

Targeting the Coronavirus SARS-CoV-2: computational insights into the mechanism of action of the protease inhibitors Lopinavir, Ritonavir, and Nelfinavir

CURRENT STATUS: UNDER REVIEW

natureresearch

Giovanni Bolcato

Molecular Modeling Section (MMS), Department of Pharmaceutical and Pharmacological Sciences, University of Padova

Maicol Bissaro

Molecular Modeling Section (MMS), Department of Pharmaceutical and Pharmacological Sciences, University of Padova

Matteo Pavan

Molecular Modeling Section (MMS), Department of Pharmaceutical and Pharmacological Sciences, University of Padova

Mattia Sturlese

Molecular Modeling Section (MMS), Department of Pharmaceutical and Pharmacological Sciences, University of Padova

Stefano Moro

Molecular Modeling Section (MMS), Department of Pharmaceutical and Pharmacological Sciences, University of Padova

✉ stefano.moro@unipd.it *Corresponding Author*

DOI:

10.21203/rs.3.rs-20948/v1

SUBJECT AREAS

Computational Biology

KEYWORDS

SARS-CoV-2, Supervised Molecular Dynamics, ligand recognition, lopinavir, ritonavir, nelfinavir

Abstract

Coronavirus SARS-CoV-2 is a recently discovered single-stranded RNA (ssRNA) betacoronavirus, responsible for a severe respiratory disease known as coronavirus disease 2019 (COVID-19), which is rapidly spreading. Chinese health authorities, as a response to the lack of an effective therapeutic strategy, started to investigate the use of lopinavir and ritonavir, previously optimized for the treatment and prevention of HIV/AIDS viral infection. Despite the clinical use of these two drugs, no information regarding their possible mechanism of action at the molecular level is still known for SARS-CoV-2. Very recently, the crystallographic structure of the SARS-CoV-2 main protease (Mpro), also known as C30 Endopeptidase, was published. Starting from this essential structural information, in the present work we have exploited Supervised Molecular Dynamics (SuMD), an emerging computational technique that allows investigating at an atomic level the recognition process of a ligand from its unbound to the final bound state. In this research, we provided molecular insight on the whole interaction pathway of lopinavir, ritonavir, and nelfinavir, three potential C30 Endopeptidase inhibitors, with the last one taken into consideration due to the promising in-vitro activity shown against the structurally related SARS-CoV protease.

Introduction

Coronavirus SARS-CoV-2, previously known as 2019-nCoV, is a recently discovered single-stranded RNA (ssRNA) betacoronavirus, responsible for a severe pathological condition known as coronavirus disease 2019 (COVID-19).¹ Since it was first identified in December 2019, this novel coronavirus has infected about 80000 people around the world, of whom more than 2000 have lost their lives due to a severe respiratory illness.²

The first outbreak of this new disease originally took place in the city of Wuhan (China), rapidly spreading in the southeast of Asia and, recently, in other continents like Europe, North America and Africa.¹ The astonishing rate at which COVID is expanding compared to previous coronavirus related diseases (SARS-CoV and MERS-CoV), in conjunction with the absence of approved drugs or effective therapeutic approaches for its treatment, have gathered the attention of the international community, which is promoting a cooperative effort to face this emergency situation.^{3,4} On January 2020 indeed,

the International Health Regulations Emergency Committee of the World Health Organization declared the outbreak a “public health emergency of international concern” in responding to SARS-CoV-2. Unfortunately, the timeline characterizing a typical drug discovery process badly couples with the urgency of finding a cure for the already infected patients as rapidly as possible. In this kind of scenario, it is of paramount importance to accelerate the early stages of the drug discovery process for COVID-19 treatment, and for all possible future emergencies.⁵

The early isolation of the SARS-CoV-2 genome from ill patients represented a first crucial outcome, making it possible to highlight an important sequence identity (~80% of conserved nucleotides) with respect to the original SARS-CoV epidemic virus.⁶ In light of this similarity, some therapeutic strategies could be inherited from other genetically related CoV diseases.

A possible target is for example represented by structural viral proteins, therefore interfering with the assembly and the internalization of the pathogen into the host, which was shown to occur also in this case through the Angiotensin-converting enzyme II (ACE2) receptor. From this perspective, the development of a vaccine is desirable, and it is foreseen that the first candidates will be advanced to clinical phase I around mid-2020.⁷⁻⁹

In the meantime, however, a great effort involves the targeting of non-structural viral proteins which are instead essential for the viral replication and the maturation processes, thus representing a crucial and specific target for anti-COVID drug development.^{3,10} In this regard, the crystallographic structure of the SARS-CoV-2 main protease (Mpro), also known as C30 Endopeptidase, was elucidated and made available to the scientific community with impressive timing, just a few weeks after the first COVID-19 outbreak (PDB ID: 6LU7). The structural characterization of the protease, which shares 96.1% of its sequence with those of SARS-CoV, has revealed a highly conserved architecture of the catalytic binding site.

As a result, Structure-Based Drug Discovery techniques (SBDD) can now be applied to efficiently speed up the rational identification of putative Mpro inhibitors or to drive the repurposing process of known therapy. This latter route is particularly attractive, as it allows to significantly shrink the time required to access the first phases of clinical trials, without compromising patient safety. A multitude

of research groups has begun to apply computational approaches, such as molecular docking based virtual screening (VS), to evaluate already approved FDA approved drugs against the aforementioned viral protease.¹¹⁻¹⁴ Many of these studies have found convergence in suggesting compounds inhibitors of the human immunodeficiency viruses (HIV) as possible anti-COVID candidates; this is surprising considering the important structural differences existing among these two homologous enzymes. The repositioning of HIV antiviral drugs for the treatment of coronavirus infections found, however, a foundation in the scientific literature of the past 20 years. Some of these compounds have therefore been experimentally investigated, showing promising activity, both in the case of SARS-CoV and MERS-CoV outbreak.^{15,16}

Moreover, at least three randomized clinical trials are currently being held in China in order to evaluate the therapeutic efficacy of Lopinavir and Ritonavir, a combination of HIV protease inhibitors, in COVID-19 treatment.⁷ In this perspective and preliminary computational research, we took advantage of the recently solved crystallographic structure of SARS-CoV-2 Mpro to perform a cutting edge *in-silico* investigation. (Figure 1)

Supervised Molecular Dynamics (SuMD), an emerging technique allowing to investigate at an atomic level of detail the molecular recognition process, was exploited to characterize the putative binding mechanism of three HIV protease inhibitors.¹⁷⁻¹⁹ In detail, along with the aforementioned combination of Lopinavir and Ritonavir, also Nelfinavir was taken into consideration, due to the promising *in-vitro* activity shown by this compound against the structurally related SARS-CoV protease.²⁰ SuMD protocol implements a tabu-like algorithm that controls the sampling of short unbiased MD trajectories, each of which hundreds of picoseconds (ps) long. In detail, simulation steps are accepted only when describing a ligand approaching a known binding site, otherwise, the simulation is discharged and restarted from the previous coordinate set. The combination of all productive SuMD simulation steps represents, therefore, a putative molecular recognition trajectory collected, differently from brute force MD, in a very competitive computational time not exceeding the nanoseconds (ns) timescale. Contrary to molecular docking, SuMD simulations fully consider both the flexibility characterizing the protein target during the binding event and the contribution played

by water molecules during the recognition. Moreover, the study is not limited to a possible final state but allows peeking dynamically at the whole process of recognition, also identifying putative metastable binding sites.

Results

The combination of the structurally related antiviral protease inhibitor Lopinavir and Ritonavir, commercially known with the name Kaletra, represent an effective therapeutic weapon ensuring an adequate and durable suppression of viral load in HIV positive patients. The synergistic coadministration of these two compounds exploits low-dosage concentration of Ritonavir which, inhibiting the metabolic inactivation of Lopinavir, acts as a pharmacokinetic enhancer.²¹ Following a preliminary favorable clinical response in SARS-CoV related diseases, the combination of the drug is currently under investigation also against SARS-CoV-2, with at least three randomized clinical trials undergoing with Chinese infected patients.¹⁵ In our computational study, we considered Lopinavir and Ritonavir as two independent inhibitors, performing separate SuMD binding simulations, which results are herein reported and analyzed.

As highlighted in Figure 2 (Panel B) about 20 ns proved to be sufficient to sample a putative Lopinavir recognition trajectory with SARS-CoV-2 protease. At a distance of about 15 Å from the binding site, the first molecular contacts are recorded (Figure 2—Panel C, D and Video 1), which guide the subsequent accommodation of the ligand into the catalytic site. The predicted final state is stabilized by a double hydrogen bond interaction with residue Glu166 backbone, tightly anchoring the inhibitor (Figure 2 -Panel A). This strong and persistent interaction (Figure 2—Panel B) is known to be crucial in many SARS-CoV complexes and moreover, was also found to stabilize the covalent peptidomimetic compound crystallized in the recently published SARS-CoV-2 Mpro structure. In addition, the cyclic urea moiety of Lopinavir mediates a hydrogen bond interaction with the side chain of Gln189, another residue whose importance has been elucidated by means of several SARS-CoV three-dimensional complexes.

Despite the modest pharmacodynamic contribution made by Ritonavir in the combined formulation under investigation by the Chinese scientific community, in which the drugs act as a pharmacokinetic

enhancer rather than a protease inhibitor, we still tried to elucidate its putative molecular recognition pathway. Also, in this case, 20 ns of SuMD simulation time were sufficient to sample a binding trajectory (Figure 3 – Panel B). Although some key interactions – i.e. hydrogen bond network with residue Glu166 and Gln189 – are appreciable also in this final state (Figure 3 —Panel A, D and Video 2), a comparative analysis of the Interaction Energy Landscape graphs (Panel C of Figure 2 and 3) suggests lower energy stability of the SuMD predicted binding mode, when compared with that characterizing Lopinavir. A reason could be seeking on the non-optimal accommodation of Ritonavir urea moiety, which floats outside the binding site exposed to the bulk solvent during all the simulation (Video 2).

In light of the promising experimental results shown by Nelfinavir, which milder the cytopathic effect induced by SARS-CoV infection strongly inhibiting the virus replication, we decided to computationally evaluate its possible molecular recognition mechanism also against SARS-CoV-2 protease. As reported in Figure 4 (Panel B), a slightly longer SuMD simulation was necessary to fully describe a putative Nelfinavir binding trajectory. Once it has approached the vestibular region of the protease catalytic site, the ligand spends the first 20 ns negotiating the accommodation with a series of polar residues with which it mediates intermittent interactions, as highlighted in the interaction energy fingerprint (Figure 4 - Panel D and Video 3). The importance of this metastable site is also depicted in the Interaction Energy Landscape (IEL) graphic (Figure 4—Panel C), from which it is possible to notice a highly populated region presenting ligand-protein interaction energy comparable to the final states previously described for the other two inhibitors. The last 10 ns of the simulation were characterized by a series of conformational rearrangements, which resulted in an optimal Nelfinavir accommodation within the protease binding cleft. In detail, the insertion of the phenyl thioether moiety in a small protein sub-pocket drives the establishment of a dense hydrogen bond network, tightly anchoring the inhibitor to the protease. As shown in Figure 4 (Panel A), SuMD predicted binding mode of Nelfinavir is characterized by great analogies with that of the originally crystallized covalent peptidomimetic compound. Residues His164, Glu166, Gln189, Thr190, and Gln196 mediate a series of directed or water-bridged hydrogen bonds interactions. Moreover, as highlighted in Figure 4 (Panel D), on the last

ns of the simulation a salt bridge interaction occurs between the side chain of residue Glu166 and the octahydro-1H-isoquinoline charged moiety of Nelfinavir. This interaction is probably responsible for the great energy stabilization that characterizes the final state which, as represented by the graphic IEL in Figure 4 (Panel C), is quantitatively greater than that of Lopinavir and Ritonavir. Intriguingly, mutagenesis studies have corroborated the crucial role played by this residue. Mutation of Glu166 correlated therefore with the block of substrate-induced dimerization of the main protease, both in SARS-CoV and in MERS-CoV.^{22,23}

Discussion

In the last two decades, three major outbreaks of coronavirus-related diseases SARS-CoV, MERS-CoV and ultimately SARS-CoV-2 have been responsible for significant public health issues, along with dramatic social-economic consequences. The process of drug discovery often undergoes timelines which are difficult to reconcile with the urgency and the need to provide an effective therapeutic response to such an emergency health situation. Drug repurposing could represent a viable possibility, and this is the case for some anti-HIV compounds targeting SARS-CoV-2 C30 Endopeptidase. The molecular basis underneath their therapeutic action remains however often obscure. In this preliminary computational investigation, we have taken advantage of the recently published crystallographic structure of SARS-CoV-2 Mpro to investigate the putative binding mechanism of three antiviral compounds, previously designed as selective HIV protease inhibitors and now under investigation as anti-COVID-19 emergency treatments. SuMD protocol was in detail exploited to collect, for each of the three inhibitors, MD simulation describing the possible mechanism of molecular recognition, thus providing an atomistic insight to interpret their data of therapeutic efficacy. An interesting aspect is represented by the speed of this approach: a few days of calculation in a modest GPU cluster allowed to collect a multitude of simulations, from which it was possible to hypothesize the recognition mechanism of Lopinavir, Ritonavir, and Nelfinavir. An approach of this type, therefore, becomes crucial in all emergency situations, making it possible to overcome the lack of structural data to guide and understand the possible repositioning of already approved drugs. In this particular case study, the SuMD protocol not only allowed to hypothesize a possible recognition

method for each antiviral but also to advance some preliminary comparative considerations.

Nelfinavir, in particular, showed the best fitting for the catalytic site of SARS-CoV-2 Mpro, establishing a tight hydrogen bonds network, similar to those elucidated in the crystallographic complex for the covalent peptidomimetic compound N3. This is in line with the promising antiviral action shown by Nelfinavir in the case of SARS-CoV, whose protease has very high sequence identity with that of the novel CoV. In addition, the methodology allowed to highlight the fundamental role played by residue Glu166 in the molecular recognition mechanism, an aspect that finds confirmation in various mutagenesis studies, thus validating the descriptive ability of the computational methodology.

Methods

Software overview

MOE suite (Molecular Operating Environment, version 2018.0101) was used to perform most of the general molecular modeling operations, such as proteins and ligands preparation.²⁴ All these operations have been performed on an 8 CPU (Intel® Xeon® CPU E5-1620 3.50 GHz) Linux workstation. Molecular dynamics (MD) simulations were performed with an ACEMD3 engine on an Nvidia GPU cluster composed of 20 NVIDIA drivers, whose models go from GTX 1080 to TitanV.²⁵ For all the simulations, the ff14SB force field was adopted to describe C30 Endopeptidase protein while general Amber force field (GAFF) was adopted to parameterize small organic molecules.²⁶⁻²⁸

Structures Preparation

The three-dimensional coordinates of C30 Endopeptidase protein in complex with a covalent peptidomimetic inhibitor (N3) were retrieved from the RCSB PDB database and prepared for SuMD simulations as herein described.²⁹ Considering the perfect symmetry that characterizes this homodimeric protein, and therefore its two catalytic binding sites, only one of the two monomers was used in this computational investigation. Once the covalent ligand was removed, residue Cys145 was restored to its reduced form. Protein was then processed by means of MOE protein structure preparation tool: residues missing atoms were built according to AMBER14 force field topology. Missing hydrogen atoms were added to X-ray derived complexes and appropriate ionization states were assigned by the Protonate-3D tool.³⁰ The coordinates of three antiviral compounds were

prepared by means of MOE builder tool and subsequently moved at least 30 Å away from the catalytic protease binding cleft, a distance bigger than the electrostatic cut-off term used in the simulation (9 Å with Amber force field) to avoid premature interaction at the initial phases of the SuMD simulations.

Solvated System Setup and Equilibration

Each system investigated by means of SuMD contains a C30 Endopeptidase target macromolecule and respectively one of the three HIV antiviral compounds taken into consideration in this study, moved far away from the protein binding site as previously described. The systems were explicitly solvated by a cubic water box with cell borders placed at least 15 Å away from any protein/ligand atom, using TIP3P as a water model. To neutralize the total charge of each system, Na⁺/Cl⁻ counterions were added to a final salt concentration of 0.154 M. The systems were energy minimized by 500 steps with the conjugate-gradient method, then 500000 steps (1 ns) of NVE followed by 500000 steps (1 ns) of NPT simulations were carried out, both using 2 fs as time step and applying harmonic positional constraints on protease and ligands heavy atoms by a force constant of 1 kcal mol⁻¹ Å⁻², gradually reduced with a scaling factor of 0.1. During this step, the temperature was maintained at 310 K by a Langevin thermostat with low dumping of 1 ps⁻¹ and the pressure at 1 atm by a Monte Carlo barostat 31. The M-SHAKE algorithm was applied to constrain the bond lengths involving hydrogen atoms. The particle-mesh Ewald (PME) method was exploited to calculate electrostatic interactions with a cubic spline interpolation and 1 Å grid spacing, and a 9.0 Å cutoff was applied for Lennard-Jones interactions 32.

Supervised Molecular Dynamics (SuMD) Simulations

SuMD code, in this implementation, is written in Python and exploits the ProDy python package to perform the geometrical ligand-target supervision process 33. SuMD protocol reduces the timescale, and consequently the computational effort, required to sample a binding event in the range of nanoseconds, instead of hundreds of nanoseconds or microseconds usually necessary with unbiased MD. Sampling is improved by applying a tabu-like algorithm that monitors the distance between the ligand center of mass with respect to the protein binding site, during short unbiased MD simulations of 600 ps. Once a SuMD step has been collected, the distance points calculated at regular time

intervals are fitted into a linear function. Only productive MD steps are maintained, those in which the computed slope is negative, indicating a ligand approach toward the protease catalytic binding site. Otherwise, the simulation is restarted by randomly assigning the atomic velocities. Supervision algorithm controlled the sampling of short simulations until the distance between the ligand and the protein binding site dropped below 5 Å, then was disabled, and a classical MD simulation was performed. For each case study up to a maximum of ten SuMD binding simulations were collected, of which only the best was thoroughly analyzed and discussed in the manuscript.

SuMD Trajectories Analysis

All the SuMD trajectories collected were analyzed by an in-house tool written in tcl and python languages, as described in the original publication 19. Briefly, the dimension of each trajectory was reduced saving MD frames at a 20 ps interval, each trajectory was then superposed and aligned on the protease C α atoms of the first frames and wrapped into an image of the system simulated under periodic boundary condition. The molecular recognition was monitored by calculating for each simulation step the distance between the catalytic binding site and the center of mass of the ligand taken into consideration (Figure 2 to 4—Panel A). A ligand-protein interaction energy estimation during the recognition process was calculated using an MMGBSA protocol as implemented in AMBER 2014, plotting MMGBSA values over time 34. The MMGBSA values were also arranged according to the distances between ligand and protease binding site mass centers in the Interaction Energy Landscape plots (Figure 2 to 4—Panel B). Here, the distances between mass centers are reported on the x-axis, while the MMGBSA values on the y-axis, and are rendered by a colorimetric scale going from blue to red for negative to positive energetic values. These graphs allow evaluating the variation of the interaction energy profile at different ligand-protein distances, helping to individuate meta-stable binding states during the binding process. Furthermore, for each target investigated in this work, the nucleotides within a distance of 4 Å from the respective ligand atoms were dynamically selected, to qualitatively and quantitatively evaluated the number of contacts during the entire binding process. The most contacted residues were thus selected, to compute a per- nucleotide electrostatic and vdW interaction energy contribution with the protease target. NAMD was used for post-processing

computation of electrostatic interactions, using AMBER ff14SB force field. The cumulative electrostatic interactions were computed for the same target residues by summing the energy values frame by frame along the trajectory, and the resulting graphs were reported at the lower-right of movies provided on supplementary material (Video V1 to V3). Representations of the molecular structures were prepared with VMD software 35.

SuMD videos

Each video is composed of four synchronized and animated panels that depict the molecular trajectory obtained by the SuMD simulation considering different aspects of the simulation. The time evolution is reported on an ns scale. In the first panel (upper-left), the molecular representation of the SARS-CoV-2 main protease is shown. The protein backbone is represented by the ribbon style (pink color) and the residues within 4 Å of each ligand investigated are shown in green, orange and cyan colors respectively for Lopinavir, Ritonavir, and Nelfinavir. In the second panel (upper-right), the dynamic distance of each ligand center of mass (CM) from the respective protein catalytic binding site during the trajectory is reported. In the third panel (lower-left), the MMGBSA energy profile is reported. The animated red circle highlights the value of the corresponding frame. The trend is depicted by a continuous black line obtained by smoothing the raw data (grey circles) using a Bezier curve procedure. In the fourth panel (lower-right) cumulative electrostatic interactions are reported for the 15 protein residues most contacted by each ligand during the whole simulation.

Declarations

Acknowledgment

MMS lab is very grateful to Chemical Computing Group, OpenEye, and Acellera for the scientific and technical partnership. MMS lab gratefully acknowledges the support of NVIDIA Corporation with the donation of the Titan V GPU, used for this research.

This scientific work has been financially supported by MIUR (PRIN2017, n. 2017MT3993).

Contributions

S. M. conceived the research idea and approach. M. B. G. B. M. S. and M. P. carried out all computational simulations. M. B. and S. M. wrote the manuscript. All authors discussed the results

and reviewed the manuscript.

Competing interests

No conflict of interest to declare

References

1. Guarner, J. Three Emerging Coronaviruses in Two DecadesThe Story of SARS, MERS, and Now COVID-19. *J. Clin. Pathol.* doi:10.1093/AJCP/AQAA029
2. Coronavirus disease 2019 (COVID-19) Situation Report – Available at: <https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200224-sitrep-35-covid-19.pdf>. (Accessed: 26th February 2020)
3. Zhang, & Liu, Y. Potential Interventions for Novel Coronavirus in China: A Systemic Review. *J. Med. Virol.* jmv.25707 (2020). doi:10.1002/jmv.25707
4. Heymann, D. , Shindo, N. & WHO Scientific and Technical Advisory Group for Infectious Hazards. COVID-19: what is next for public health? *Lancet (London, England)* **395**, 542–545 (2020).
5. Mani, , Wadhvani, A. & Krishnamurthy, P. T. Drug Repurposing in Antiviral Research: A Current Scenario. *J. Young Pharm.* **11**, 117–121 (2019).
6. Gralinski, E. & Menachery, V. D. Return of the Coronavirus: 2019-nCoV. *Viruses* **12**, 135 (2020).
7. Keener, A. B. Four ways researchers are responding to the COVID-19 outbreak. *Med.* (2020). doi:10.1038/d41591-020-00002-4
8. Letko, M. & Munster, V. Functional assessment of cell entry and receptor usage for lineage B β - coronaviruses, including 2019-nCoV. *bioRxiv* 01.22.915660 (2020). doi:10.1101/2020.01.22.915660
9. Wrapp, D. et Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* (2020). doi:10.1126/science.abb2507

10. Anand, K., Yang, H., Bartlam, M., Rao, Z. & Hilgenfeld, R. Coronavirus main proteinase: target for antiviral drug therapy. in *Coronaviruses with Special Emphasis on First Insights Concerning SARS* 173–199 (Birkhäuser-Verlag, 2005). doi:10.1007/3-7643-7339-3_9
11. Li, *et al.* Therapeutic Drugs Targeting 2019-nCoV Main Protease by High-Throughput Screening. *bioRxiv* 2020.01.28.922922 (2020). doi:10.1101/2020.01.28.922922
12. Xu, *et al.* Nelfinavir was predicted to be a potential inhibitor of 2019-nCoV main protease by an integrative approach combining homology modelling, molecular docking and binding free energy calculation. *bioRxiv* 2020.01.27.921627 (2020). doi:10.1101/2020.01.27.921627
13. Liu, X. & Wang, X.-J. Potential inhibitors for 2019-nCoV coronavirus M protease from clinically approved medicines. *bioRxiv* 01.29.924100 (2020). doi:10.1101/2020.01.29.924100
14. Contini, Virtual Screening of an FDA Approved Drugs Database on Two COVID-19 Coronavirus Proteins. (2020). doi:10.26434/CHEMRXIV.11847381.V1
15. Chu, C. M. *et al.* Role of lopinavir/ritonavir in the treatment of SARS: Initial virological and clinical findings. *Thorax* **59**, 252–256 (2004).
16. Sheahan, T. P. *et al.* Comparative therapeutic efficacy of remdesivir and combination lopinavir, ritonavir, and interferon beta against MERS-CoV. *Commun.* **11**, 1–14 (2020).
17. Sabbadin, & Moro, S. Supervised molecular dynamics (SuMD) as a helpful tool to depict GPCR– ligand recognition pathway in a nanosecond time scale. *J. Chem. Inf. Model.* **54**, 372–376 (2014).
18. Cuzzolin, A. *et al.* Deciphering the Complexity of Ligand-Protein Recognition Pathways Using Supervised Molecular Dynamics (SuMD) *J. Chem. Inf. Model.* **56**, 687–

- 705 (2016).
19. Salmaso, , Sturlese, M., Cuzzolin, A. & Moro, S. Exploring Protein-Peptide Recognition Pathways Using a Supervised Molecular Dynamics Approach. *Structure* **25**, 655-662.e2 (2017).
 20. Yamamoto, N. *et al.* HIV protease inhibitor nelfinavir inhibits replication of SARS-associated coronavirus. *Biophys. Res. Commun.* **318**, 719–725 (2004).
 21. Cvetkovic, R. S. & Goa, K. Lopinavir/ritonavir: A review of its use in the management of HIV infection. *Drugs* **63**, 769–802 (2003).
 22. Cheng, S. C., Chang, G. G. & Chou, C. Y. Mutation of glu-166 blocks the substrate-induced dimerization of SARS coronavirus main *Biophys. J.* **98**, 1327–1336 (2010).
 23. Ho, B. *et al.* Critical assessment of the important residues involved in the dimerization and catalysis of MERS Coronavirus Main Protease. *PLoS One* **10**, (2015).
 24. Chemical Computing Group (CCG) Molecular Operating Environment (MOE). (2018).
 25. Harvey, M. J., Giupponi, G. & Fabritiis, G. De. ACEMD: accelerating biomolecular dynamics in the microsecond time scale. *Chem. Theory Comput.* **5**, 1632–1639 (2009).
 26. Tan, , Piana, S., Dirks, R. M. & Shaw, D. E. RNA force field with accuracy comparable to state-of- the-art protein force fields. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E1346–E1355 (2018).
 27. Wang, , Wang, W., Kollman, P. A. & Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* **25**, 247–260 (2006).
 28. Sprenger, K. G., Jaeger, V. W. & Pfaendtner, J. The general AMBER force field (GAFF) can accurately predict thermodynamic and transport properties of many ionic *J Phys Chem B* **119**, 5882–5895 (2015).

29. Berman, M. *et al.* The protein data bank. *Nucleic Acids Res.* **28**, 235–242 (2000).
30. Labute, Protonate3D: assignment of ionization states and hydrogen coordinates to macromolecular structures. *Proteins* **75**, 187–205 (2009).
31. Loncharich, R. J., Brooks, B. R. & Pastor, R. W. Langevin dynamics of peptides: The frictional dependence of isomerization rates of N-acetylalanine-N'-methylamide. *Biopolymers* **32**, 523–535 (1992).
32. Essmann, *et al.* A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577–8593 (1995).
33. Bakan, A., Meireles, M. & Bahar, I. ProDy: protein dynamics inferred from theory and experiments. *Bioinformatics* **27**, 1575–1577 (2011).
34. Miller, B. R. *et al.* py: An Efficient Program for End-State Free Energy Calculations. *J Chem Theory Comput* **8**, 3314–3321 (2012).
35. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *Mol. Graph.* **14**, 33–38 (1996).

Supplementary Video Details

Supplementary Videos Video 1, 2, and 3. Each video is composed of four synchronized and animated panels that depict the molecular trajectory obtained by the SuMD simulation considering different aspects of the simulation. The time evolution is reported in a ns scale. In the first panel (upper-left), the molecular representation of the SARS-CoV-2 main protease is shown. The protein backbone is represented by the ribbon style (pink colour) and the residues within 4 Å of each ligand investigated are shown in green, orange and cyan colours respectively for Lopinavir, Ritonavir and Nelfinavir. In the second panel (upper-right), the dynamic distance of each ligand center of mass (CM) from the respective protein catalytic binding site during the trajectory is reported. In the third panel (lower-left), the MMGBSA energy profile is reported. The animated red circle highlights the value of the corresponding frame. The trend is depicted by a continuous black line obtained by smoothing the raw data (grey circles) using a Bezier curve procedure. In the fourth panel (lower-right) cumulative

electrostatic interactions are reported for the 15 protein residues most contacted by each ligand during the whole simulation.

Figures

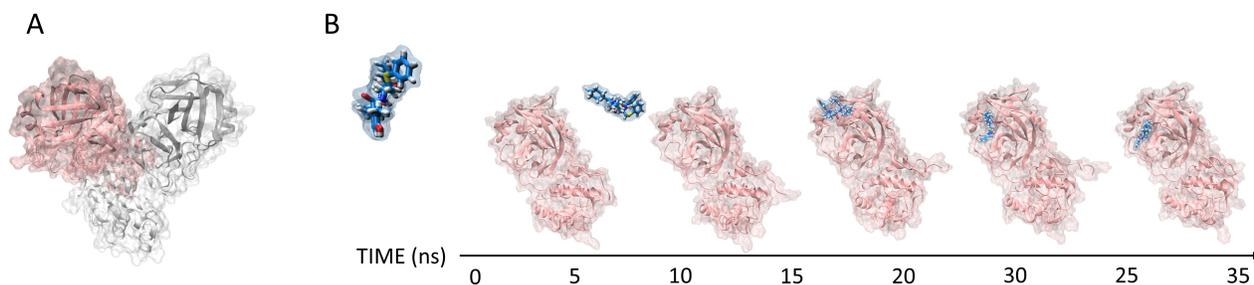


Figure 1

The crystallographic structure of SARS-CoV-2 C30 Endopeptidase exploited in our computational investigation (PDB ID : 6LU7) is reported in Panel A. The two different monomers composing the homodimeric proteases are depicted using different colors (i.e. pink and white respectively for monomer A and B). As represented on Panel B, only one chain (monomer A) was exploited in our SuMD protocol to describe the putative inhibitor binding mechanism.

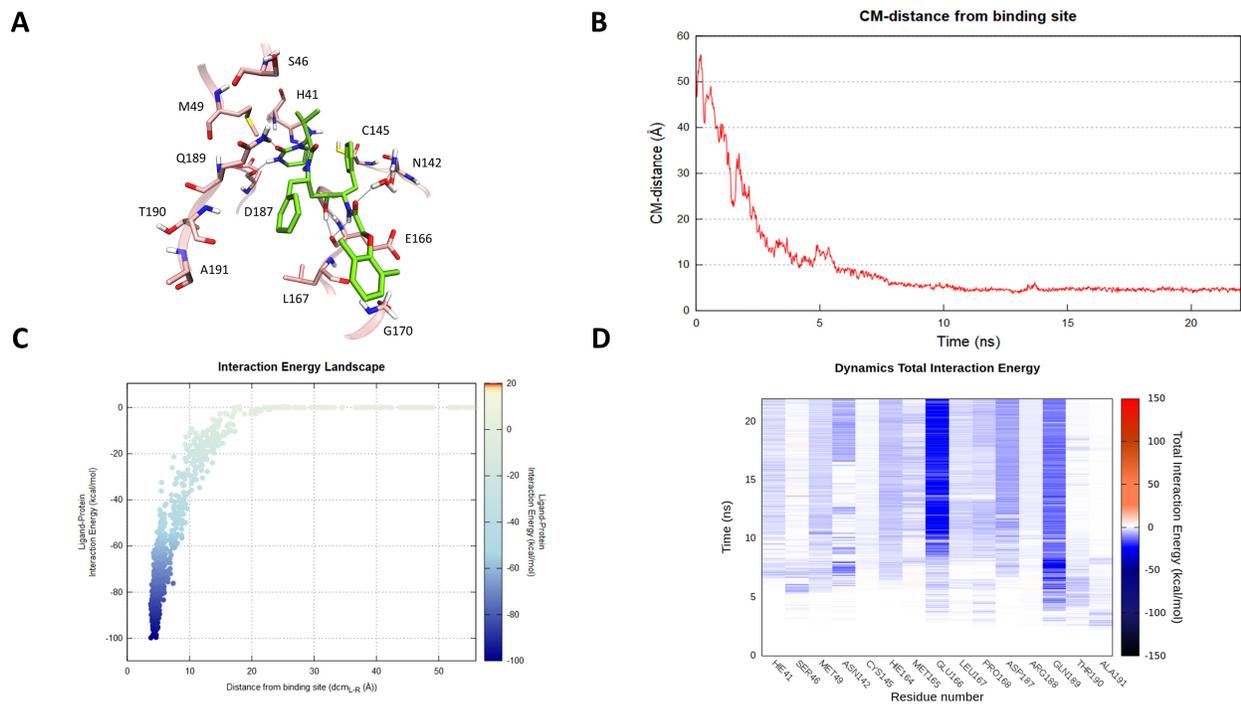


Figure 2

This panel summarizes the recognition pathway of Lopinavir against the SARS-CoV-2 main protease. (A) Lopinavir conformation sampled in the last frame of the SuMD trajectory (green-colored molecule). The residues surrounding the binding site are reported in pink color. (B) Distance between the ligand center of mass (Cm) and the catalytic binding site of the Mpro during the SuMD simulation. (C) Interaction Energy Landscape describing the protein-ligand recognition process; values are arranged according to the distances between ligand and protein target mass centers. (D) Dynamic total interaction energy (electrostatic + vdW) computed for most contacted M pro residues.

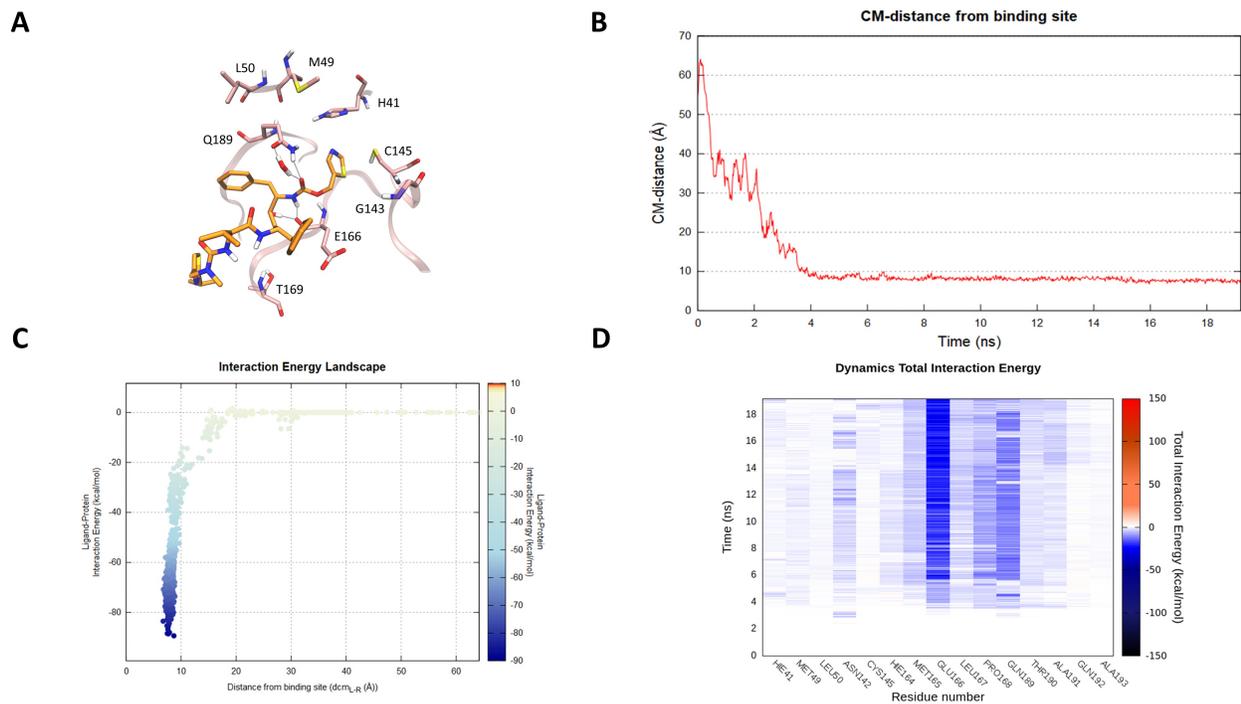


Figure 3

This panel summarizes the recognition pathway of Ritonavir against the SARS-CoV-2 main protease. (A) Ritonavir conformation sampled in the last frame of the SuMD trajectory (orange-colored molecule). The residues surrounding the binding site are reported in pink color. (B) Distance between the ligand center of mass (Cm) and the catalytic binding site of the Mpro during the SuMD simulation. (C) Interaction Energy Landscape describing the protein-ligand recognition process; values are arranged according to the distances between ligand and protein target mass centers. (D) Dynamic total interaction energy (electrostatic + vdW) computed for most contacted M pro residues.

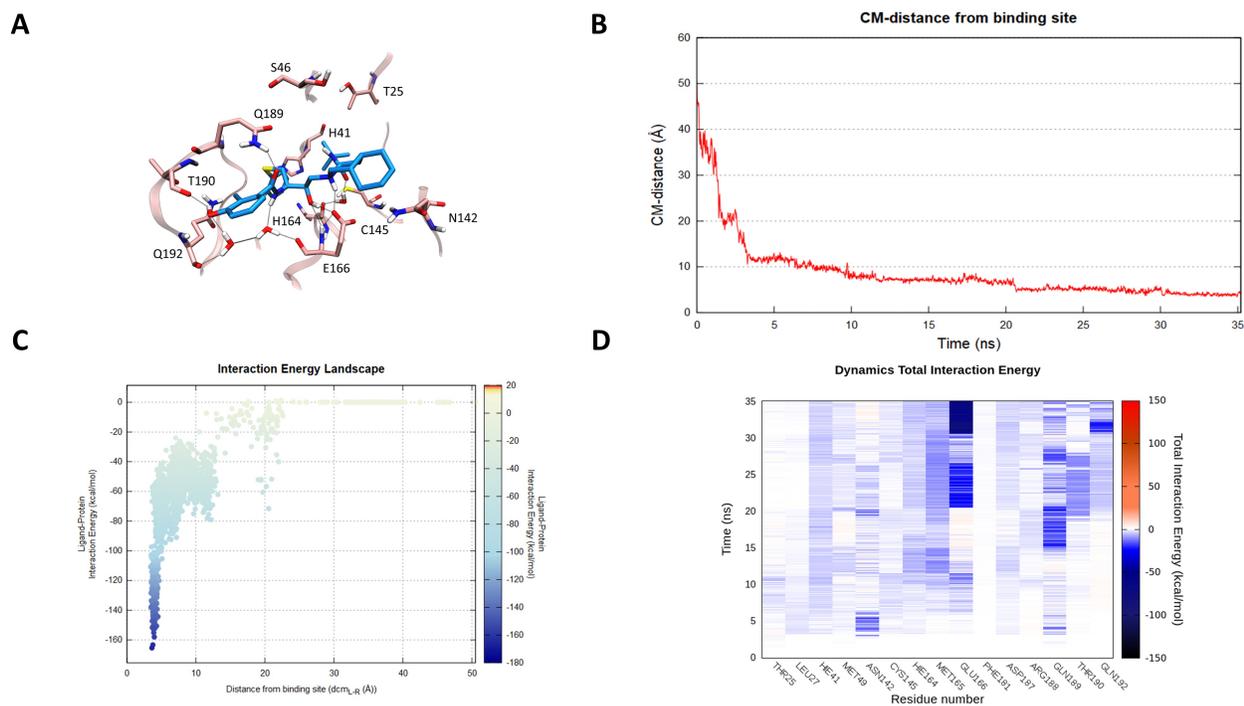


Figure 4

This panel summarizes the recognition pathway of Nelfinavir against the SARS-CoV-2 main protease. (A) Nelfinavir conformation sampled in the last frame of the SuMD trajectory (cyan-colored molecule). The residues surrounding the binding site are reported in pink color. (B) Distance between the ligand center of mass (Cm) and the catalytic binding site of the M pro during the SuMD simulation. (C) Interaction Energy Landscape describing the protein-ligand recognition process; values are arranged according to the distances between ligand and protein target mass centers. (D) Dynamic total interaction energy (electrostatic + vdW) computed for most contacted M pro residues.

Supplementary Files

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

- Video_1.mp4
- Video_2.mp4
- Video_3.mp4
- Video_3.mp4