

Exploring the effectiveness of SARS-CoV-2 3CLpro inhibitors through molecular dynamics and binding free energy simulations

ANDREY PINHEIRO (✉ andrey.pinheiro@icb.com.br)

universidade federal do Pará <https://orcid.org/0000-0002-4972-980X>

GABRIEL XAVIER

universidade federal do Pará

ANDREI SIQUEIRA

universidade federal do Pará

ALEX LIMA

universidade federal do Pará

DÉLIA AGUIAR

universidade federal do Pará

EVONNILDO GONÇALVES

universidade federal do Pará

Short Report

Keywords: Molecular Dynamics, Mpro, 3CLpro, COVID-19

Posted Date: October 2nd, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-84858/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

In December 2019, in the city of Wuhan in China, a novel Coronavirus was identified as the causative agent of a Severe Acute Respiratory Syndrome, later called Corona Virus Disease 2019 (COVID-19). Since the identification of the agent and sequencing of its genome, proteases such as 3CLpro and PLpro that participate in the viral cycle have been identified as possible pharmacological targets. This study aimed to evaluate the affinity of SARS-CoV-2 3CLpro (PDB 6LU7) concerning promising binders identified by other studies using virtual screening against the ZINC database and other molecules within the possibility of inhibiting the protease, such as Hydroxychloroquine, Chloroquine, and Remdesivir. Around 1,140 ns of molecular dynamics simulations were performed to evaluate stability and binding free energy values of protein-ligand complexes (60 ns for each ligand). The estimated affinity, based on 5000 frames trajectories, revealed that N3, 11b, remdesivir and a ZINC database ligand are the most promising inhibitors of 3CLpro. Given that most studies present energy results based on docking runs and one-frame coordinates analyses, the results found may present a more accurate energy value and may help experimental approaches to developing a drug against COVID-19. Despite the importance of vaccine development, alternative strategies, such as specific viral inhibitors, are important to reduce the impact of the disease on people.

Introduction

In December 2019 in Wuhan city in China, local doctors identified a new coronavirus as the causative agent of a severe acute respiratory syndrome type [1] later called Coronavirus Disease 2019 (COVID-19). COVID-19 is a disease caused by SARS-CoV-2 [2], SARS are commonly characterized by high fever, malaise and non-productive cough and in some severe cases can cause generalized interstitial infiltration of the lungs, leading to the cases of intubation and mechanical ventilation of those affected [3]. From the appearance of the first cases in China in December 2019 up to the month of July 2020, about 15,785,641 cases and 640,016 deaths had been reported [4]. Because of that, several studies have been performed seeking to find a treatment or vaccine for the virus.

The genome of the human coronavirus has approximately 29.8 to 29.9 kb ssRNA [5], which decodes various polypeptides that give rise to structural and non-structural proteins; these polypeptides are proteolytically cleaved to form several proteins. Some of these proteins are of fundamental importance for the viral cycle [6], such as two overlapping proteins called pp1a and pp1ab or replicases 1a and 1ab, with molecular weights of 450 kD and 750 kD respectively [7]. These proteolytic processes are mediated by 3-chymotrypsin-like protease (3CLpro) also called Mpro and papain-like protease (PLpro) [8]. This information makes these proteases potential targets for anti-coronavirus inhibitors [2, 6, 9]. The crystallographic structure of 3CLpro revealed that the protein is composed of three domains: residues 8 to 99 make up the first domain, 100 to 183 the second domain, both composed of beta antiparallel sheets; the third domain consists of five helices that are responsible for the proteolytic activity of the protein (residues 200 to 300). A large looping of 16 residues (184 to 199) connects domains II and III [10].

After comparing with sequences of other viruses of the same family, such as mouse hepatitis virus (MHV), bovine coronavirus (BCoV) and SARS-CoV, it was observed that in both species the 3-chymotrypsin-like protease had a high identity indices, 50%, 49% and 40% respectively [11]. This fact allowed the identification of the substrate binding site, located in a cleft between domains I and II, similar to cysteine proteases and serine proteases [10, 11]. The catalytic site of HCoV 229E Mpro has histidine and cysteine (His41 and Cys 144) differing only because it does not have the catalytic triad characteristic of this type of proteases, since that in the third position it has valine (Val84) instead of cysteine [6, 10, 11].

Virtual screening studies were realized to find inhibitor candidates among thousands of molecules on compounds databases [12, 13] and the results were promising. Here, the best candidates were selected based on docking results and toxicity predictions, for simulation by molecular dynamics and binding free energy calculations for corroborating their function as SARS-CoV-2 protease inhibitors.

Materials And Methods

Protein Structure

Currently, there are 57 structures of Mpro placed in the Protein Data Bank. This study, used the Mpro structure of SARS-CoV-2, achieved by X-Ray Diffraction, with 2.16 Å resolution (PDB code: 6LU7) [10]. The ligand selection was performed based on their toxicity levels according to the classification made by HOFMARCHER [14]. Ligands with toxicity values less than 10 were subjected to molecular docking. Additionally, the docking and molecular dynamics were performed for ligands such as Hydroxychloroquine, Chloroquine [15, 16], Remdesivir [17–19], N3 [20] and 11b [21], to serve as an energy parameter to assess the affinity of the other ligands – positive control. The structure of the ligands with possible 3CLpro inhibitory activities, derived from the Virtual Screening carried out by HOFMARCHER [14] were obtained from the ZINC database (<https://zinc.docking.org/>). The 3D structure of ligands, such as Hydroxychloroquine (PubChem CID:3652), Chloroquine (PubChem CID:2719), and Remdesivir (PubChem CID:121304016) were obtained from the PubChem database (<https://pubchem.ncbi.nlm.nih.gov/>).

Molecular Docking

The Template Docking function of Molegro Virtual Docking 5 software was used to determine the initial conformation of the ligands at the active site of the protease. MolDock scoring function with a grid resolution of 0.30 was used to rank the solutions, toggling on Internal ES, Internal HBond, and Sp2-Sp2 torsions. Docking results were further refined by energy minimization and H-bonds to optimize hydrogen bonds after docking. Each docking had 10 runs, a population size of 100 and max iterations of 2000.

Molecular Dynamics

To determine the protonation state of the residues, we used the H++ server (<http://biophysics.cs.vt.edu/>) considering a pH of 7.0. The AMBER 18 package was used to perform all steps of preparation and

molecular dynamics [22]. Cl⁻ and Na⁺ ions were used to neutralize possible system charges and the system was solvated using the TIP3P water model in a cubic box of 8 Å. Sander performed the energy minimization in 5 stages. The first four stages used 3000 cycles of steepest descent and 5000 cycles of conjugate gradients. In the remaining stage, 5000 cycles of steepest descent and 30000 cycles of conjugate gradients were used. The heavy atoms had their movement restrained by a harmonic potential of 1,000 Kcal/mol*Å², while in the final step all atoms were free. For heating and equilibration, 14 steps were used. A harmonic potential of 25 Kcal/mol*Å² was employed to restrain the heavy atoms in the initial stages. In step 13, the restraining force was set to 0. The temperature was gradually increased until it reached 300 K. Langevin Dynamics (thermostat) was employed with a collision frequency of 3.0 ps⁻¹. The heating procedure lasted 650 picoseconds until step 13 and was performed using an NVT ensemble. Afterwards, a 2-nanosecond equilibration phase was employed in an NPT ensemble. The SHAKE algorithm was employed to restrict the vibration of the ligations of all hydrogen atoms. The Particle Mesh Ewald method was used for calculating electrostatic interactions using a cutoff value of 8.0 Å. Thus, 60 ns MD was produced in an NVT ensemble, for each system. The cpptraj module was used to compute the Root Mean Square Deviation (RMSD) of both trajectories, considering the heavy atoms of the main chain. For this study, five thousand snapshots from the last 10 ns of the molecular dynamics simulations were used to calculate the free energies using MM-GBSA and MM-PBSA methods. For this study, five thousand snapshots from the last 10 ns of the molecular dynamics simulations were used to calculate the free energies using MM-GBSA and MM-PBSA methods.

Results And Discussion

The dynamics results are available in figure 1, with residues that interacting with ligand, and the RMSD information on the N3, Remdesivir, and 11b ligands. The results show that the stabilizing control ligands have high stability with respect to the displacement of alpha carbons along the trajectory; the remdesivir ligand showed homogeneity in its RMSD along with the molecular dynamics simulation; however, it presented high values of RMSD.

The values of free binding energy calculated by the MMGBSA and MMPBSA methods are shown in table 1, the values demonstrate that the control ligands show a considerable affinity for the active protease site, while the virtual screening ligands did not present values that were so considerable.

Among the ligands selected by virtual screening and those analyzed by this study, the most promising was ligand ZINC000016317677, with energies of -38.5395 MMGBSA and -36.3917 MMPBSA. However, it does not have affinity significant enough to be considered an effective protease inhibitor and the other ligands coming from the virtual screening also did not show enough significant affinity to be considered Mpro inhibitors. This fact confirms the importance of using combined bioinformatic together since this is a purely *in silico* area, so the combined use of the available tools is of utmost importance in order for the results to be more robustly ratified, providing solid results for beginning or not beginning in *vitro* studies, thus optimizing research time and reducing costs.

N3, 11b,chloroquine and hydroxychloroquine

The values of free energy generated based on the MMGBSA and MMPBSA methods from molecular dynamics revealed that the irreversible Mpro N3 and 11b inhibitor [10] has a significant affinity to block the active protease system, which in the energy decomposition showed significant interactions with residues HIS41, MET49, CYS145, MET165 and HIST163, HIE164, MET165 respectively, as shown in figure 1. Similar studies proposed N3 connection energy of approximately -64.32 kcal/mol, using the prime MMGBSA tool [13], while the one found by this study was -44.6905 kcal/mol in the MMGBSA method and -39,4006 kcal/mol in the MMPBSA method, slightly discordant values. It is noteworthy that the AMBER module responsible for computing the free energy performs the calculations based on trajectory data for 5000 snapshots, being able to perform calculations with more precision because it takes into account the conformational variations of the ligand and receiver over time [23]. Another method used to perform the solvation-free energy calculation is the Prime MMGBSA module [24]. However, it does not use information from Molecular Dynamics trajectories, and performs the calculation without taking any into consideration. Because the DM simulations are closer to reality, the method performed by the AMBER module can therefore be considered more accurate.

Furthermore, other ligands with possible *in vitro* antiviral activity such as HCQ and HC [16] did not show enough significant values of free energy to be considered enzyme inhibitors, which means that they are inefficient against 3CLpro. However, there are still other mechanisms proposed for the antiviral action of these drugs, such as an adsorption inhibitor [15], which prevents Spike proteins of the viral envelope from binding with the angiotensin-converting proteins ACE2 [25], or acting on intracellular acidic organelles that are important for membrane fusion, lysosomes/endosomes [16].

Remdesivir

The drug remdesivir has already been used against MERS and SARS and its mechanism of action is known to be a basic analog that interrupts the extension of the RNA strand in viral replication [25]. In this study, it was proven to be promising as an inhibitor of the studied protease, since it was the molecule that presented free binding energy values closest to the binding energy of the N3 and 11b inhibitor[10].The binder interacts significantly with residues ASN142, GLY143, HIS163, and MET165, and the binder interacts weakly with the catalytic residues of the HIS41 and SER144 active sites, but according to the free energy values of -40.2158 in the MMGBSA and -29.3688 MMPBSA, bringing up the hypothesis that it can be used as a drug to inhibit 3CLpro, significantly restrain viral replication and thus decrease the host's viral titration. *In vitro* [18] and *in vivo* [25] studies confirm the effectiveness of remdesivir as a viral inhibitor, such as the study published by CAO [25], in which there was a significant decrease of viral titration in mouse lung cells.

Conclusion

For this study 19 promising candidates to be SARS-CoV-2 Mpro inhibitors were chosen to be evaluated by molecular dynamics and an extensive binding free energy protocol. Because most studies present energy

results based on docking runs and one frame coordinate analyses, the results found may present a more accurate energy value and may help experimental approaches to develop a drug against COVID-19.N3, 11b and remdesivir were confirmed as the best 3CLpro inhibitor candidates ;additionally, a ligand discovered by molecular screening seems may be a possible alternative. Chloroquine and hydroxychloroquinedid not present energy values that justify their use in 3CLpro. Therefore, MMPBSA.py script was shown to be a valuable tool for estimating the effectiveness of bioactive compounds.

Declarations

CONFLICT OF INTEREST: The authors Andrey Henrique Gama Pinheiro, Gabriel Albuquerque Xavier, Andrei Santos Siqueira, Alex Ranieri Jerônimo Lima, Délia Cristina Figueira Aguiar, and Evonnildo Costa Gonçalves declare that he/she has no conflict of interest.

FUNDING: This study was funded by Fundação Amazônia de Amparo a Estudos e Pesquisas do Pará (FAPESPA) (ICAAF 099/2014), and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) 311686/2015-0 (ECG).

ETHICAL APPROVAL: This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Hui DS, I Azhar E, Madani TA, et al (2020) The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health – The latest 2019 novel coronavirus outbreak in Wuhan, China. International Journal of Infectious Diseases 91:264–266. <https://doi.org/10.1016/j.ijid.2020.01.009>
2. Tahir M, Alqahtani SM, Alamri MA, Chen L (2020) Structural basis of SARS-CoV-2 3CLpro and anti-COVID-19 drug discovery from medicinal plants. Journal of Pharmaceutical Analysis. <https://doi.org/10.1016/j.jpha.2020.03.009>
3. Lee N, Hui D, Wu A, et al (2003) A Major Outbreak of Severe Acute Respiratory Syndrome in Hong Kong. 1986–1994
4. world health organization (2020) Coronavirus disease (COVID-2019) situation reports. https://www.who.int/docs/default-source/coronavirus/situation-reports/20200726-covid-19-sitrep-188.pdf?sfvrsn=f177c3fa_2
5. Khailany RA, Safdar M, Ozaslan M (2020) Genomic characterization of a novel SARS-CoV-2. Gene Reports 19:. <https://doi.org/10.1016/j.genrep.2020.100682>
6. Chen YW, Yiu CB, Wong K (2020) Prediction of the SARS-CoV-2 (2019-nCoV) 3C-like protease (3CL pro) structure: virtual screening reveals velpatasvir , ledipasvir , and other drug repurposing candidates [version 2; peer review : 3 approved]. 2:1–17
7. Herold J, Raabe T, Schelle-Prinz B, Siddell SG (1993) Nucleotide Sequence of the Human Coronavirus 229E RNA Polymerase Locus. Virology 195:680–691. <https://doi.org/10.1006/viro.1993.1419>

8. Anand K, Ziebuhr J, Wadhwani P, et al (2003) (3CL pro) Structure: Basis for Design of Anti-SARS Drugs. *Nature*. 300:1763–1768
9. Gentile D, Patamia V, Scala A, et al (2020) Inhibitors of SARS-CoV-2 Main Protease from a Library of Marine Natural Products: A Virtual Screening and Molecular Modeling Study. 2020:. <https://doi.org/10.20944/preprints202003.0372.v1>
10. Jin Z, Du X, Xu Y, et al (2020) Structure of Mpro from COVID-19 virus and discovery of its inhibitors. *Nature*. <https://doi.org/10.1038/s41586-020-2223-y>
11. Anand K, Ziebuhr J, Wadhwani P, et al (2003) Coronavirus Main Proteinase (3CL^{pro}) Structure: Basis for Design of Anti-SARS Drugs. *Science* 300:1763 LP – 1767. <https://doi.org/10.1126/science.1085658>
12. Hofmarcher M, Mayr A, Rumetshofer E, et al (2020) Large-Scale Ligand-Based Virtual Screening for SARS-CoV-2 Inhibitors Using Deep Neural Networks. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3561442>
13. Mittal L, Kumari A, Srivastava M, et al (2020) Identification of potential molecules against COVID-19 main protease through structure-guided virtual screening approach. *Journal of Biomolecular Structure and Dynamics* 0:1–19. <https://doi.org/10.1080/07391102.2020.1768151>
14. Hofmarcher M, Mayr A, Rumetshofer E, et al (2020) Large-scale ligand-based virtual screening for SARS-CoV-2 inhibitors using deep neural networks
15. Fantini J, Di Scala C, Chahinian H, Yahi N (2020) Structural and molecular modelling studies reveal a new mechanism of action of chloroquine and hydroxychloroquine against SARS-CoV-2 infection. *International Journal of Antimicrobial Agents* 55:105960. <https://doi.org/10.1016/j.ijantimicag.2020.105960>
16. Liu J, Cao R, Xu M, et al (2020) Hydroxychloroquine, a less toxic derivative of chloroquine, is effective in inhibiting SARS-CoV-2 infection in vitro. *Cell discovery* 6:16. <https://doi.org/10.1038/s41421-020-0156-0>
17. Raju V, Munikumar M, Ramakrishna U, et al (2020) Remdesivir (GS-5734) as a therapeutic option of 2019-nCOV main protease – in silico approach. *Journal of Biomolecular Structure and Dynamics* 0:1–14. <https://doi.org/10.1080/07391102.2020.1781694>
18. Wang M, Cao R, Zhang L, et al (2020) Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Research* 30:269–271. <https://doi.org/10.1038/s41422-020-0282-0>
19. Naik VR, Munikumar M, Ramakrishna U, et al (2020) Remdesivir (GS-5734) as a therapeutic option of 2019-nCOV main protease–in silico approach. *Journal of Biomolecular Structure and Dynamics* 0:1–14. <https://doi.org/10.1080/07391102.2020.1781694>
20. Jin Z, Du X, Xu Y, et al (2020) Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature* 582:289–293. <https://doi.org/10.1038/s41586-020-2223-y>
21. Dai W, Zhang B, Jiang XM, et al (2020) Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. *Science* 368:1331–1335. <https://doi.org/10.1126/science.abb4489>

22. D.A. Case, K. Belfon, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, G. Giambasu, M.K. Gilson, H. Gohlke, A.W. Goetz, R. Harris, S. Izadi, S.A. Izmailov, K. Kasavajhala, A. Kovalenko, R. Krasny, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, V. Man, K.M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, A. Onufriev, F.Pan, S. Pantano, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, N.R. Skrynnikov, J. Smith, J. Swails, R.C. Walker, J. Wang, L. Wilson, R.M. Wolf, X. Wu, Y. Xiong, Y. Xue, D.M. York and P.A. Kollman (2018) Amber 18. University of California, San Francisco
23. Wang E, Sun H, Wang J, et al (2019) End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. *Chem Rev* 119:9478–9508.
<https://doi.org/10.1021/acs.chemrev.9b00055>
24. Li J, Abel R, Zhu K, et al (2011) The VSGB 2.0 model: a next generation energy model for high resolution protein structure modeling. *Proteins* 79:2794–2812. <https://doi.org/10.1002/prot.23106>
25. Cao Y chen, Deng Q xin, Dai S xue (2020) Remdesivir for severe acute respiratory syndrome coronavirus 2 causing COVID-19: An evaluation of the evidence. *Travel Medicine and Infectious Disease* 35:101647

Table

Due to technical limitations, table 1 is only available as a download in the supplemental files section.

Figures

FIGURE 1

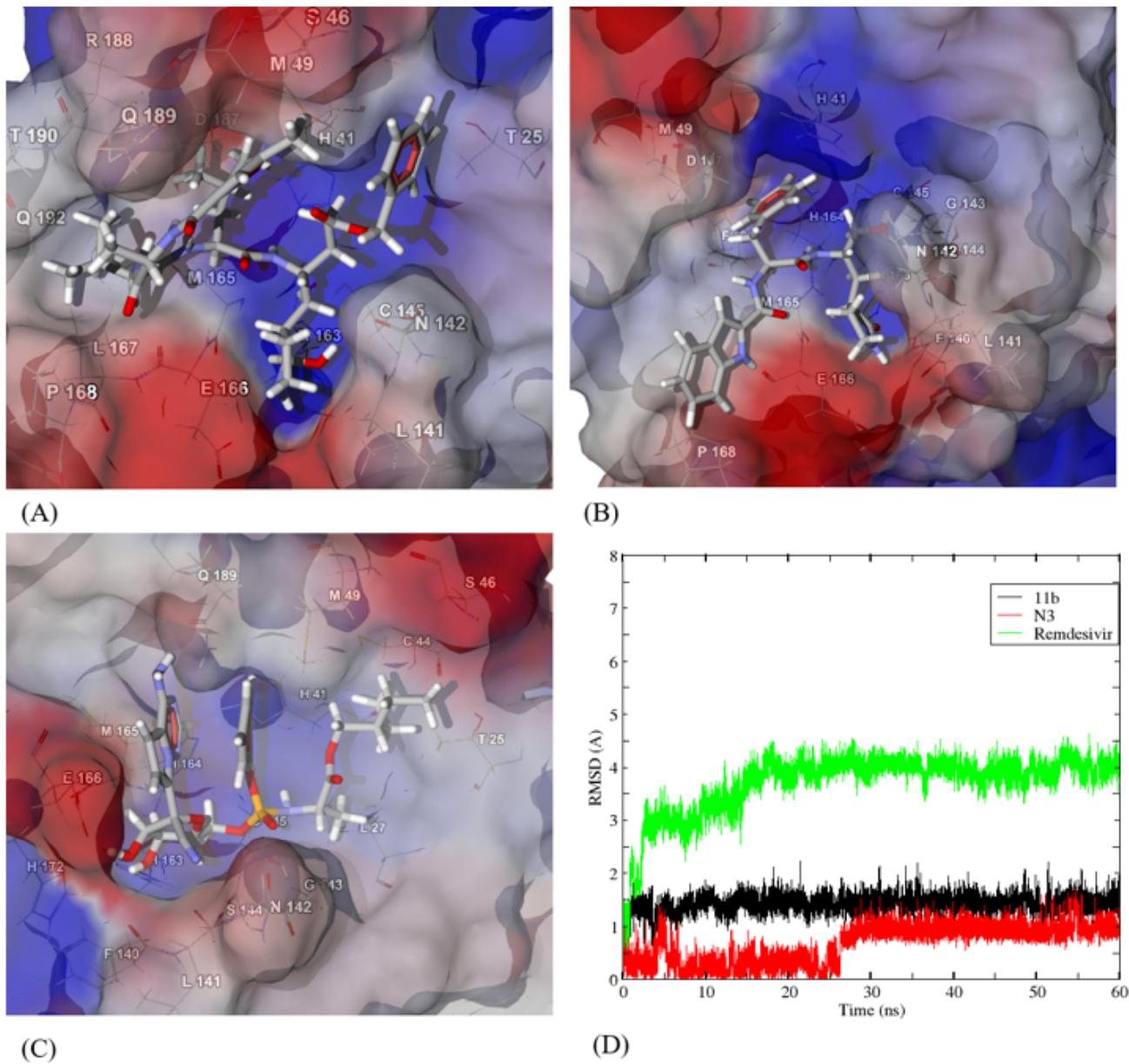


Figure 1

Representing the conformation of the ligands in the active site of the protease and the residues that contribute significantly to the binding and the RMSD values of the alpha carbons over the 60ns; (A) the position of ligand N3 in the active site interacting with residues ASN142, GLY143, HIS163, and MET165; (B) ligand 11b in the active site interacting with residues HIST163, HIE164, and MET165; (C) the residues from the active site interacting with the drug Remdesivir, the residues that interact are ASN142, GLY143, HIS163, and MET165; (D) RMSD of binders Remdesivir, 11b and N3.

Supplementary Files

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

- [Table1.docx](#)