Microscopic factors modulating the interactions between the SARS-Cov-2 main protease and $\alpha-$ketoamide inhibitors

Luigi Genovese,[†] William Dawson,[‡] Takahito Nakajima,[‡] Viviana Cristiglio,[¶] Valérie Vallet,[§] and Michel Masella[*,‖]

†*Univ. Grenoble Alpes, CEA, IRIG-MEM-L_Sim, 38000 Grenoble, France*

‡*RIKEN Center for Computational Science, Kobe, Japan*

¶*Institut Laue Langevin, 71 Av. des Martyrs, 38000 Grenoble France*

§*Université de Lille, CNRS, UMR 8523 - PhLAM - Physique des Lasers Atomes et Molécules, F-59000 Lille, France*

‖*Laboratoire de Biologie Structurale et Radiobiologie, Service de Bioénergétique, Biologie Structurale et Mécanismes, Institut Joliot, CEA Saclay, F-91191 Gif sur Yvette Cedex, France*

E-mail: michel.masella@cea.fr

Abstract

We performed 10 ns scale molecular dynamics simulations of 6 SARS-Cov-2 main protease/$\alpha-$ketoamide inhibitor complexes in aqueous solution, in the phase before the inhibitor covalently binds to the protease's catalytic cysteine, using a polarizable multi-scale molecular modeling approach. For each simulation, 100 $\mathbf{M^{pro}}$/inhibitor snapshots (about 4 800 atoms) were extracted along the last 2 ns simulation segments. They were post processed using a fully quantum mechanical $O(N)$ approach to decompose the protease in sets of fragments from which we computed the mean local interaction energies between the inhibitors and the different pockets of the protease catalytic domain. Contrary to earlier results, our analysis shows that the protease pocket S2 to be a key anchoring site able to lock within the catalytic domain an $\alpha-$ketoamide inhibitor even before covalent bonding to the protease catalytic cysteine occurs. To target that pocket our computations suggest to consider hydrophobic groups, like cyclo-propyl or cyclo-hexyl.

Several new drugs targeting the SARS-Cov-2 main viral protease $\mathbf{M^{pro}}$ (also called $\mathbf{3CL^{pro}}$) have been shown to be promising (or promising lead) compounds to develop new antiviral treatments for COVID-19.[1–5] All of these inhibitors have been built from standard docking approaches by inferring the microscopic factors modulating the strength of their interaction with $\mathbf{M^{pro}}$ from experimental structural data and then selected from experimental trial-and-error approaches. To speed up the optimization process or the development of new $\mathbf{M^{pro}}$ inhibitors, a pivotal information is the knowledge of reliable data that accurately quantify the strength of the microscopic interactions at the origin of the stability of $\mathbf{M^{pro}}$/inhibitor complexes.

Despite being heavily computationally demanding, Quantum chemical Methods, QM, are the most reliable theoretical approaches to investigate microscopic systems. Recently one of us proposed an efficient quantum $O(N)$ method based on Daubechies wavelets[6,7] allowing one to investigate molecular systems comprising thousands of atoms on modern supercomputing systems. An interesting feature of that approach is its ability to decompose a molecular system into fragments from which one may compute a map summarizing the microscopic local interactions occurring within a molecular complex.[8,9] In the present study, we combine such a QM approach with a simulation stage based on a multi-scale polarizable Molecular Modeling, MM, one[10,11] to investigate the Potential Energy Surface, PES, of $\mathbf{M^{pro}}$/inhibitor complexes from Molecular Dynamics, MD, simulations in the aqueous phase. A set of complex structures are extracted from the MD trajectories and post processed using the QM approach both to assess the reliability of the MM approach and to compute mean $\mathbf{M^{pro}}$/inhibitor complex local interaction maps that account for structural fluctuations.

Recently, theoretical studies have appeared investigating the stability of $\mathbf{M^{pro}}$ complexes with marketed drugs[12,13] (from chloroquine to curcumin and including the new peptide-like inhibitor "N3"[2]), with marine natural product putative inhibitors[14] or with hydroyethylamine analogs[15] using standard pairwise force field-based MD approaches or the QM-based Fragment Molecular Orbital scheme (applied to the $\mathbf{M^{pro}}$/N3 complex X-ray crystallographic structure[16]). Instead we employ here our sequential MM – QM (not to be confused with traditional QM/MM methods) scheme to investigate the interaction of $\mathbf{M^{pro}}$ with four strong peptidomimetic $\alpha-$ketoamide inhibitors recently synthesized by the Hilgenfelds team[1] (*i.e.* inhibitors **13a**, **13b**, **14b** and **11r**) as well as two such inhibitors characterized by a weak or no inhibitory potency for the main proteases of a large set of coronaviruses[17] (inhibitors **11p** and **11f**, see Figure 1). Note that the lack of inhibitory potency does not prejudge of the ability of an "inhibitor" to interact with the $\mathbf{M^{pro}}$ catalytic site or to form a stable complex.[17] In addition to being unlikely toxic, $\alpha-$ketoamide inhibitors are able to form a covalent bond with the $\mathbf{M^{pro}}$ catalytic cysteine residue (based on a reversible mechanism[18]) that stabilizes $\mathbf{M^{pro}}$/inhibitor complexes. However we focus our study to complexes corresponding to the pre-reactive state before the formation of a covalent bond between $\mathbf{M^{pro}}$ and the inhibitors, a state that is also pivotal to understanding the stability of $\mathbf{M^{pro}}$/inhibitor complexes.[19,20] Together with earlier theoretical studies, our present data will allow us to promote a more global understanding of the microscopic factors modulating the interaction of $\mathbf{M^{pro}}$ with putative inhibitors.

$\mathbf{M^{pro}}$/inhibitor complex starting structures are built from the X-ray PDB structure 6Y7M[21] corresponding to the Hilgenfelds inhibitor **13a**[1] by manually docking the other inhibitors to best fit the **13a** conformation. The $\mathbf{M^{pro}}$ catalytic pair $\mathrm{His}^{41}/\mathrm{Cys}^{145}$ is considered in its ionic form (its standard protonation state in $\mathbf{M^{pro}}$ as shown by a joint neutron/X-ray study[22] - see Supplementary Information): their side chains are an imidazolium and a thiolate group, respectively. The inhibitors are not covalently bonded to Cys^{145}. For each complex, we performed 10 independent NPT MD simulations (corresponding to different starting velocity sets) in the aqueous phase at the 10 ns scale using the code POLARIS(MD).[23] Protease dimerization is needed for catalytic activity as local dimer interactions stabilizes the catalytic pocket shape.[24] We consider here a single $\mathbf{M^{pro}}$ monomer and we account for the presence of a second unit by harmonically restraining the positions of the $\mathbf{M^{pro}}$ backbone carbons C_α that are more than 5 Å from any inhibitor atom.

$\mathbf{M^{pro}}$/inhibitors complexes are modeled using an updated version of the polarizable all atoms force field TCPEp.[25] Besides standard additive potentials like Coulombic and dispersion energy terms, TCPEp also includes a many-body polarization term (based on an induced dipole moment approach including short-range damping effects) and many-body anisotropic terms to model hydrogen bond networks. The TCPEp parameters are assigned to reproduce high-end quantum *ab initio* computations regarding a training set of molecular clusters (see Ref. 26 for instance). Water is simulated using an updated version of the coarse-grained Polarizable Pseudo Particle, PPP, approach[10,11] that improves ion hydration modeling. Protease/inhibitor complexes are embedded in rectangular boxes comprising about 57 000 PPP particles. The force field and the accuracy of the coupled TCPEp/PPP approach to model hydrated proteins and $\alpha-$ketoamide inhibitors are discussed in the Supporting Information.

We also computed local $\mathbf{M^{pro}}$/inhibitors Potentials of Mean Force, PMFs, in aqueous phase corresponding to the distance r between (1) the carbon atom connecting the ketoamide moiety to the inhibitor main chain and (2) the $\mathbf{M^{pro}}$ His^{164} backbone carbon C using standard Umbrella Sampling techniques coupled to our MD protocol. The r distance was scanned from 4 to 8 Å: within that distance domain we assume our MD protocol to provide a sampling of the local $\mathbf{M^{pro}}$/inhibitor PES that is sufficiently accurate.

Along the last 2 ns MD segments, we extracted 100 $\mathbf{M^{pro}}$/inhibitor regularly spaced snapshots (each comprising about 4 800 atoms) that were further investigated using a full QM $O(N)$ approach based on Density Functional Theory with the Perdew-Burke-Erzerhof (PBE) functional implemented by Daubechies wavelets formalism.[6,7] On modern supercomputing systems the computation of a $\mathbf{M^{pro}}$/inhibitor complex PBE single energy point is achieved within less than 2 hours using 1 024 computational cores. Note that the PBE energies discussed below (unless otherwise stated) have been corrected by including D3 dispersion terms.[27] The localized basis functions used in that QM scheme allow one to readily gather system atoms into "fragments" (f) and to approximate the system density matrix F as a sum of fragment density matrices F^f up to a desired level of accuracy measured by the "fragment purity" index $\Pi^f = \mathbf{Tr}\left((F^f)^2 - F^f\right)$.[8,9] A fragmentation is physically meaningful when all the $|\Pi^f|$s are small, typically about 5 % as shown by an earlier study[9] and as set here. Such a fragmentation can be defined common to all the snapshots of an entire MD trajectory. A quantum $\mathbf{M^{pro}}$/inhibitor interaction map may be thus readily drawn from the quantum energies $\delta\bar{U}^{fi}$ measuring the magnitude of the interactions between the $\mathbf{M^{pro}}$ fragments and the inhibitor (taken as a single fragment i) averaged over our MD simulations to account for $\mathbf{M^{pro}}$/inhibitor complex structural fluctuations. Note that, the $\delta\bar{U}^{fi}$s computed in the present study do not include the D3 dispersion correction, they measure the strength of the local $\mathbf{M^{pro}}$/inhibitor repulsion-exchange and electrostatic interactions. They have thus to be considered to compare $\mathbf{M^{pro}}$/inhibitor interaction patterns among them and not to discuss the global strength of $\mathbf{M^{pro}}$ fragment/inhibitor interactions.

Along all independent MD simulations, inhibitors **11r**, **13a**, **13b** and **14b** interact strongly with the $\mathbf{M^{pro}}$ catalytic pair His^{41}/Cys^{145}, see Figure 1. Precisely the His^{41} imidazolium is hydrogen bonded to one of the ketoamide oxygen and the Cys^{145} sulfur strongly interacts with both the ketoamide sp^2 carbons (the corresponding mean distances are about 2.8 Å). The mean $\mathbf{M^{pro}}$/inhibitor distance r defined above is about 5.3 Å and these inhibitors are also hydrogen bonded to residues His^{163}, His^{165}, Glu^{166} and Asn^{189} backbones or side chains along all simulations in agreement with experiments.[1] However large structural fluctuations of the inhibitor side chains can be observed, in particular the Boc group within the $\mathbf{M^{pro}}$ S4 pocket.

The above "standard" $\mathbf{M^{pro}}$/inhibitor interaction pattern is also observed along 6 $\mathbf{M^{pro}}$/**11p** simulations (we recall that **11p** is the non-inhibitory substrate). Along them the mean dihedral angle $\bar{\psi}^{\mathbf{11p}} = \langle \mathrm{N-C-C-C} \rangle$ corresponding to atoms connecting the acetonitryl moiety to the inhibitor main chain is about 180°. That moiety does not reside within pocket S2 and it is not hydrogen bonded to any $\mathbf{M^{pro}}$ residue. Along 3 other simulations, this interaction pattern is strongly altered: $\bar{\psi}^{\mathbf{11p}} \approx -60°$, $r \approx 7.2$ Å and the acetonitryl moiety establishes hydrogen bonds with the Tyr^{54} hydroxyl group and/or with the Gln^{189} backbone.

Inhibitor **11f** maintains a "standard" interaction pattern only along 3 simulations: the **11f** Boc group resides then at the 'entrance' of pocket S2 (it is not as deeply buried in that pocket as the cyclo-hexyl and cyclo-propyl groups of inhibitors **13a** and **13b**, see Supporting Information), in agreement with experiments regarding the SARS-Cov main protease.[17] Along all the other **11p/11f** simulations, the inhibitor leaves the catalytic site.

The propensity of inhibitors to maintain a standard interaction pattern may be assessed from our PMF(r) profiles, see Figure 2. The four strong inhibitors PMFs are close: they present a single minimum at $r \approx 5.3$ Å and they increase then by up to 5 kcal mol^{-1} at $r \approx 7.5$ Å. The PMF of **11f** also presents a minimum at $r \approx 5.0$ Å but it increases until it reaches a weak energy barrier of 2 kcal mol^{-1} at $r \approx 6.0$ Å explaining the propensity of **11f** to escape from the catalytic site along our simulations. We computed three PMF profiles for **11p** corresponding to the $\psi^{\mathbf{11p}}$ dihedral angle restrained harmonically to a value of 60 ($\mathrm{PMF}^{\mathbf{11p}}_{60}$), -60 ($\mathrm{PMF}^{\mathbf{11p}}_{-60}$) and 180° ($\mathrm{PMF}^{\mathbf{11p}}_{180}$), respectively. These dihedral angle values correspond to minimum locations on the acetronitryl dihedral energy profile discussed in the Supporting Information. $\mathrm{PMF}^{\mathbf{11p}}_{180^\circ}$ is close to the **11f** one with, however, with an even weaker energy barrier (about 1 kcal mol^{-1}) at about 6.0 Å, whereas the lowest minima of $\mathrm{PMF}^{\mathbf{11p}}_{60^\circ}$ and $\mathrm{PMF}^{\mathbf{11p}}_{-60^\circ}$ are located at 6.8-7.0 Å. While we have not exhaustively sampled the $\mathbf{M^{pro}}$/inhibitor PESs,[28] the computed PMFs support the inhibitor behaviors along our independent MD simulations.

We computed the MM and QM/PBE+D3 $\mathbf{M^{pro}}$/inhibitor mean interaction energies $\Delta\bar{U}$ from 100 $\mathbf{M^{pro}}$/inhibitor complex snapshots extracted along a single MD simulation for each strong inhibitor **11r**, **13a**, **13b** and **14b**, and along three and five simulations corresponding to different $\mathbf{M^{pro}}$/inhibitors interaction patterns for **11p** and **11f**, respectively. $\Delta\bar{U}$s are computed as the difference between the $\mathbf{M^{pro}}$/inhibitor complex energy and the energies of $\mathbf{M^{pro}}$ and of the inhibitor alone in their complex geometry. Both these sets of $\Delta\bar{U}$ values are linearly correlated. However the MM energies are more stable than the QM ones by about 20 %, see Figure 2. We have identified (see Supporting Information) that this discrepancy does not arise from differences in MM and PBE+D3 descriptions of $\mathbf{M^{pro}}$/inhibitor short-range interactions, but rather from a PBE under-polarization of the $\mathbf{M^{pro}}$ chemical bonds yielding a PBE electrostatic potential within the $\mathbf{M^{pro}}$ catalytic pocket weaker by 20 % as compared to MM. A similar PBE bond under-polarization compared to a polarizable force field approach was already observed for liquid water.[29] Lastly, D3 and MM dispersions represent from 40 to 80 % of the absolute $\Delta\bar{U}$ values.

MM and PBE+D3 $\Delta\bar{U}$ values of the four strong inhibitors differ at most by 10% but they are all smaller than the **11p** value corresponding to a simulation along which that inhibitor maintain a standard interaction pattern. The inhibitor potency is thus not dominated by the sole strength of the $\mathbf{M^{pro}}$/inhibitor interaction but it results from complex interaction competitions between the inhibitor, the $\mathbf{M^{pro}}$ enzyme and their chemical environment, as discussed in Ref. 1 and as suggested by our PMFs. Along the simulations where inhibitors **11p** and **11f** left the $\mathbf{M^{pro}}$ catalytic site, MM and PBE+D3 $\Delta\bar{U}$ values are twice to three times weaker compared to simulations where the inhibitors maintain a standard interaction pattern with $\mathbf{M^{pro}}$.

Focusing now on the snapshot sets corresponding to a standard $\mathbf{M^{pro}}$/inhibitor interaction pattern, *i.e.* the four strong inhibitors sets and those corresponding to simulations labeled 8 and 9 in Figure 2 for inhibitors **11f** and **11p**, respectively, our QM fragmentation yields temporally stable and almost identical $\mathbf{M^{pro}}$ fragment patterns. Among the about 200 fragments identified, about 20 of them (located at the inhibitor vicinity) interact noticeably with it. Most of these fragments correspond to a single residue at the remarkable exception of the fragment $\mathrm{Gly^{143}}$-$\mathrm{Ser^{144}}$-$\mathrm{Cys^{145}}$ that gathers the oxyanion residues and the catalytic cysteine. We assigned the fragments to the $\mathbf{M^{pro}}$ catalytic domain pockets from distance arguments, see Figure 3 where we also plot the mean energies $\Delta\bar{U}^{\mathrm{pocket}}_{\mathrm{inhi}}$ corresponding to the sum of the $\delta\bar{U}^{fi}$ energies running on the fragments belonging to a given pocket. For the four strong inhibitors and inhibitor **11f**, their $\Delta\bar{U}^{\mathrm{pocket}}_{\mathrm{inhi}}$ profiles are close: they interact the strongest with pocket S1 (which comprises the $\mathbf{M^{pro}}$ catalytic pair) and in a negligible way with pockets S3 and S4. The $\Delta\bar{U}^{\text{pocket}}_{\text{inhi}}$ profile of inhibitor **11p** differs noticeably compared to the latter ones: **11p** interacts the strongest with pocket S2 and noticeably with pocket S3. The dihedral angle $\bar{\psi}^{\mathbf{11p}}$ is about 180° along the **11p** simulation 9: the **11p** acetonitryl moiety is outside of pocket S2 and the inhibitor phenyl group initially residing within the pocket S4 core evolves to interact with the backbone CO groups of Val^{186}, Arg^{188} and Thr^{190} at the pockets S4/S2 boundary (see Supporting Information). The strongest interaction of **11p** with pocket S2 may thus be considered as an artifact but this does not lead to change the above conclusion about the difference in the $\Delta\bar{U}^{\text{pocket}}_{\text{inhi}}$ profiles. The energies $\Delta\bar{U}^{\text{pocket}}_{\text{inhi}}$ measure only the strength of the $\mathbf{M^{pro}}$/inhibitor local repulsion-exchange and electrostatic interactions. Because of the weight of dispersion in $\mathbf{M^{pro}}$/inhibitor interactions, a $\Delta\bar{U}^{\text{pocket}}_{\text{inhi}}$ profile close to those of the four strong inhibitors **13a** to **11r** does not correspond necessarily to a potent inhibitor, as for **11f**.

By employing a computational scheme which sequentially couples a multi-scale polarizable MM approach and a particularly efficient QM method, we have shown that the most promising $\alpha-$ketoamide inhibitors developed the Hilgenfeld's team target mainly three pockets of the $\mathbf{M^{pro}}$ catalytic domain, namely S1, S1 and S2. Contrary to the recent large-scale MD simulations study of Huynh et *al*,[12] both our PMFs (in particular those corresponding to **11p** and **11f**) and fragmentation computations unambiguously show the inhibitor potency not to be tied to the inhibitor capacity to strongly interact with the $\mathbf{M^{pro}}$ pocket S4 but with pocket S2. Note that besides simulation artifacts underlined by Huynh et *al* (like the accuracy of their scoring functions and of their standard pairwise force fields), the conclusion of the latter authors may arise from their set of investigated drugs that were not able to specifically target pocket S2.

Hence our computations strongly suggest $\mathbf{M^{pro}}$ pocket S2 to be a key anchoring site that is able to lock within the $\mathbf{M^{pro}}$ catalytic domain an $\alpha-$ketoamide inhibitor even before to be covalently bonded to the $\mathbf{M^{pro}}$ catalytic cysteine, which warrants the generality of that conclusion. We may also note here that other promising $\mathbf{M^{pro}}$ inhibitors like the carmofur one[3] and the inhibitor N3[2] and related[4] all target pocket S2 by means of hydrophobic groups like cyclo-propyl or cyclo-hexyl, and even using a fluoro-phenyl group as shown by Dai et *al.*[4] New inhibitors must target that pocket to provide thermodynamically stable $\mathbf{M^{pro}}$/inhibitor complexes, preferentially using the latter chemical groups.

Supplementary Material Available

This material is available free of charge via the Internet at http://pubs.acs.org. It provides more detailed discussions regarding the accuracy of the our multi scale MM approach to model proteins and $\alpha-$ketoamide inhibitors in aqueous phase as well as of both the MM and quantum PBE level of theory to describe these inhibitors in gas phase as compared to high end quantum CCSD(T) computations. Discussions regarding the PBE under-polarization of the enzyme covalent bonds are also discussed and the details of the $\mathbf{M^{pro}}$/inhibitors structures, quantum and MM interaction energies and fragmentation data (like the fragment assignment to $\mathbf{M^{pro}}$ pockets) are also provided. The final structures of the $\mathbf{M^{pro}}$/inhibitor complexes along all our MD simulations (in PDB format) are freely available at http://biodev.cea.fr/polaris/download.html.

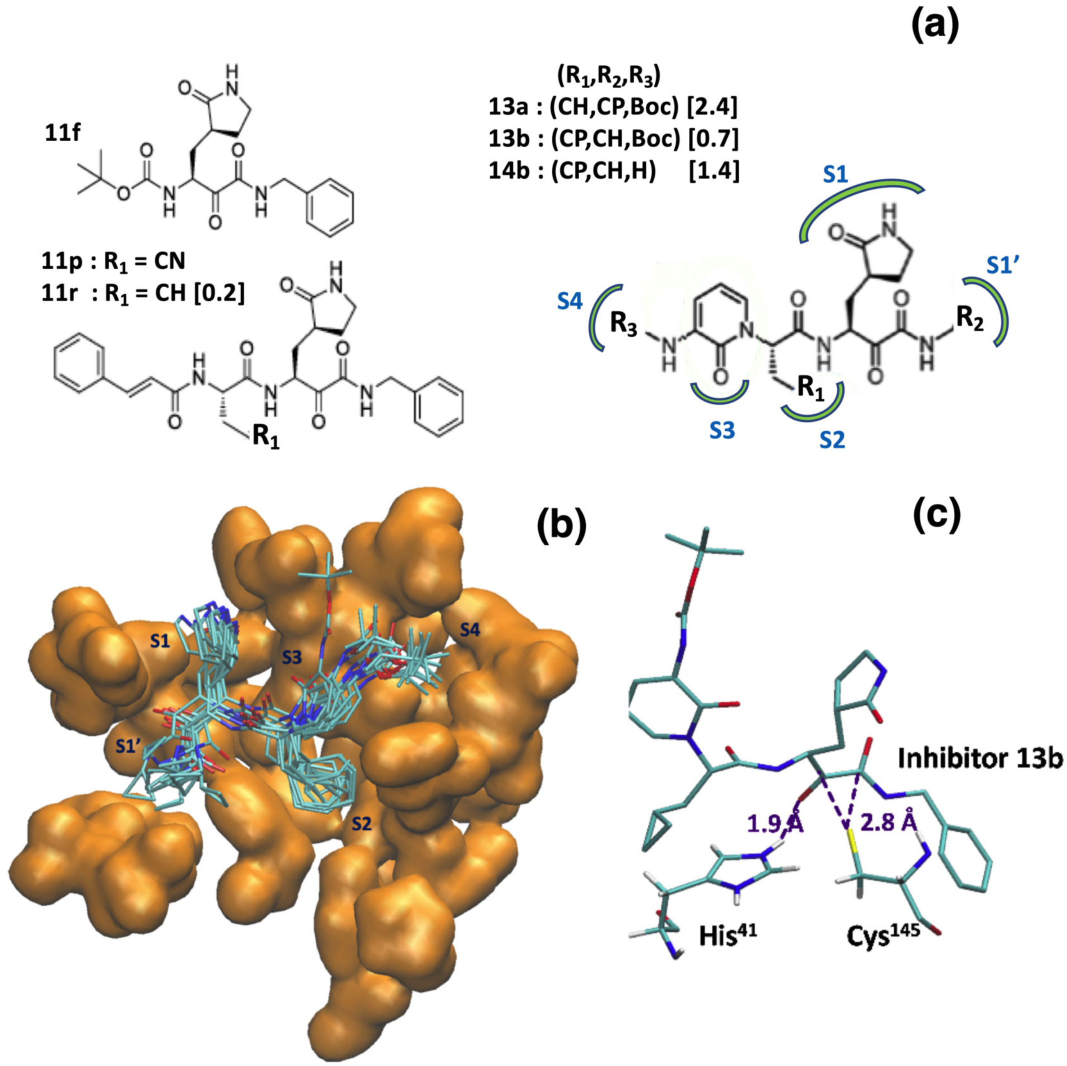


Figure 1: (a) Definition of the six $\alpha-$ketoamide inhibitors considered in the present study. CN, CP, CH and Boc are the acetonytril, cyclo propyl, cyclo hexyl and tert-butyloxycarbonyl groups, respectively. In the brackets the EC_{50} values in μM unit that measure the inhibitor potency for the SARS-Cov-2 $\mathbf{M^{pro}}$ main protease.[1] For inhibitors **13a**, **13b**, and **14b**, the interaction between their side chains and the $\mathbf{M^{pro}}$ pockets S1 to S4 (as defined in Refs.[4,17]) are shown. (b) Superimposition of the final inhibitor **13a** structures along the 10 independent MD simulations within the $\mathbf{M^{pro}}$ catalytic binding site. (c) Detail of the $\mathbf{M^{pro}}$/inhibitor **13b** final structure along one of the simulations. Dashed lines are the strong hydrogen bond between the His^{41} imidazolium and the inhibitor ketoamide group and the Cys^{145} sulfur/inhibitor ketoamide carbon sp^2 distances.

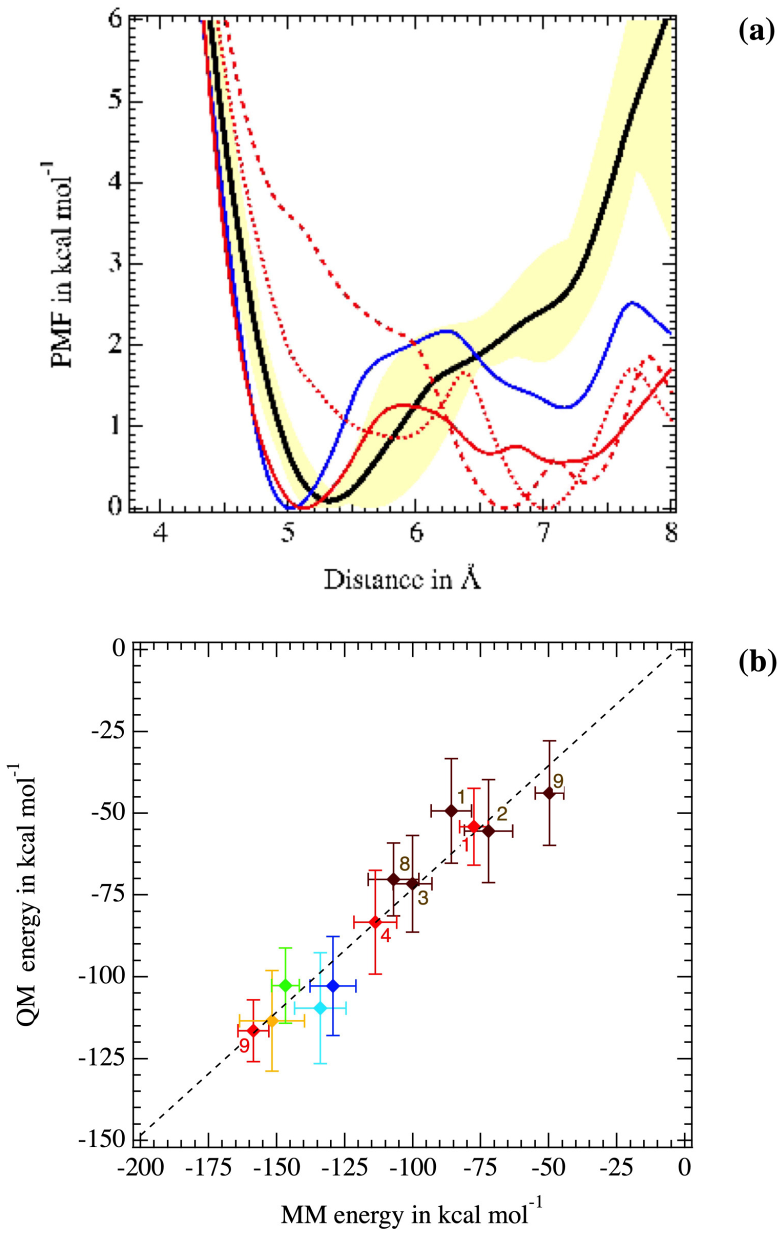
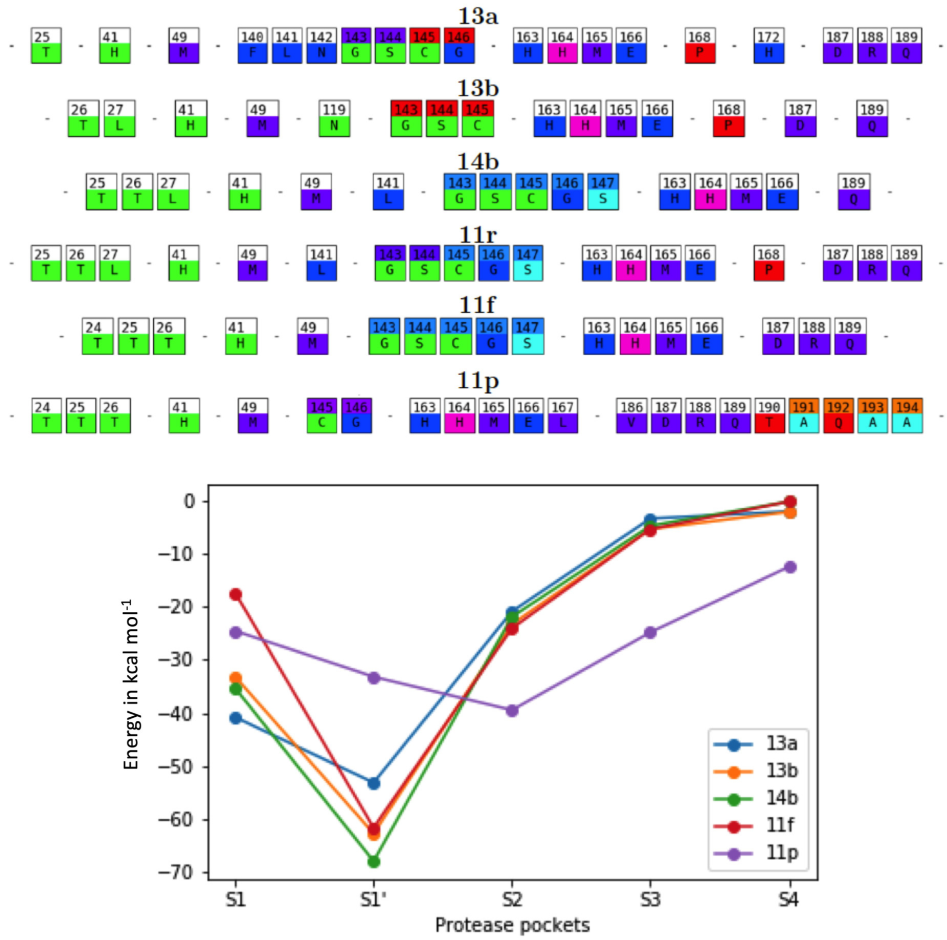


Figure 2: (a) Local PMF(r) profiles corresponding to the $\mathbf{M^{pro}}$/inhibitor complexes. Black line: mean PMF corresponding to inhibitors **11r**, **13a**, **13b** and **14b** (the yellow domain is defined by the minimum and maximum values of each single PMF). Blue: PMF of **11f**. Red: $\mathrm{PMF}^{\mathbf{11p}}_{180^\circ}$ (full line), $\mathrm{PMF}^{\mathbf{11p}}_{-60^\circ}$ (dashed line) and $\mathrm{PMF}^{\mathbf{11p}}_{60^\circ}$ (dotted line) corresponding to inhibitor **11p**, see text for definition. (b) Mean QM $\mathbf{M^{pro}}$/inhibitor interaction energies $\Delta\bar{U}$ as a function of their MM counter parts (brown **11f**, red **11fp**, orange **13a**, green **11r**, light blue **13b**, dark blue **14b**). The error bars correspond the standard deviations of these mean energy values. For **11f** and **11p** data, the numbers shown corresponds to the simulation labels (see Supporting Information).



(a)

(b)

Figure 3: (a) $\mathbf{M^{pro}}$ fragments interacting noticeably with the inhibitors (*i.e.* their corresponding fragment/inhibitor energy $\delta \bar{U}^{fi}$ is larger than $\mathrm{k_B T} = 0.6$ kcal mol^{-1}). Each box corresponds to a $\mathbf{M^{pro}}$ residue, the box colors show, in the upper part, the regroupment in the QM fragments. Amino acids with similar colors belong to the same fragment. A white top color indicates that the amino acid itself constitutes a good fragment. The bottom colors indicate to which $\mathbf{M^{pro}}$ catalytic pocket the corresponding amino-acid has been (arbitrarily) assigned. The pocket S1, S1', S2, S3 and S4 are colored, respectively, in (dark) blue, green, purple, pink and red. Because of the common residues delimiting the different pockets, we define pocket S3 from the single residue His^{164}. The cyan aminoacids have not been assigned to a pocket. (b) $\Delta \bar{U}^{\mathrm{pocket}}_{\mathrm{inhi}}$ inhibitor/protease pocket energies for the MD snapshot sets where the inhibitors maintain a standard interaction pattern. The corresponding standard deviations represent about 20 % of the strong mean $\Delta \bar{U}^{\mathrm{pocket}}_{\mathrm{inhi}}$ values up to 80 % of the weak ones. We can observe that the **11p** inhibitor has an outstanding pattern both on the fragmentation and on the pocket interactions. The strong $\Delta \bar{U}^{\mathrm{S3}}_{\mathrm{inhi}}$ value reported here for inhibitor **11p** arises from a hydrogen bond network between that inhibitor, the backbone CO and side chain imidazole group of His^{164}, and the amide side chain group of Gln^{189} (see Supporting Information).

Acknowledgments

M.M. thanks Emmanuel Oseret (Exascale Computing Research Laboratory, a joint INTEL/CEA/UVSQ/GENCI laboratory) for his help in optimizing the code POLARIS(MD). This work was granted access to the TGCC HPC resources under the allocation 2019-2020 [A0070307078] and the Grand Challenge allocation [GC0429] made by GENCI. L.G. and W.D. also gratefully acknowledge the joint CEA-RIKEN collaboration action. We also acknowledge support from the MaX EU Center of Excellence. We are also grateful to the financial support by the Labex CaPPA (contract ANR-11-LABX-0005-01) and I-SITE ULNE project OVERSEE (contract ANR-16-IDEX-0004), and by the CPER CLIMIBIO (European Regional Development Fund, Hauts de France council, French Ministry of Higher Education and Research).

References

(1) Zhang, L.; Lin, D.; Sun, X.; Curth, U.; Drosten, C.; Sauerhering, L.; Becker, S.; Rox, K.; Hilgenfeld, R. Crystal Structure of SARS-CoV-2 Main Protease Provides a Basis for Design of Improved α-Ketoamide Inhibitors. *Science* **2020**, *368*, 409–412, DOI: 10.1126/science.abb3405.

(2) Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C. et al. Structure of Mpro from SARS-CoV-2 and Discovery of its Inhibitors. *Nature* **2020**, *582*, 289–293, DOI: 10.1038/s41586-020-2223-y.

(3) Jin, Z.; Zhao, Y.; Sun, Y.; Zhang, B.; Wang, H.; Wu, Y.; Zhu, Y.; Zhu, C.; Hu, T.; Du, X. et al. Structural Basis for the Inhibition of SARS-CoV-2 Main Protease by Antineoplastic Drug Carmofur. *Nat. Struct. Mol. Biol.* **2020**, *27*, 529–532, DOI: 10.1038/s41594-020-0440-6.

(4) Dai, W.; Zhang, B.; Jiang, X.-M.; Su, H.; Li, J.; Zhao, Y.; Xie, X.; Jin, Z.; Peng, J.; Liu, F. et al. Structure-Based Design of Antiviral Drug Candidates Targeting the SARS-CoV-2 Main Protease. *Science* **2020**, *368*, 1331–1335, DOI: 10.1126/science.abb4489.

(5) VanPatten, S.; He, M.; Altiti, A.; F Cheng, K.; Ghanem, M. H.; Al-Abed, Y. Evidence Supporting the Use of Peptides and Peptidomimetics as Potential SARS-CoV-2 (COVID-19) Therapeutics. *Future Med. Chem.* **0**, *0*, null, DOI: 10.4155/fmc-2020-0180.

(6) Ratcliff, L. E.; Dawson, W.; Fisicaro, G.; Caliste, D.; Mohr, S.; Degomme, A.; Videau, B.; Cristiglio, V.; Stella, M.; D 'Alessandro, M. et al. Flexibilities of Wavelets as a Computational Basis Set for Large-Scale Electronic Structure Calculations. *J. Chem. Phys.* **2020**, *152*, 194110, DOI: 10.1063/5.0004792.

(7) Genovese, L.; Neelov, A.; Goedecker, S.; Deutsch, T.; Ghasemi, S. A.; Willand, A.; Caliste, D.; Zilberberg, O.; Rayson, M.; Bergman, A. et al. Daubechies Wavelets as a Basis Set for Density Functional Pseudopotential Calculations. *J. Chem. Phys.* **2008**, *129*, 014109, DOI: 10.1063/1.2949547.

(8) Dawson, W.; Mohr, S.; Ratcliff, L. E.; Nakajima, T.; Genovese, L. Complexity Reduction in Density Functional Theory Calculations of Large Systems: System Partitioning and Fragment Embedding. *J. Chem. Theory Comput.* **2020**, *16*, 2952–2964, DOI: 10.1021/acs.jctc.9b01152.

(9) Mohr, S.; Masella, M.; Ratcliff, L. E.; Genovese, L. Complexity Reduction in Large Quantum Systems: Fragment Identification and Population Analysis via a Local Optimized Minimal Basis. *J. Chem. Theory Comput.* **2017**, *13*, 4079–4088, DOI: 10.1021/acs.jctc.7b00291.

(10) Masella, M.; Borgis, D.; Cuniasse, P. Combining a Polarizable Force-Field and a Coarse-Grained Polarizable Solvent Model: Application to Long Dynamics Simulations of Bovine Pancreatic Trypsin Inhibitor. *J. Comput. Chem.* **2008**, *29*, 1707–1724, DOI: 10.1002/jcc.20932.

(11) Masella, M.; Borgis, D.; Cuniasse, P. Combining a Polarizable Force-Field and a Coarse-Grained Polarizable Solvent Model. II. Accounting for Hydrophobic Effects. *J. Comput. Chem.* **2011**, *32*, 2664–2678, DOI: 10.1002/jcc.21846.

(12) Huynh, T.; Wang, H.; Luan, B. In Silico Exploration of the Molecular Mechanism of Clinically Oriented Drugs for Possibly Inhibiting SARS-CoV-2's Main Protease. *J. Phys. Chem. Lett.* **2020**, *11*, 4413–4420, DOI: 10.1021/acs.jpclett.0c00994.

(13) Khan, S. A.; Zia, K.; Ashraf, S.; Uddin, R.; Ul-Haq, Z. Identification of Chymotrypsin-Like Protease Inhibitors of SARS-CoV-2 via Integrated Computational Approach. *J. Biomol. Struct. Dyn.* **2020**, *0*, 1–10, DOI: 10.1080/07391102.2020.1751298.

(14) Gentile, D.; Patamia, V.; Scala, A.; Sciortino, M.; Piperno, A.; Rescifina, A. Putative Inhibitors of SARS-CoV-2 Main Protease from A Library of Marine Natural Products: A Virtual Screening and Molecular Modeling Study. *Mar. Drugs* **2020**, *18*, 225, DOI: 10.3390/md18040225.

(15) Kumar, S.; Sharma, P. P.; Shankar, U.; Kumar, D.; Joshi, S. K.; Pena, L.; Durvasula, R.; Kumar, A.; Kempaiah, P.; Poonam, et al. Discovery of New Hydroxyethylamine Analogs against 3CLpro Protein Target of SARS-CoV-2: Molecular Docking, Molecular Dynamics Simulation, and Structure-Activity Relationship Studies. *J. Chem. Inf. Model.* **0**, *0*, null, DOI: 10.1021/acs.jcim.0c00326.

(16) Hatada, R.; Okuwaki, K.; Mochizuki, Y.; Handa, Y.; Fukuzawa, K.; Komeiji, Y.; Okiyama, Y.; Tanaka, S. Fragment Molecular Orbital Based Interaction Analyses on COVID-19 Main Protease - Inhibitor N3 Complex (PDB ID: 6LU7). *J. Chem. Inf. Model.* **2020**, *60*, 3593–3602, DOI: 10.1021/acs.jcim.0c00283.

(17) Zhang, L.; Lin, D.; Kusov, Y.; Nian, Y.; Ma, Q.; Wang, J.; von Brunn, A.; Leyssen, P.; Lanko, K.; Neyts, J. et al. α-Ketoamides as Broad-Spectrum Inhibitors of Coronavirus and Enterovirus Replication: Structure-Based Design, Synthesis, and Activity Assessment. *J. Med. Chem.* **2020**, *63*, 4562–4578, DOI: 10.1021/acs.jmedchem.9b01828.

(18) Tichá, A.; Stanchev, S.; Vinothkumar, K. R.; Mikles, D. C.; Pachl, P.; Began, J.; Škerle, J.; Švehlova, K.; Nguyen, M.; Verhelst, S. et al. General and Modular Strategy for Designing Potent, Selective, and Pharmacologically Compliant Inhibitors of Rhomboid Proteases. *Cell Chem. Biol.* **2017**, *24*, 1523 – 1536.e4, DOI: 10.1016/j.chembiol.2017.09.007.

(19) Świderek, K.; Moliner, V. Revealing the Molecular Mechanisms of Proteolysis of SARS-CoV-2 Mpro by QM/MM Computational Methods. *Chem. Sci.* **2020**, –, DOI: 10.1039/D0SC02823A, Advance Article.

(20) Pillaiyar, T.; Manickam, M.; Namasivayam, V.; Hayashi, Y.; Jung, S.-H. An Overview of Severe Acute Respiratory Syndrome–Coronavirus (SARS-CoV) 3CL Protease Inhibitors: Peptidomimetics and Small Molecule Chemotherapy. *J. Med. Chem.* **2016**, *59*, 6595–6628, DOI: 10.1021/acs.jmedchem.5b01461.

(21) Zhang, L.; Lin, D.; Hilgenfeld, R. Crystal structure of the complex resulting from the reaction between the SARS-CoV main protease and tert-butyl (1-((S)-3-cyclohexyl-1-(((S)-4-(cyclopropylamino)-3,4-dioxo-1-((S)-2-oxopyrrolidin-3-yl)butan-2-yl)amino)-1-oxopropan-2-yl)-2-oxo-1,2-dihydropyridin-3-yl)carbamate. DOI: 10.2210/pdb6Y7M/pdb.

(22) Kovalevsky, A.; Kneller, D.; Coates, L. Joint neutron/X-ray structure of SARS-CoV-2 3CL Mpro at room temperature. DOI:10.2210/pdb7JUN/pdb.

(23) http://biodev.cea.fr/polaris/.

(24) Anand, K.; Palm, G. J.; Mesters, J. R.; Siddell, S. G.; Ziebuhr, J.; Hilgenfeld, R. Structure of Coronavirus Main Proteinase Reveals Combination of a Chymotrypsin Fold with an extra $\alpha-$Helical Domain. *EMBO J.* **2002**, *21*, 3213–3224, DOI: 10.1093/emboj/cdf327.

(25) Masella, M.; Cuniasse, P. A Many-Body Model to Study Proteins. I. Applications to MLnm+ Complexes, M^{m+}=Li+, Na+, K+, Mg^{2+}, Ca^{2+}, and Zn^{2+}, L=H2O, CH3OH, HCONH2, n=1-6, and to Small Hydrogen Bonded Systems. *J. Chem. Phys.* **2003**, *119*, 1866–1873, DOI: 10.1063/1.1579478.

(26) Houriez, C.; Meot-Ner (Mautner), M.; Masella, M. Solvation of the Guanidinium Ion in Pure Aqueous Environments: A Theoretical Study from an Ab Initio-Based Polarizable Force Field. *J. Phys. Chem. B* **2017**, *121*, 11219–11228, DOI: 10.1021/acs.jpcb.7b07874.

(27) Schröder, H.; Creon, A.; Schwabe, T. Reformulation of the D3(Becke-Johnson) Dispersion Correction without Resorting to Higher than C6 Dispersion Coefficients. *J. Chem. Theory Comput.* **2015**, *11*, 3163–3170, DOI: 10.1021/acs.jctc.5b00400.

(28) You, W.; Tang, Z.; Chang, C.-e. A. Potential Mean Force from Umbrella Sampling Simulations: What Can We Learn and What Is Missed? *J. Chem. Theory Comput.* **2019**, *15*, 2433–2443, DOI: 10.1021/acs.jctc.8b01142.

(29) Ratcliff, L. E.; Mohr, S.; Huhs, G.; Deutsch, T.; Masella, M.; Genovese, L. Challenges in Large Scale Quantum Mechanical Calculations. *WIREs Computational Molecular Science* **2017**, *7*, e1290, DOI: 10.1002/wcms.1290.