differ significantly (Table 2). This is due to fact that the shape of the active site does not change significantly throughout the MD simulation. Additionally, given that the active pocket itself is shallow and wide, many ligands can bind here in different conformations with similar affinities (Fig. 5). Therefore, the lack of a more strictly defined active site makes it not a good target for non-covalent inhibitors. On the other hand, since the 3CLpro role and its catalytic mechanism are known and well described [4, 16, 17], it is a good target for covalent inhibitors.

Table 2. Docking binging energies (in kcal/mol) of 10 hit molecules in the active site of SARS-CoV-2 3CLpro. Weighted binding constant (Kw) was calculated using eq. 1.

Ligand name	K_1	K_2	K_3	K_4	K_5	K_w	Drug status
phthalocyanine	-8.9	-8.6	-9.7	-9.2	-10.1	-9.1	investigational for the treatment of actinic keratosis, Bowen's disease, skin cancer, and stage I or stage II mycosis fungoides
fenebrutinib	-9.4	-8.9	-9.2	-8.2	-8.5	-9.0	investigational for the treatment of non-Hodgkin's lymphoma and chronic lymphocytic leukemia

R-428	-9.2	-8.7	-8.7	-9.0	-8.9	-8.9	investigational for the treatment of non-small cell lung cancer
rimegepant	-9.2	-8.4	-9.2	-8.5	-8.7	-8.8	approved for the treatment of acute migraine headache
DB01897[a]	-8.9	-8.9	-8.8	-8.9	-8.4	-8.8	experimental as a hematopoietic prostaglandin D synthase inhibitor
Zk-806450	-9.1	-8.5	-9.3	-8.7	-8.2	-8.8	experimental as a serine-type endopeptidase inhibitor
radotinib	-8.7	-9.0	-8.9	-8.7	-8.3	-8.8	investigational for the treatment of myelogenous, chronic, BCR-ABL positive leukemia
nilotinib	-8.7	-8.9	-8.9	-8.8	-8.4	-8.8	investigational for the treatment of chronic myelogenous leukemia (CML)
tegobuvir	-8.8	-9.1	-8.3	-8.8	-8.3	-8.8	investigational for the treatment of chronic hepatitis C
PF-5190457	-8.5	-9	-8.9	-8.4	-8.7	-8.7	investigational as a ghrelin receptor antagonist

[a] 2-(2f-benzothiazolyl)-5-styryl-3-(4f-phthalhydrazidyl)tetrazolium chloride

For 10 best inhibitors, a blind docking study was also performed in the search of a more favorable binding site; we investigated a possibility of an alternative, allosteric binding site, which could impede the function of the SARS-CoV-2 3CLpro. In the screening, the ligands have shown preferred binding in the groove between domains II and III, and in most cases with higher affinities than for the active site. The groove of the representative conformation of the cluster 1 (groove 1) is delimited by residues Arg 4, Lys 5, Ala 7, Ser 10, Gln 127, Lys 137, Gly 138, Ser 139, Glu 166, Gly 170, Trp 207, Ser 284, and Gln 306. (Table 3 and Fig. 6a). For all other clusters, due to conformational changes, this groove 1 is non-existent. However, another groove (groove 2) forms, where ligands bind on the side opposite to the Cys 145 residue, delimited by amino acid residues Met 6, Val 104, Gly 109, Gln 110, Gln 127, Asn 151, Tyr 154, Asn 203, Asp 245, His 246, Thr 292, Phe 294, and Arg 298 (Table 3 and Fig. 6b). By conducting the RMSD analysis on the entire backbone, it was determined that, in fact, there are two main conformations of the SARS-CoV-2 3CLpro, represented by clusters 1 and 2 (backbone RMSD is 8.761 Å), while the backbone RMSD between representative conformations of cluster 2 and clusters 3, 4, and 5 is much lower (2.264, 1.406, and 1.178 Å, respectively). We compared the conformations of clusters 1 and 2 to the crystallographic structures of the monomeric (PDB entry 6LU7) and its dimeric form (PDB entry 6Y2G). The analysis revealed that the conformation of cluster 2 is the most similar to the crystallographic conformations (with RMSD of 2.149 Å in case of 6LU7 and 1.588 Å in case of 6Y2G). The similarity between cluster 1 and crystallographic conformations is lower (with RMSD of 7.378 Å and 8.172 Å for 6LU7 and 6Y2G, respectively), while the difference between the monomeric and dimeric crystallographic forms is 0.721 Å.

In both cases, the groove is located in the vicinity of amino acids crucial for the dimerization process. As it was determined by Lim et al. [18], a single Arg298Ala mutation completely stops the dimerization, resulting in an inactive monomeric form of the enzyme. Since the role of the N-finger in correct dimerization is also well known [19, 20], binding of ligands between domains II and III (most notably, the N-finger and Arg 298) could allosterically inhibit the activity of SARS-CoV-2 3CLpro by preventing it to dimerize into the active form [7, 19].

Table 3. Docking binging energies (in kcal/mol) of 10 hit molecules in the groove between the domains II and III of SARS-CoV-2 3CLpro. Weighted binding constant (Kw) was calculated using eq. 1.

Ligand name	K_1	K_2	K_3	K_4	K_5	K_w
phthalocyanine	-11	-10.8	-11.4	-10	-11.2	-10.9
R-428	-10.3	-9.6	-10.2	-9.7	-10.7	-10
Zk-806450	-10.3	-9.2	-10.3	-9.8	-10.1	-9.9
nilotinib	-9.6	-9.6	-9.8	-10.2	-10.6	-9.8
fenebrutinib	-10.4	-8.8	-9.4	-9.3	-9.4	-9.5
DB01897[a]	-9.8	-8.5	-9.8	-9.0	-10.3	-9.3
radotinib	-9.4	-8.7	-9.6	-10	-10.5	-9.3
tegobuvir	-9.4	-8.9	-9.3	-9.8	-10.3	-9.3
PF-5190457	-8.9	-8.6	-8.8	-8.4	-10.4	-8.9
rimegepant	-8.4	-8.8	-9.6	-8.6	-9.1	-8.8

[a] 2-(2f-benzothiazolyl)-5-styryl-3-(4f-phthalhydrazidyl)tetrazolium chloride

This groove represents a possible alternative inhibition target to inhibition of the active site. Additionally, considering the shallowness of the active site and the lack of discrimination between ligands' binding energies when binding to it, this study suggests that binding of ligands inside the groove (as opposed to the active site) might even be more favorable for non-covalent inhibitors. Therefore, we propose considering the groove as a target binding site in future 3CLpro inhibition studies and in drug repurposing.

Conclusion

In this article ensemble virtual screening study of DrugBank library was conducted using both the crystallographic conformation of SARS-CoV-2 3CLpro, as well as its five representative conformations which were obtained after 300 ns MD simulation and *k*-means clustering. The study has shown that the 3CLpro active site is not the best target for non-covalent inhibitors due to its shallowness and wideness. However, we propose targeting the additional, allosteric binding site, located in the groove between domains II and III. This groove is partially made of amino acids in the N-finger region (residues 1–7) and crucial Arg 298 residue, which are involved in dimerization and activation of SARS-CoV-2 3CLpro. Therefore, binding of ligands to this binding site could inhibit protein dimerization, and consequently, SARS-CoV-2 3CLpro activation.

References

1. Chen Y, Liu Q, Guo D (2020) Emerging coronaviruses: Genome structure, replication, and pathogenesis. J Med Virol 92:418–423. https://doi.org/10.1002/jmv.25681
2. Zhu N, Zhang D, Wang W, et al (2020) A Novel Coronavirus from Patients with Pneumonia in China, 2019. N Engl J Med 382:727–733. https://doi.org/10.1056/nejmoa2001017
3. Wu F, Zhao S, Yu B, et al (2020) A new coronavirus associated with human respiratory disease in China. Nature 579:. https://doi.org/10.1038/s41586-020-2008-3
4. Anand K, Ziebuhr J, Wadhwani P, et al (2003) Coronavirus Main Proteinase (3CLpro) Structure: Basis

for Design of Anti-SARS Drugs. Science (80-) 300:1763–1767. https://doi.org/10.1126/science.1085658

5. Zhou P, Yang X-L, Wang X-G, et al (2020) A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 579:. https://doi.org/10.1038/s41586-020-2012-7
6. Xu X, Chen P, Wang J, et al (2020) Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission. Sci China Life Sci 63:457–460. https://doi.org/10.1007/s11427-020-1637-5
7. Yang H, Yang M, Ding Y, et al (2003) 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. https://doi.org/10.1073/pnas.1835675100
8. Chen H, Wei P, Huang C, et al (2006) Only one protomer is active in the dimer of SARS 3C-like proteinase. J Biol Chem 281:13894–13898. https://doi.org/10.1074/jbc.M510745200
9. Shi J, Song J (2006) The catalysis of the SARS 3C-like protease is under extensive regulation by its extra domain. FEBS J 273:1035–1045. https://doi.org/10.1111/j.1742-4658.2006.05130.x
10. Ashburn TT, Thor KB (2004) Drug repositioning: Identifying and developing new uses for existing drugs. Nat Rev Drug Discov 3:673–683. https://doi.org/10.1038/nrd1468
11. Pushpakom S, Iorio F, Eyers PA, et al (2018) Drug repurposing: Progress, challenges and recommendations. Nat Rev Drug Discov 18:41–58. https://doi.org/10.1038/nrd.2018.168
12. Vanhaelen Q (2019) Computational Methods for Drug Repurposing
13. Harrison C (2020) Coronavirus puts drug repurposing on the fast track. Nat Biotechnol February: https://doi.org/10.1038/d41587-020-00003-1
14. A Randomized, Open-Label, Multi-Centre Clinical Trial Evaluating and Comparing the Safety and Efficiency of ASC09/Ritonavir and Lopinavir/Ritonavir for Confirmed Cases of Novel Coronavirus Pneumonia (COVID-19). http://www.chictr.org.cn/showprojen.aspx?proj=49075
15. ClinicalTrials.gov. https://clinicaltrials.gov/ct2/home
16. Nukoolkarn V, Lee VS, Malaisree M, et al (2008) Molecular dynamic simulations analysis of ritronavir and lopinavir as SARS-CoV 3CLpro inhibitors. J Theor Biol 254:861–867. https://doi.org/10.1016/j.jtbi.2008.07.030
17. Li C, Teng X, Qi Y, et al (2016) Conformational Flexibility of a Short Loop near the Active Site of the SARS-3CLpro is Essential to Maintain Catalytic Activity. Sci Rep 6:1–9. https://doi.org/10.1038/srep20918
18. Lim L, Shi J, Mu Y, Song J (2014) Dynamically-driven enhancement of the catalytic machinery of the SARS 3C-like protease by the S284-T285-I286/A mutations on the extra domain. PLoS One 9:. https://doi.org/10.1371/journal.pone.0101941
19. Zhong N, Zhang S, Zou P, et al (2008) Without Its N-Finger, the Main Protease of Severe Acute Respiratory Syndrome Coronavirus Can Form a Novel Dimer through Its C-Terminal Domain. J Virol 82:4227–4234. https://doi.org/10.1128/jvi.02612-07
20. Wei P, Fan K, Chen H, et al (2006) The N-terminal octapeptide acts as a dimerization inhibitor of SARS coronavirus 3C-like proteinase. Biochem Biophys Res Commun 339:865–872. https://doi.org/10.1016/j.bbrc.2005.11.102

21. Case DA, Betz RM, Cerutti DS, et al (2016) Amber 2016. Univ California, San Fr
22. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 23:327–341. https://doi.org/10.1016/0021-9991(77)90098-5
23. Darden T, York D, Pedersen L (1993) Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. J Chem Phys 98:10089–10092. https://doi.org/10.1063/1.464397
24. Roe DR, Cheatham TE (2013) PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. J Chem Theory Comput 9:3084–3095. https://doi.org/10.1021/ct400341p
25. DrugBank. https://www.drugbank.ca/
26. O'Boyle NM, Banck M, James CA, et al (2011) Open Babel: An open chemical toolbox. J Cheminform 3:33. https://doi.org/10.1186/1758-2946-3-33
27. Morris GM, Huey R, Lindstrom W, et al (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J Comput Chem 30:2785–2791
28. Handoko SD, Ouyang X, Su CTT, et al (2012) QuickVina: Accelerating AutoDock Vina using gradient-based heuristics for global optimization. IEEE/ACM Trans Comput Biol Bioinforma 9:1266–1272. https://doi.org/10.1109/TCBB.2012.82

Figures

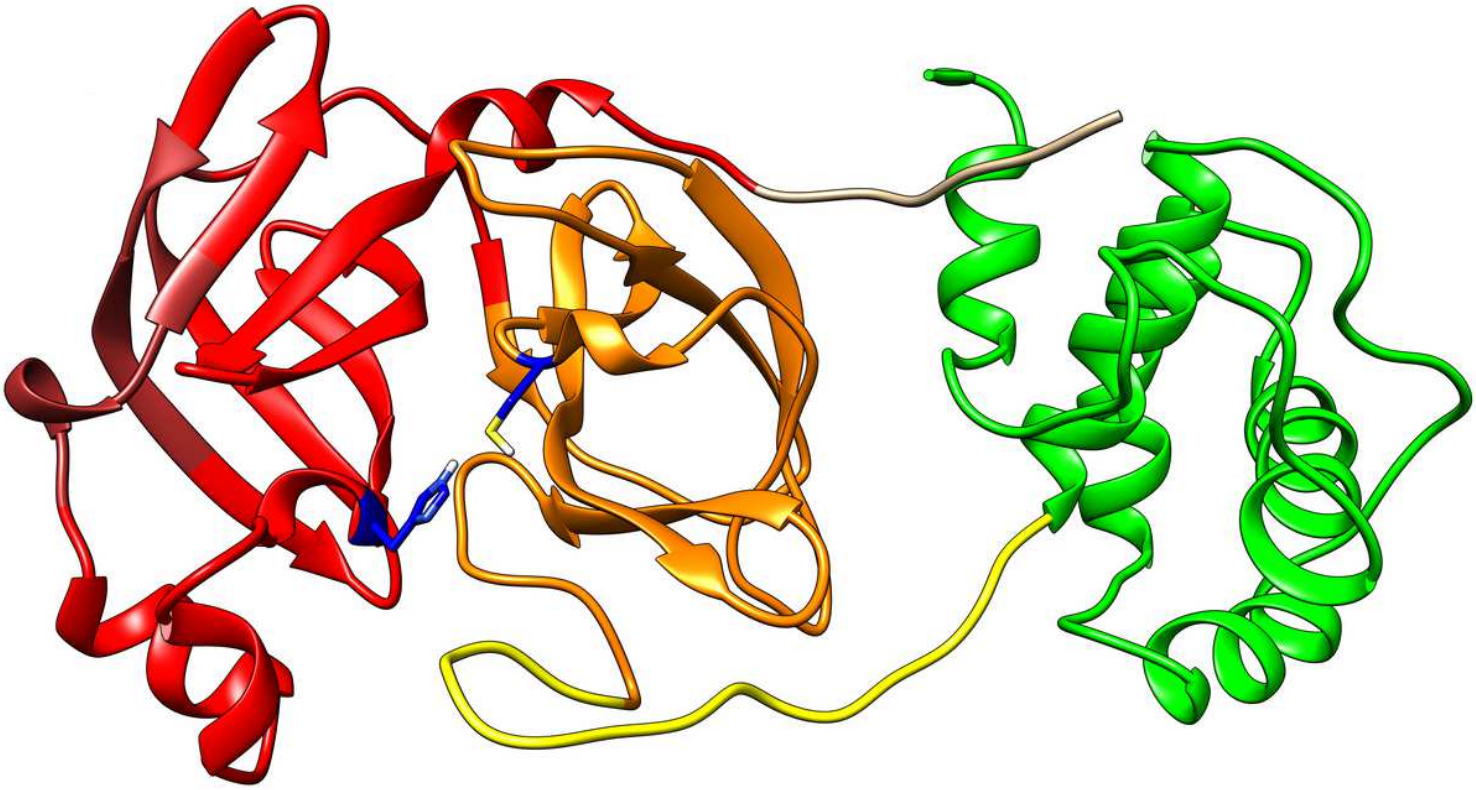

Figure 1

Structure of SARS-CoV-2 3CLpro (PDB ID: 6LU7). The N-finger (residues 1–7) is depicted in tan, domain I (residues 8–101) in red, domain II (residues 102–184) in orange, loop region (residues 185–200) in yellow, domain III (residues 201–306) in green, and conserved His 41 and Cys 145 in dark blue.

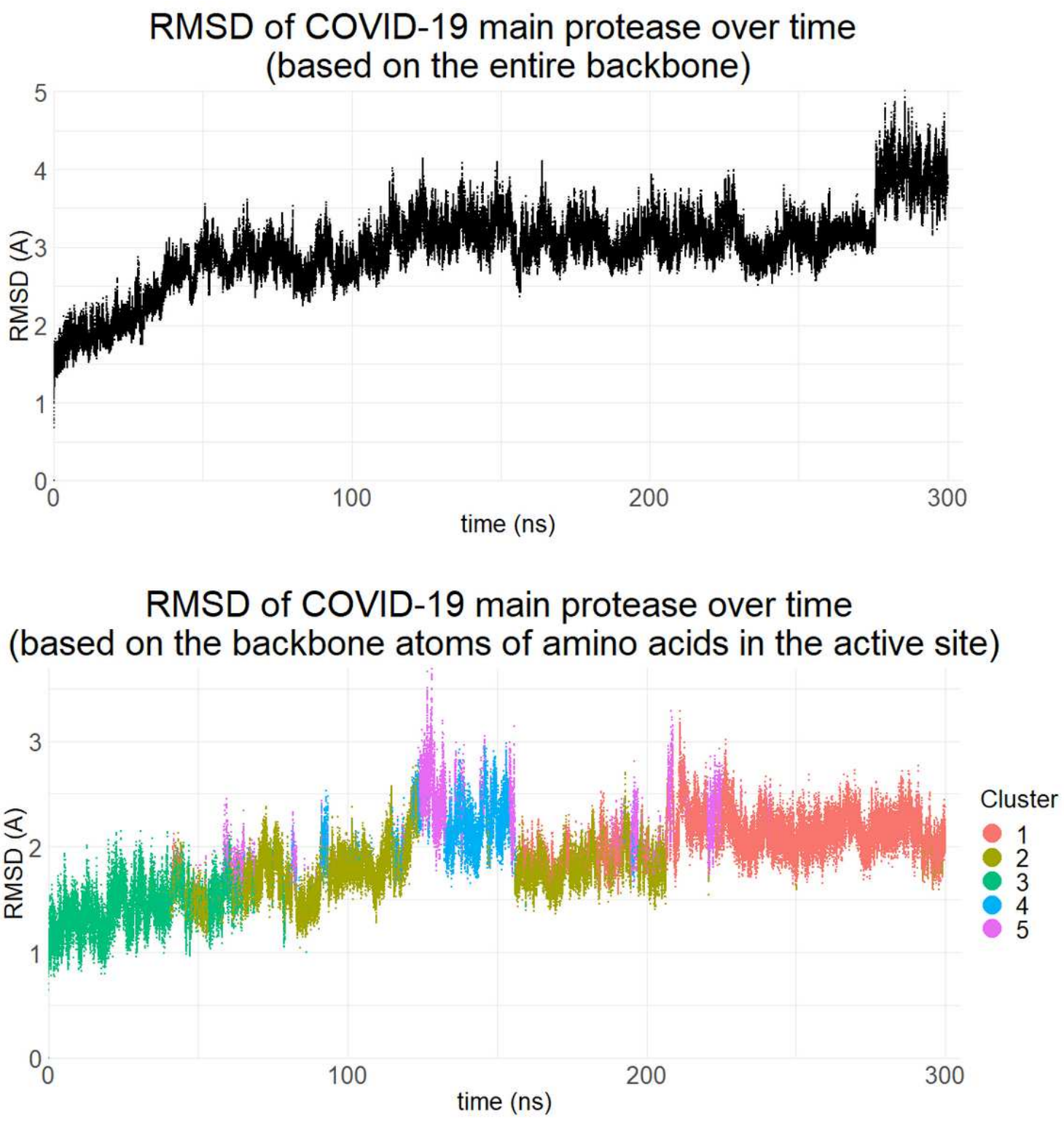

Figure 2

Root-Mean-Square Deviation (RMSD) of SARS-CoV-2 3CLpro (PDB entry 6LU7) over time, in relation to the first frame: a) based on the whole backbone, b) based on the backbone atoms of amino acids located within 5.5 Å of the Cys 145 residue.

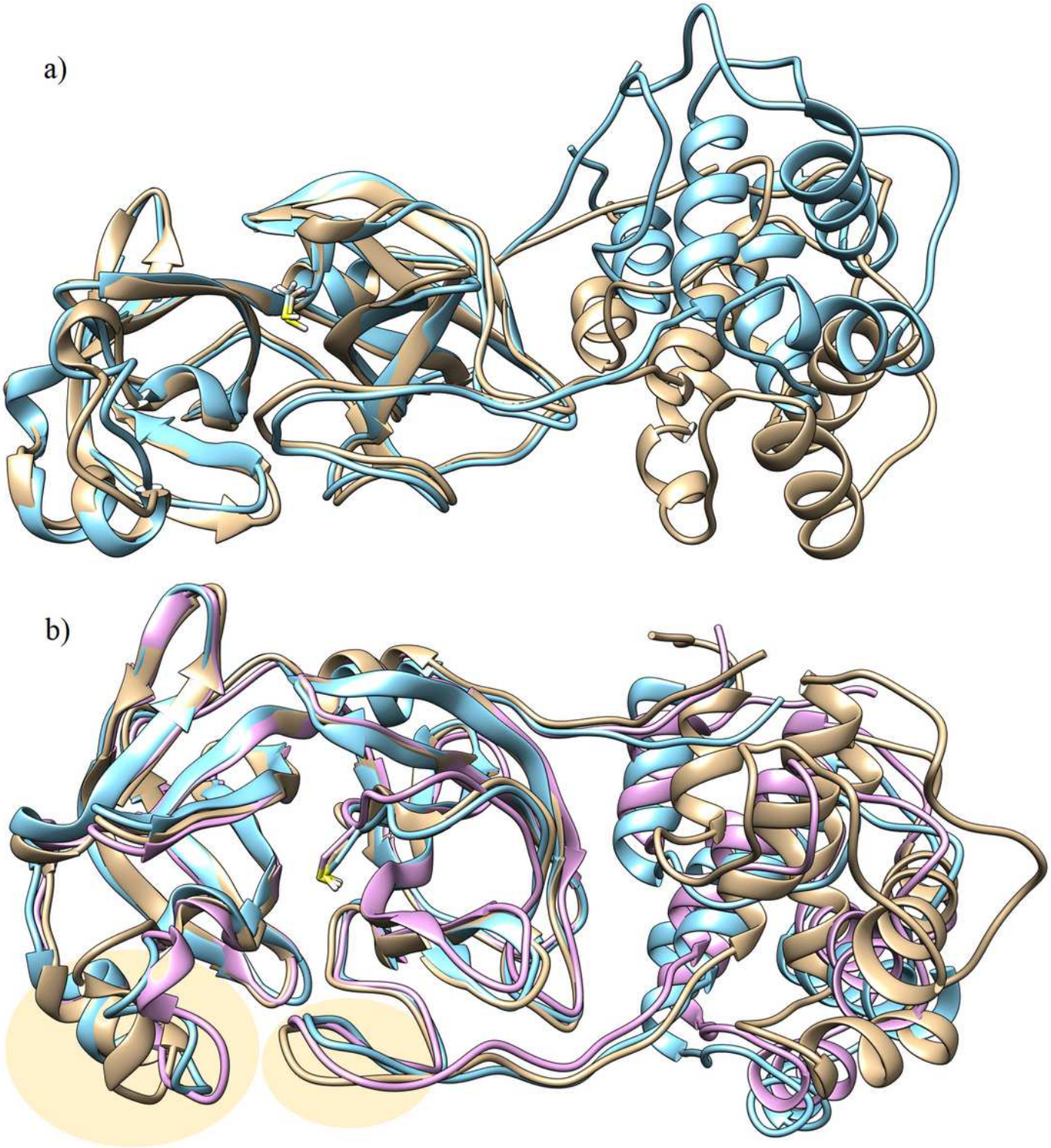

Figure 3

Overlay of SARS-CoV-2 3CLpro (PDB entry 6LU7) with depicted Cys 145 residue. a) after 250 (brown) and 300 ns (light blue) of MD simulation, b) representative conformations of clusters 1 (brown), 2 (light blue), and 3 (pink) with the most significant conformational differences highlighted.

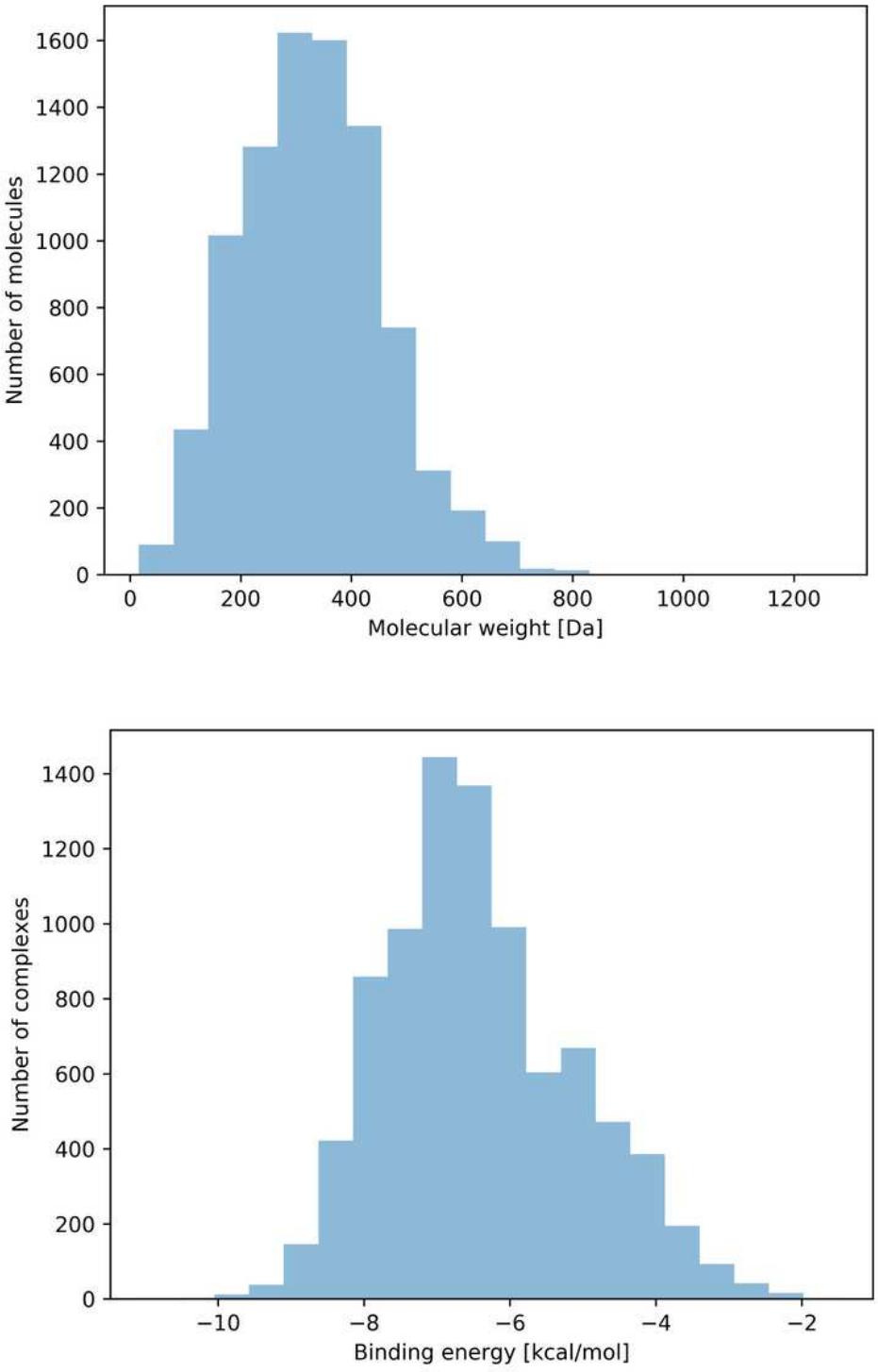

Figure 4

Distribution of molecular weights of ligands in the DrugBank database (top) and of binding energies as a result of virtual screening to SARS-CoV-2 3CLpro (bottom) (PDB entry 6LU7).

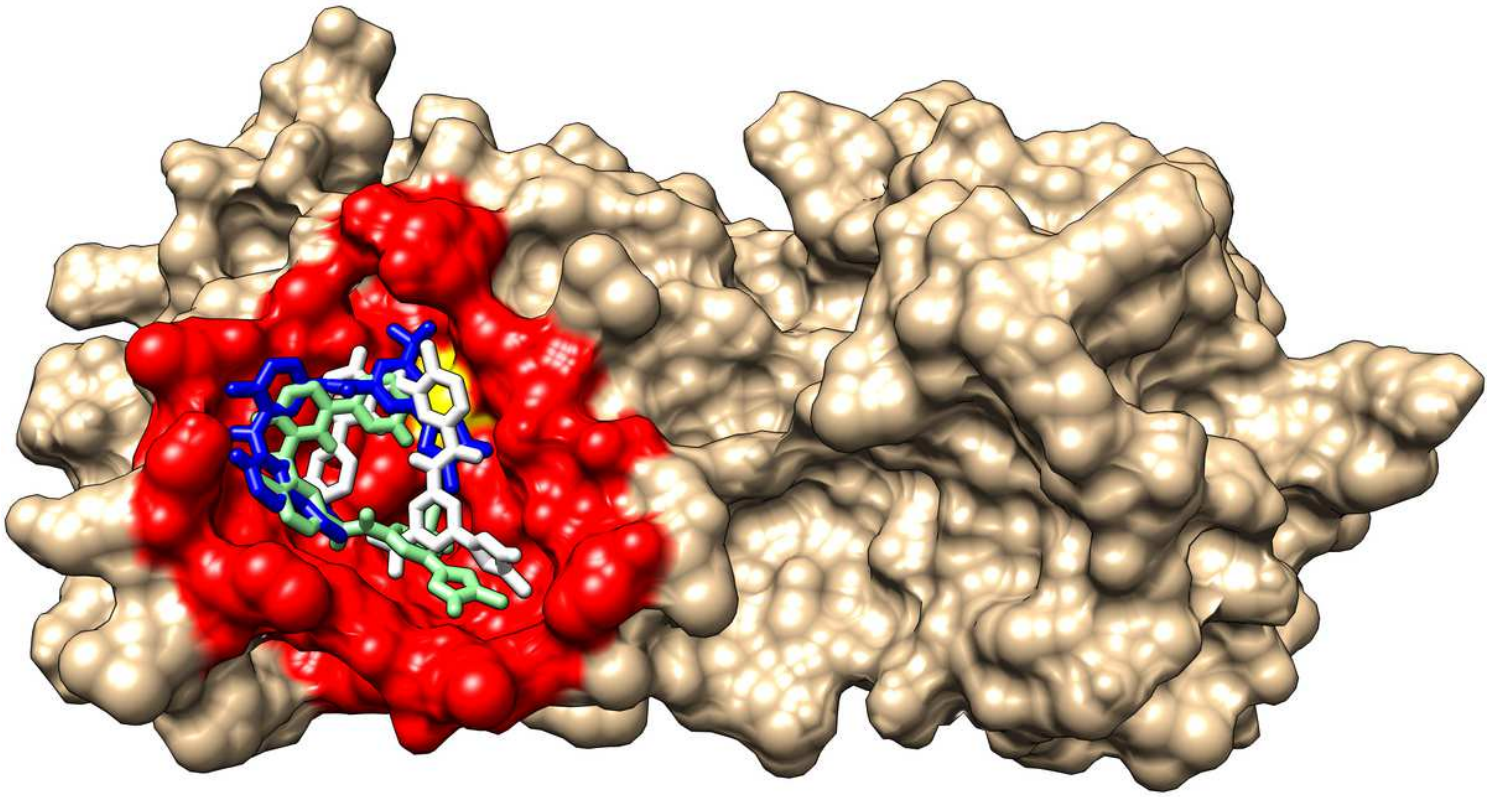

Figure 5

Active site of SARS-CoV-2 3CLpro (PDB entry 6LU7) (red) with Cys 145 (yellow) with best three conformations of nilotinib (conformations with binding energies of -8.7 (blue), -8.7 (light green), -8.5 kcal/mol (white).

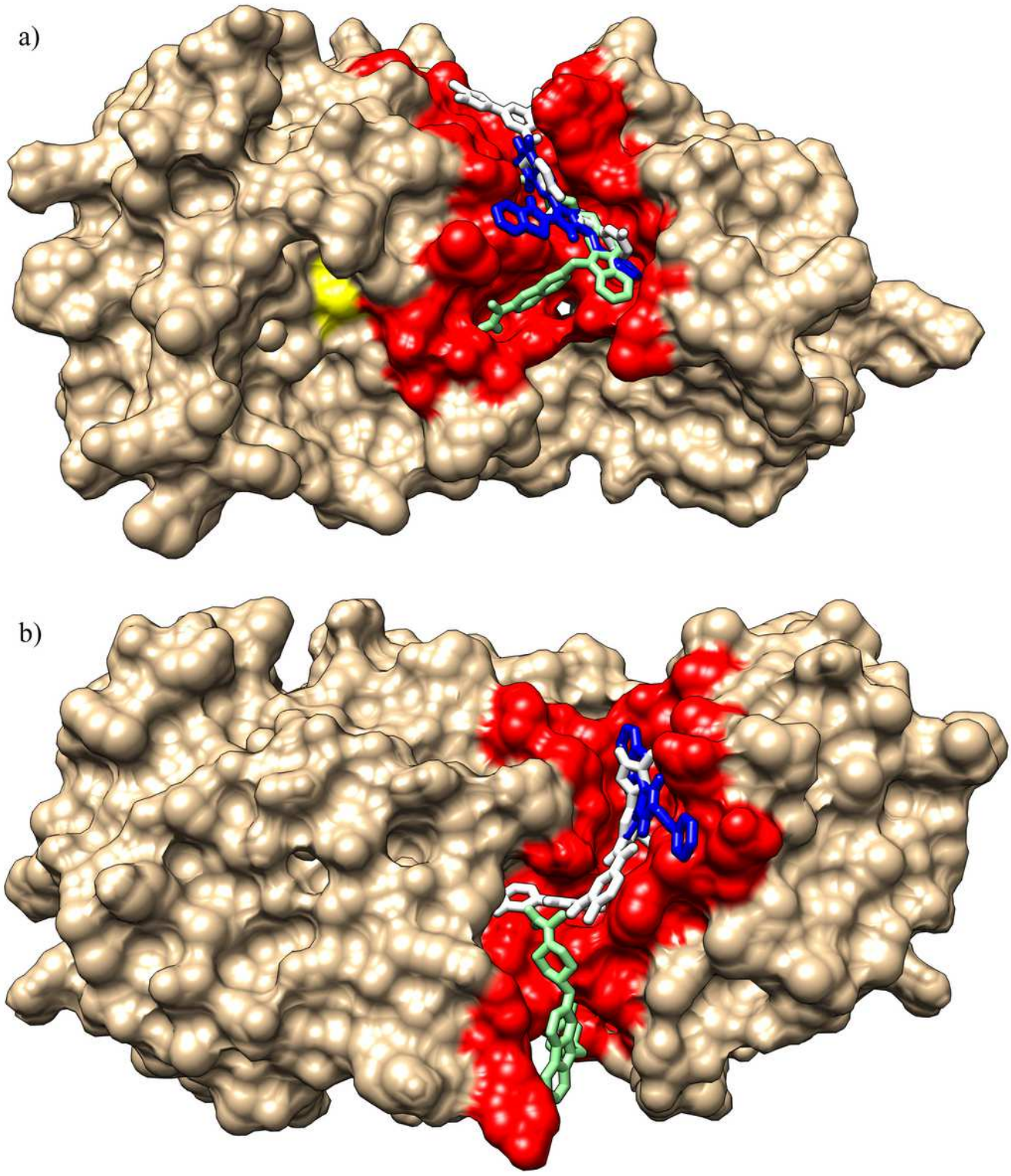

Figure 6

SARS-CoV-2 3CLpro (PDB entry 6LU7) with depicted Cys 145 (yellow) and bound ligands DB01897 (blue), Zk-806450 (light green), and nilotinib (white) in grooves of the two highest populated conformations. a) conformation 1, groove 1, b) conformation 2, groove 2.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupportingInformation.zip