QM/MM Simulations of the Covalent Inhibition of the SARS-CoV-2 Main Protease: Four Compounds and Three Reaction Mechanisms

Bella L. Grigorenko [a,b], Igor V. Polyakov [a,b], Maria G. Khrenova [a,c], Goran Giudetti [d], Shirin Faraji [e], Anna I. Krylov [d,*], and Alexander V. Nemukhin [a,b,*]

[a] Department of Chemistry, Lomonosov Moscow State University, Moscow, 119991, Russia

[b] Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, Moscow, 119334, Russia

[c] Bach Institute of Biochemistry, Federal Research Centre "Fundamentals of Biotechnology", Russian Academy of Sciences, Moscow, 119071, Russia

[d] Department of Chemistry, University of Southern California, Los Angeles, California, 90089-0482, United States

[e] Zernike Institute for Advanced Materials, University of Groningen, Groningen, 9747 AG, The Netherlands

* Corresponding authors: krylov@usc.edu and anemukhin@yahoo.com

ABSTRACT: The search for efficient inhibitors of the SARS-CoV-2 enzymes is ongoing due to the continuing COVID-19 pandemic. We report the results of computational modeling of the reactions of the SARS-CoV-2 main protease (M^{Pro}) with four potential covalent inhibitors. Two of them, carmofur and nirmatrelvir, have been shown experimentally the ability to inhibit M^{Pro}. Two other compounds, X77A and X77C, were designed computationally in this work, derived from the structure of X77, a non-covalent inhibitor forming a tight surface complex with M^{Pro}. We modified the X77 structure by introducing warheads capable of efficient chemical reactions with the catalytic cysteine residue in the M^{Pro} active site. The reaction mechanisms of the four molecules with M^{Pro} were investigated by quantum mechanics/molecular mechanics (QM/MM) calculations using large quantum subsystems. First, at the QM/MM level, we optimized structures of stationary points on the potential energy surfaces corresponding to the reactants, products, intermediates, and transition states along the hypothesized reaction coordinates. Analysis of these structures has informed the selection of collective variables for the subsequent calculations of the Gibbs energy profiles using molecular dynamics simulations with QM/MM potentials (QM/MM MD). In these simulations, the QM part was treated by DFT with the PBE0 functional. The results show that all four compounds form covalent adducts with the catalytic cysteine Cys 145 of M^{Pro}. From the chemical perspective, the reactions of these four compounds with M^{Pro} follow three distinct mechanisms. In all cases, the reaction is initiated by a nucleophilic attack of the thiolate group of the deprotonated cysteine residue from the catalytic dyad Cys145-His41 of M^{Pro}. In the case of carmofur and X77A, the covalent binding of the thiolate to the ligand is accompanied by the formation of the fluoro-uracil leaving group. The reaction with X77C follows the nucleophilic aromatic substitution S_NAr mechanism. The reaction of M^{Pro} with nirmatrelvir, which has a reactive nitrile group, leads to the formation of the covalent thioimidate adduct with the thiolate of the Cys145 residue in the enzyme active site.

INTRODUCTION

The quantum mechanics/molecular mechanics (QM/MM) methods are indispensable tools for modeling biochemical reactions in complex environment.[1–9] QM/MM-based calculations enable construction of potential energy and free energy profiles of enzyme-catalyzed reactions and reactions of the covalent inhibition of enzymes. The latter are of particular interest due to the COVID-19 pandemics, which stimulated massive efforts, including computer simulations, aiming to reveal molecular-level mechanisms of the action of SARS-CoV-2 enzymes and to design efficient non-covalent and covalent inhibitors to inactivate target enzymes.[10–13]

This work contributes to this effort by modeling reactions of four compounds with the critical SARS-CoV-2 enzyme, the main protease (M^{Pro}), also known as the 3 chymotrypsin-like protease ($3CL^{Pro}$).[14,15] This enzyme, encoded by the viral genome, plays an important role in cleaving viral polyproteins into functional proteins. Thus, inhibiting this enzyme [14–21] blocks viral replication, making M^{Pro} an attractive drug target.

QM/MM-based computer simulations provide insights into cysteine protease reaction mechanisms and can be used to predict novel compounds as prospective drugs.[22-35] Numerous studies investigated irreversible (or covalent) inhibitors of cysteine proteases (a class to which M^{Pro} belongs).[12,28-45] The list of prospective inhibitors is growing, but mechanisms of their interaction with the enzyme are not yet fully elucidated. In this work, we consider two compounds, carmofur[21,46,47] and nirmatrelvir,[48,49] which have already been identified as the irreversible inhibitors of M^{Pro}. We also introduce two novel compounds, called X77A and X77C. We designed these molecules computationally, starting from the structure of X77, a potent noncovalent inhibitor[50,51] of M^{Pro}. X77 is capable of forming a tight surface complex with SARS-CoV-2 M^{Pro}, whose structure has been deposited in the Protein Data Bank[52] (PDB ID: 6W63).

Fig. 1 shows molecular models of the compounds considered in this work, and their chemical formulae are given in Fig. 2. As we discuss below, the reactions of all four compounds with the catalytic amino acid residue Cys145 of M^{Pro} involve the nucleophilic attack of the Cys145 thiolate on the target carbon atom of the inhibitor (red asterisks in Fig. 1 mark these target carbon atoms); however, the detailed mechanisms are different. In the figures and in the text, we refer to these carbon atoms and their chemically bound partners (see Fig. 1) oxygen, nitrogen, fluorine without additional indices (in the files in the Supporting Information (SI), these atoms have specific indices in each compound).

Carmofur, 1-hexylcarbamoyl-5-fluorouracil, is a known drug for the treatment of colorectal cancer.[47] Nirmatrelvir, (1R,2S,5S)-N-[(1S)-1-cyano-2-[(3S)-2-oxopyrrolidin-3-yl]ethyl]-3-[(2S)-3,3dimethyl-2-[(2,2,2-trifluoroacetyl)amino]butanoyl]-6,6-dimethyl-3-azabicyclo[3.1.0]hexane2-carboxamide, is also known as the substance PF-07321332 developed by Pfizer. This compound is an active component of the approved oral drug Paxlovid for the treatment of COVID-19.[48]

The two new molecules designed in this work were derived computationally from the structure of X77, N-(4-tert-butylphenyl)-N-[(1R)-2-(cyclohexylamine)-2-oxo-1-(pyridin3-yl)ethyl]-1H-imidazole-4-carboxamid. Several studies [50,51] described this compound and its mimetics as a promising *non-covalent* inhibitors of M^{Pro}. According to the results of molecular docking,[51] the binding energy of X77 to MPro is high (ΔG~-10 kcal/mol), giving rise to the dissociation constant of 0.057 μM. Here, we follow a different strategy, aiming to develop effective *covalent* inhibitors. Specifically, we propose to modify X77 by introducing warhead groups capable of efficient chemical reactions with the catalytic cysteine residue in the M^{Pro} active site (see Fig. 1). In other words, we propose to turn an efficient non-covalent inhibitor into a covalent inhibitor.

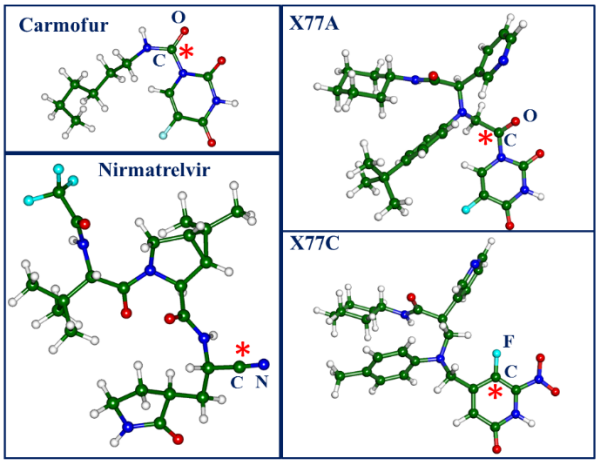


Figure 1. Molecular models of the compounds considered in this work as covalent inhibitors of M^{Pro}. Here and in all figures, carbon atoms are colored green, oxygen—red, nitrogen—blue, sulfur—yellow, fluorine—cyan, hydrogen—white. Red asterisks mark the target carbon atoms of the nucleophilic attack of the Cys145 thiolate of M^{Pro}.

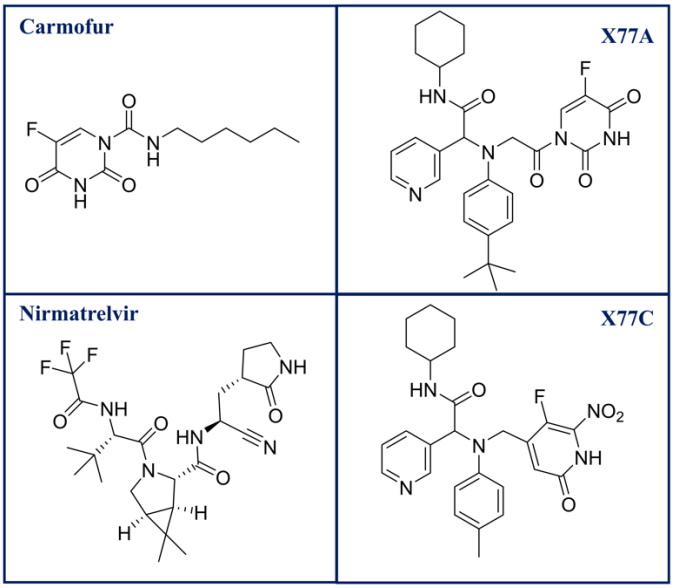


Figure 2. The chemical formulae of the compounds shown in Fig. 1.

According to the current knowledge, reactions of the covalent binding of the catalytic Cys145 of M^{Pro} are initiated by the proton transfer from cysteine to its partner in the catalytic dyad, His41, followed by the nucleophile attack of the sulfur ion on the target carbon atom of the ligand[23–45,53–57] (e.g., see discussion in Ref. 41). The emerging negative charge on the atom chemically bound to the target carbon atom (oxygen in carmofur and X77A, nitrogen in nirmatrelvir, fluorine in X77C) is stabilized by the oxyanion hole formed by the peptide chain Gly143-Ser144-Cys145. Thus, the following structural elements are important for modeling the inhibition reaction: the side chains in the

catalytic dyad (Cys145/His41) and the oxyanion hole (Gly143-Ser144-Cys145); this is common for all four compounds considered in this work. The differences are as follows. In reactions with carmofur, X77A, and X77C, the formation of the covalent bond between the sulfur and carbon atoms is accompanied by the leaving group (fluoro-uracil for carmofur and X77A, fluorine ion for X77C), whereas in the reaction with nirmatrelvir, there is a proton transfer pathway, which saturates the emerging valency in the nitrile nitrogen. Mechanisms of the creation of the leaving groups may also follow different scenarios. These important details of the reaction mechanisms are the focus of our study.

SYSTEM PREPARATION AND COMPUTATIONAL PROTOCOLS

We used the following strategy to simulate mechanisms of selected reactions. First, at the QM/MM level, we optimized structures of stationary points on the potential energy surfaces (PES) corresponding to the reactants, products, intermediates, and transition states along the hypothesized reaction coordinates. The analysis of the corresponding structures informed the selection of collective variables for the subsequent calculations of the Gibbs energy profiles using molecular dynamics simulations with QM/MM potentials (QM/MM MD).

The crystal structure PDB ID: 7BUY[21] of M^{Pro} with the aliphatic tail of the carmofur molecule attached to Cys145 served as a template to construct all model systems. The M^{Pro}–carmofur model for QM/MM calculations of the reaction of covalent inhibition of M^{Pro} by carmofur using the NWChem[58] and Q-Chem[59,60] software packages, was reported in our previous study.[61] We prepared model systems of the three other compounds by inserting the corresponding substrates into the protein structure using molecular mechanics tools. To validate the structures produced by molecular mechanics, we used the crystal structures of M^{Pro} complexed with the relevant ligands, i.e., PDB ID: 7VH8 for the product of the M^{Pro}-nirmatrelvir reaction and PDB ID: 6W63 for the complex of the non-covalent inhibitor X77 with M^{Pro}.

Fig. 3 shows the fragments of the active site of M^{Pro}, as they appear in the PDB structures relevant to the present simulations. We pay attention to the position of the amino acid residues of the catalytic, Cys145 and His41, and of the oxyanion hole side chains, Gly143-Ser144-Cys145, which is directly related to the chemical reactions of the selected compounds (Figs. 1,2) with Cys145 in the protein cavity. To design prospective covalent inhibitors, we replaced the molecular group of the non-covalent M^{Pro} inhibitor X77 (highlighted in yellow in Fig. 3) by the reactive warheads (see panels X77A and X77C in Figs. 1,2). Importantly, our molecular docking calculations show that the X77A and X77C molecules have binding energies with M^{Pro} similar to those of the parent X77 species: -8.9 kcal/mol for X77A and -9.4 kcal/mol for X77C, to be compared with our computed value for X77 -9.74 kcal/mol, or with the literature value of -10.2 kcal/mol.[51] Therefore, the proposed molecules X77A and X77C exhibit a high affinity to the catalytic site of M^{Pro}.

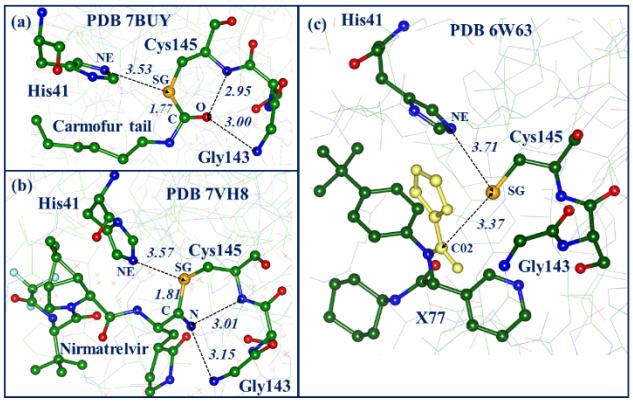


Figure 3. Fragments of the M^{Pro} active site from the selected PDB structures. Panels (a) and (b): fragments of the active site of M^{Pro} relevant to the reactions of selected compounds with Cis145 as they appear in the PDB structures. We focus on the chain Gly143-Ser144-Cys145 with the reactive Cys145 and the oxyanion hole groups and to the location of the His41 side chain relative to Cys145. Panel (c): the PDB structure 6W63, a non-covalent complex of M^{Pro} with X77. To design covalent inhibitors, we replaced the fragment of X77 (highlighted in yellow) by the reactive warheads (see panels X77A and X77C in Fig. 1). Here and below the distances are given in Å.

The partitioning of the model systems into the QM and MM parts for each compound is explained in Results and in the SI. As emphasized in our QM/MM study of the M^{Pro}–carmofur model,[61] reporting only relevant structures from PDB as initial coordinates of heavy atoms and the partitioning of the system into QM and MM parts is not sufficient to ensure the reproducibility of QM/MM-based calculations of the energy profiles of enzymatic reactions. More details need to be reported for others to be able to evaluate the results and to reproduce the findings. In the SI, we provide details of the preparation of the model systems, including addition of hydrogen atoms and the protonation states of the amino acid side chains, solvation of proteins, initial relaxation of the model structures using classical MD, link-atom schemes in the QM/MM boundary treatment, embedding protocols, and optimization algorithms.

We use the following notations for the computed structures: REAC (reactant state), IP (ion pair state), PROD (product state), TS (transition state). For each system, the REAC structure refers to the neutral state, Cys145/His41, in the catalytic dyad; IP corresponds to the structure with the ion-pair state, Cys^-/His^+, PROD corresponds to the structure with the covalently bound Cys145 with a leaving group kept in the active site. Reaction intermediates, besides the IP state, are described in the corresponding subsections; in particular, TI means the tetrahedral intermediate and MC means the Meisenheimer complex.

The QM/MM optimization of the stationary points was carried out using the density functional theory with the PBE0 functional[62] with the dispersion correction ($D3$[63]) to describe the QM part. The performance of the PBE0

functional has been extensively documented and benchmarked (see, for example, Ref. 64). It has been shown to perform well for computing reaction profiles for organic molecules.[64] Our groups used this functional for simulating other biological systems. In our previous paper (Ref. 61) we compared this functional against more advanced ones (wB97X-D) and found that the energy profile is insensitive to the functional choice. Energies and forces in the MM part were computed with the AMBER force-field parameters.[65] These QM(PBE0-D3/6-31G*)/MM(AMBER) calculations were performed using the NWChem program [58] with the electrostatic embedding scheme. The QM/MM-optimized minimum-energy structures were obtained in series of unconstrained minimizations. The TS structures were optimized in series of constrained minimizations, assuming appropriate reaction coordinates. The structures of TSs, separating the corresponding minimum-energy points, were verified by performing forward and backward descent from the located saddle points. Additional details are given in Results and in the SI. The optimized coordinates of all structures were deposited to the COVID-19 hub repository supported by MolSSI (see the SI for the complete list of deposited files).

To compute the Gibbs free energy profiles, which is an essential step in modeling protein-ligand systems,[66] we employed the algorithms based on biased MD trajectories.[67,68] The recent implementation[69] interfacing the MD program NAMD[70] with quantum chemistry packages allowed us to apply computational protocols with the QM/MM potentials, as in our previous studies of enzyme-catalyzed reactions.[71–73] Energies and gradients in QM were computed at the PBE0-D3/6-31G** level using the TeraChem program.[74] The CHARMM36 force-field[75] was used in the MM subsystems. Sizes of QM subsystems in these calculations were somewhat smaller than in the QM/MM optimization, but considerably larger than, for example, in previous studies[25-27] of the M^{Pro}–nirmatrelvir reaction. We performed umbrella sampling simulations with additional harmonic potentials centered at different collective variable values. Trajectories were 5-10 ps long; force constants were 10-40 kcal/mol/Å^2; the umbrella integration and weighted histogram analysis were used. Further details, such as selection of collective variables and the QM-MM partitioning, are described in Results and in the SI.

RESULTS AND DISCUSSION

Reaction of M^{Pro} with carmofur

As explained above, we used the M^{Pro}-carmofur model system, which was characterized in Ref. 61, as a template for modeling reactions of covalent inhibition of the enzyme by all compounds considered in the present work. In Ref. 61, we focused only on the reaction step from IP to PROD (using the QM/MM scheme with a slightly different QM-MM partitioning). Here, we also consider the step of the ion pair ($Cys145^-$/$His41^+$) formation. In the present calculations, the QM subsystem consisted of 155 atoms, including the entire carmofur molecule, the side chains of His41, Gly143, Ser144, Cys145 side chains; the detailed description of the computational protocol is given in the SI.

Fig. 4 shows the QM/MM optimized structures of REAC, IP, and PROD. The computed structures of all stationary points on the PES, including the transition states, are given in the SI. At the step of IP formation, REAC→TS1→IP, a two-dimensional energy plot along the distances between the transferring proton HS and the SG atom of Cys145 and the NE atom of His41, d(SG-HS) and d(NE-HS), allowed us to estimate the TS1 point separating the REAC and IP structures. Along this pathway we observed a gradual decrease of the distance of the nucleophilic attack, d(SG-C), from the initial value 3.62 Å in REAC to 3.15 Å in IP. At the next step IP→TS2→PROD, the distance d(SG-C) served as a reaction coordinate in the QM/MM constrained optimization. After passing the TS2 structure, the covalent bond SG-C is formed, and the leaving group, the fluoro-uracil warhead, is separated from the formed covalent adduct of M^{Pro} with the aliphatic tail of the carmofur molecule. The adduct is firmly accommodated in the protein cavity and the C-O group is captured in the oxyanion hole.

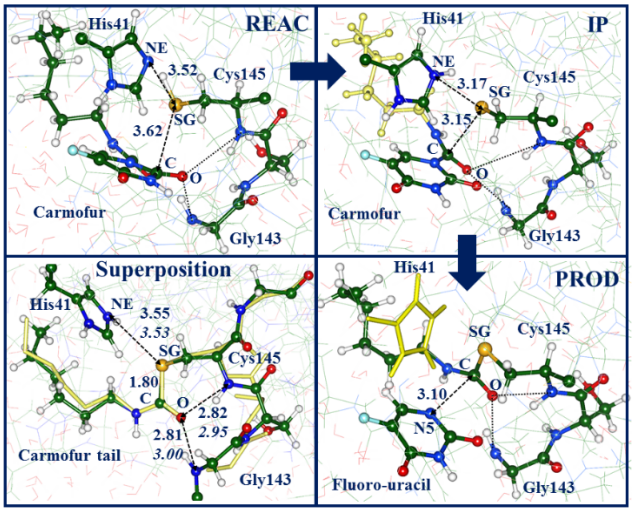


Figure 4. The QM/MM optimized structures of REAC, IP, PROD for the M^{Pro}—carmofur reaction. The left bottom panel shows the superposition of PROD (colored balls and sticks) and the crystal structure 7BUY (yellow sticks). The side chain of His41 in the right bottom panel is shown in yellow sticks. The distances in italics correspond to the crystal structure.

We use the only available piece of experimental information, the crystal structure of the reaction products, to validate the computational protocol. The bottom left panel in Fig. 4 ('Superposition') shows the active site of the computed PROD structure (colored balls and sticks) and compares it with the relevant fragment (yellow sticks) of the crystal structure (PDB ID: 7BUY[21]). We note a good agreement for the key distances between the computationally derived structure and the crystal structure (see Fig. 3). In the adduct of the protein with the

carmofur tail, the distances between atoms SG and NE in the catalytic dyad, as well as the distances in the oxyanion hole region between the nitrogen atoms of Cys145 and Gly143 and the oxygen atom (O) in the adduct of the protein with the carmofur tail are close in the experimental and computationally derived structures, even though the leaving group (the fluoro-uracil warhead) is not present in the crystal structure, but it is kept in the active site of the model system.

The computed QM(PBE0-D3/6-31G*)/MM(AMBER) energies of all stationary points are as follows: REAC→TS1 (+12) →IP (+9) →TS1 (+16) →PROD (-13). Here and below, the values in parentheses gives the energy of the corresponding stationary point in kcal/mol relative to the level of REAC. According to these results, the highest energy barrier in the M^{Pro}–carmofur reaction leading to a stable covalent adduct corresponds to the formation of the IP state. To estimate the activation energy in this reaction, we carried out calculations of the free energy profile.

We computed the Gibbs free energy profile with the QM(PBE0-D3/6-31G**)/MM(CHARMM36) potentials used in MD simulations. The QM part included the carmofur molecule, the Cys145 and His41 side chains, and a nearby water molecule. We defined the reaction coordinate (the collective variable (CV)) as the combination of the relevant distances: CV=d(SG-HS)-d(NE-HS); the details are presented in the SI. The computed profile (Fig. 5) shows the activation barrier of 10.4 kcal/mol and the position of the IP state 9.3 kcal/mol higher than that of the REAC state, which is consistent with the energies of the stationary points optimized in the QM/MM calculations.

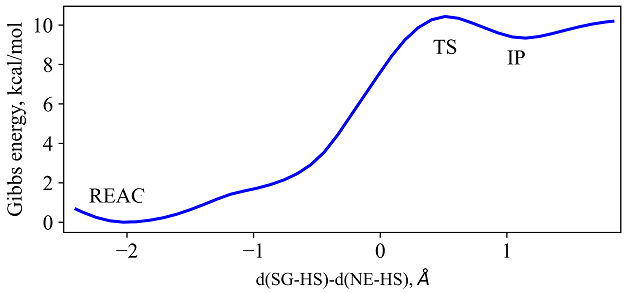


Figure 5. The computed Gibbs free energy profile for the REAC→TS1→IP step of the ion-pair formation in the M^{Pro} - carmofur reaction.

We note that the step of the ion-pair formation is common for the catalytic cycle of cysteine proteases;[36,54,55] however, different computational studies evaluating the corresponding free energy surface resulted in different free-energy profiles. For example, for the ion-pair formation in the reaction of M^{Pro} covalent inhibition by the N3 peptidyl Michael acceptor, two research groups almost simultaneously reported the Gibbs free energy activation barriers of 1.4 kcal/mol[33] and 10.7 kcal/mol.[34]

To conclude this subsection, we note that the simulations describe the formation of the covalent adduct in the M^{Pro}–carmofur reaction[21] consistently with the experimental observations. However, no attempts to use carmofur as the COVID-19 drug have been reported.

Reaction of M^{Pro} with X77A

We designed the X77A compound by introducing the warhead with the fluoro-uracil moiety, resembling that in the carmofur molecule. The target atom for the nucleophile attack of the Cys145 thiolate is the similar carbonyl carbon atom marked by the asterisk in Fig. 1. Thus, it is reasonable to expect that the mechanism of the reaction M^{Pro} with X77A resembles that in the M^{Pro} reaction with carmofur with the same leaving group. Although the basic features are common, we located an additional reaction intermediate (besides IP)—a tetrahedral intermediate (TI)—on the route from IP to PROD.

In the QM/MM optimization, the large QM part included the entire X77A molecule, the molecular groups of His41, Cys145, Ser144, Gly143, Thr25, Thr26, Leu27, Leu141, Asn142, Gly146, His164, Met165, Asp187, and 7 water molecules (208 atoms in total).

The panels in Fig. 6 illustrate the minimum energy points optimized in the QM(PBE0-D3/6-31G*)/MM(AMBER) calculations; the structures of all stationary points including TSs, are shown in the SI. The first step—REAC→TS1(+8)→IP(+4)—shares similar features with the M^{Pro} - carmofur reaction (cf. upper panels in Figs. 4, 6), but with a slightly lower energy barrier. Scans along the gradually decreasing coordinate d(SG-C) allowed us to locate the stationary points at the subsequent reaction steps: IP(+4)→TS2 (+5) →TI (-15) and TI (-15) →TS3 (-14) →PROD (-23). According to these QM(PBE0-D3/6-31G*)/MM(AMBER) calculations, the highest energy barrier corresponds to the step of the IP formation, whereas the energy barriers at the subsequent steps are low, 1-4 kcal/mol. The structure of the products (the bottom left panel in Fig. 6) shows that the covalent adduct is firmly trapped in the protein cavity; the C-O bond is captured by the oxyanion hole.

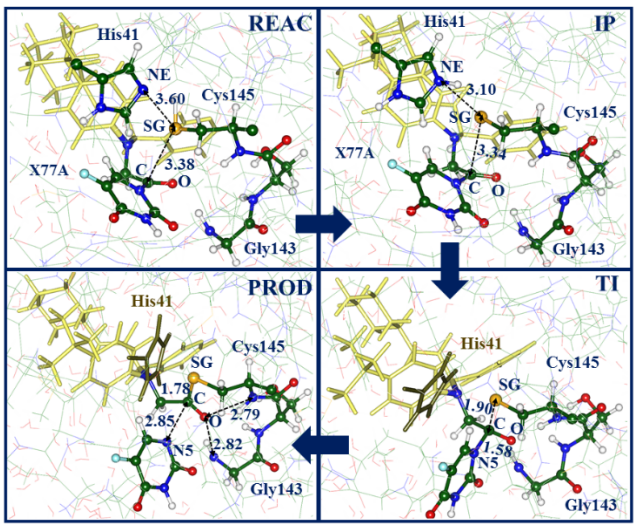


Figure 6. The QM/MM optimized structures for the M^{Pro} – X77A reaction. A large part of the X77A molecule is shown in

light yellow sticks. The side chain of His41 in the bottom panels is shown in goldish yellow sticks.

The QM(PBE0-D3/6-31G**)/MM(CHARMM36) MD simulations resulted in the Gibbs energy profiles for the M^{Pro} – X77A reaction illustrated in Fig. 7. The QM part included the fragment of the substrate, His41, Leu141, Asn142, Gly143, Ser144, Cys145, Gly146, His164, and Met165, a water molecule that interacts with the His41. Collective variables were selected as follows: CV1=d(NE-HS)-d(SG-HS) at the reaction step of the ion pair formation, and CV2=d(SG-C)-d(C-N5) for the subsequent steps. The upper part in Fig. 7 summarizes the data showing that activation barriers along the reaction pathway are low; the highest energy barrier corresponds to the formation of the ion pair state.

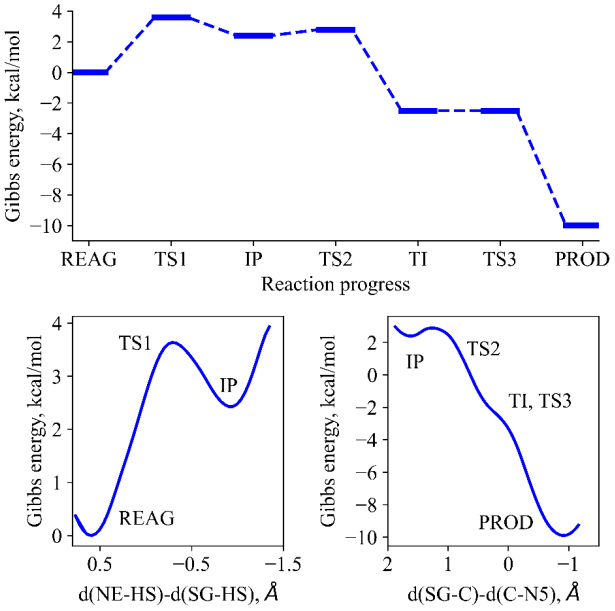


Figure 7. The computed Gibbs free energy profiles for the M^{Pro} – X77A reaction. The upper panel shows the diagram combining the results at the two reaction steps illustrated in the bottom panels. The collective variables are as follows: CV1 = d(NE-HS)-d(SG-HS), CV2 = d(SG-C)-d(C-N5).

Thus, according to the present simulations, the X77A compound should be an efficient covalent inhibitor of M^{Pro}.

Reaction of M^{Pro} with X77C

Klein et al.[45] proposed to use aromatic compounds that can react with the catalytic cysteine by the S_NAr addition/elimination mechanism as a new class of covalent inhibitors of cysteine proteases. Several such compounds have been tested as prospective inhibitors of the protease rhodesain.[45] Inspired by this idea, we introduced the 5-fluoro-6-nitro-pyrimidine2,4(1H,3H)-dione warhead into the X77 template to create compound X77C. Upon deprotonation, the sulfur ion of Cys145 attacks the carbon atom C initially bound to fluorine (see the X77C panel in Fig. 1).

We used the same strategy as for the M^{Pro} – X77A reaction to characterize the energy profiles for the M^{Pro} – X77C reaction and to dissect the reaction mechanism: the QM/MM calculations of the structures on the PES followed by the QM/MM MD calculations of the Gibbs free energy profiles.

The results of QM(PBE0-D3/6-31G*)/MM(AMBER) optimization of the minimum energy structures are shown in Fig. 8; the structures of all stationary points including TSs are given in the SI. In QM/MM optimization, the large QM part included the entire X77C molecule, the molecular groups of His41, Cys145, Ser144, Gly143, Thr25, Thr26, Leu27, Leu141, Asn142, Gly146, His164, Met165, Asp187, and 7 water molecules (203 atoms in total).

According to these results (REAC→TS1(+4)→IP(0)→TS2 (+2) →MC (-15) →TS3 (-14) →PROD (-25), we located two reaction intermediates—the ion pair state (IP) and the Meisenheimer complex (MC), which are separated by fairly low energy barriers (not exceeding 4 kcal/mol). The structure of PROD confirms the formation of the covalent adduct; the leaving group (F^-) is captured by the anion hole.

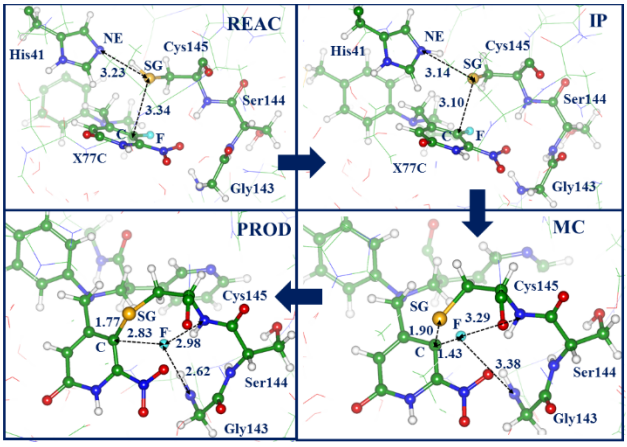


Figure 8. Structures of the QM/MM optimized structures for the M^{Pro} – X77C reaction.

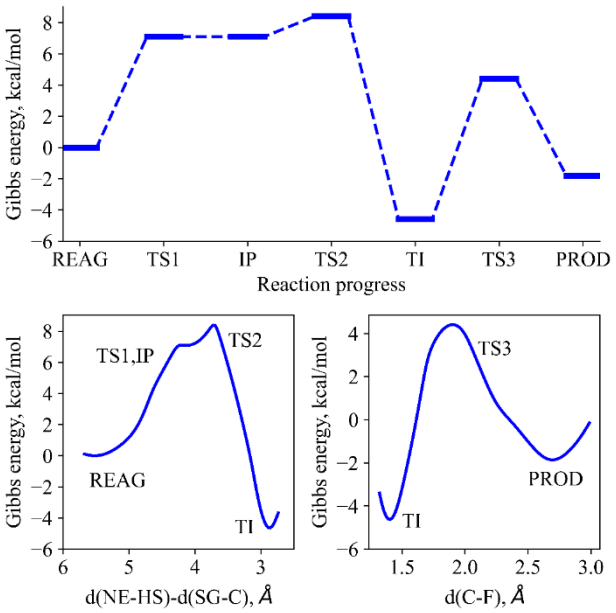


Figure 9. The computed Gibbs free energy profiles for the M^{Pro}–X77C reaction. The upper panel shows the diagram combining the results at the two reaction steps illustrated in the bottom panels. The collective variables are as follows: CV1 = d(NE-HS)-d(SG-C), CV2 = d(C-F).

The results of the QM(PBE0-D3/6-31G**)/MM(CHARMM36) MD simulations of the Gibbs free energy profiles are shown in Fig. 9. The QM part included the fragment of the substrate, His41, Leu141, Asn142, Gly143, Ser144, Cys145, Gly146, His164, and Met165, a water molecule that interacts with the His41. Collective variables were selected after several trials: CV1=d(NE-HS)-d(SG-C) up to the MC formation and CV2=d(C-F) at the subsequent step. In this case, the IP intermediate is not clearly visible on the free energy surface; the free energy profile at the step resembles the features reported by Ramos-Guzmán et al. in modeling other reactions of covalent inhibition of M^{Pro}.[24-26,34] In contrast, the MC intermediate corresponds to the minimum-energy structure in both the QM/MM and QM/MM MD calculations. We note that the nature of the Meisenheimer complex in the S_NAr reactions[45,76,77] is still debated, in particular, whether it represents a reaction intermediate or a transition state. In our case, the results favor the formation of the minimum-energy structure separated from the reactants and products by the free energy barriers of 13 and 4.4 kcal/mol.

We conclude that the compound X77C can react with M^{Pro} with low activation barriers, leading to the covalent binding of the catalytic Cys145.

Reaction of M^{Pro} with nirmatrelvir

The molecular model of nirmatrelvir is shown in the lower left panels in Fig. 1. The covalent binding of nirmatrelvir by M^{Pro} is confirmed by several crystal structures in the Protein Data Bank (e.g., PDB IDs: 7VH8, 7MLG, 7MLF). The reaction of M^{Pro} with a nitrile-based ligand, such as nirmatrelvir, should lead to a covalent thioimidate adduct after deprotonation of Cys145 and the nucleophilic attack of the thiolate making the SG-C covalent bond.[33,34,44]

The interaction of the nirmatrelvir molecule with M^{Pro} studied by classical molecular dynamics simulations[49] shows a tight binding of this compound at the protein surface. Three computational papers [33-35] reported results on the mechanism of the M^{Pro}– nirmatrelvir reaction based on QM/MM calculations.

Ramos-Guzmán et al.[25,26] computed the minimum free energy path for the nirmatrelvir covalent binding to M^{Pro} using the adaptive string method with QM/MM potentials. The QM subsystem composed of the fragments of Cys145 and His41, a water molecule and the warhead of the inhibitor (about 50 atoms in total), was described at the B3LYP-D3/6-31+G* level in Ref. 25 and at the M06-2X-D3/6-31+G* level in Ref. 26. The path was determined through biased MD simulations using 7 collective variables that included the distances of all the bonds being broken, formed or whose formal order changed during the process. The computed profiles show single TS of 14-16 kcal/mol and the reaction energy of 10-14 kcal/mol. [25,26] No clear stabilization of the IP state was found.

Ngo et al.[27] used the ONIOM version of QM/MM to evaluate the M^{Pro}–nirmatrelvir reaction energy profile. The authors assumed that Cys145, His41 and the nearby residue Asp187 form a catalytic triad to facilitate covalent binding of the ligand to the protein. The QM part comprised 49 atoms, including small fractions of the ligand and of the amino acid triad, described at the B3LYP-D3/6-31G(d) level upon QM/MM optimization followed by single point calculations at the M06-2X/6-311+G(2d,2p) level. The constructed energy diagram corresponds to a flat profile within 3.4 kcal/mol at the first reaction steps showing no formation of the ion-pair state. At the final step, describing proton transfer with a participation of a mediated water molecule, an energy barrier of 19 kcal/mol was reported. The computed energy of reaction products was about 9 kcal/mol below the level of reactants.

The QM part in QM(PBE0-D3/6-31G*)/MM(AMBER) optimization included the reactive part of the nirmatrelvir molecule, the molecular groups of His41, Gly143, Ser144, Cys145, Thr25, Thr26, Leu27, Gly146, Ser147, Val148, Met162, His163, His64, Met165, and 10 water molecules. The results of QM/MM optimization of the minimum-energy points on PES are shown in Fig. 10; the relative energies are as follows: REAC→TS1(+2)→INT(-18)→TS2 (-12)→PROD (-29).

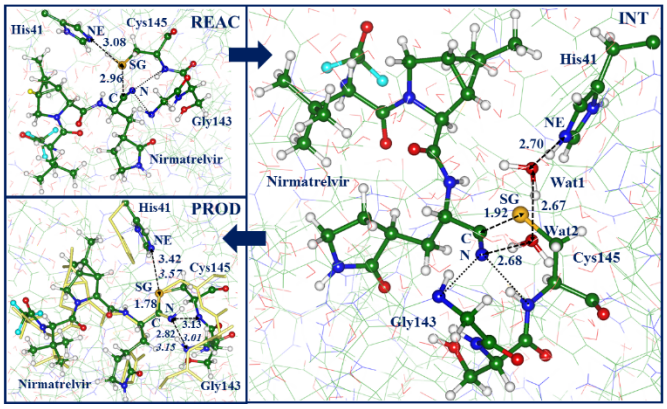


Figure 10. The QM/MM optimized structures for the M^{Pro}–nirmatrelvir reaction.

We located a reaction intermediate (INT). Its structure has a short distance d(SG-C) of 1.92 Å (much shorter than that in the IP state in the reactions of M^{Pro} with camofur, X77A, and X77C), whereas the His41 side chain remains protonated (positively charged). To complete the reaction, i.e., to protonate the N atom of the ligand, two water molecules, Wat1 and Wat2 shown in the right part in Fig. 10, form a proton wire from the N_ε atom of His41 with typical distances about 2.7 Å between heavy atoms.

The Gibbs free energy profiles were computed with the QM(PBE0-D3/6-31G*)/MM(CHARMM36) potentials describing the 97-atomic subsystem composed of the nirmatrelvir molecule, Cys145 and His41 side chains, and water molecules. The obtained profile is shown in Fig. 11.

The results of the present simulations agree in part with the previous modeling of the M^{Pro}–nirmatrelvir reaction.[25-27] In particular, all approaches do not favor the formation of the IP state as an energy minimum. All approaches obtain a considerable reaction energy, e.g., -14 kcal/mol in the present Gibbs free energy calculations. On the other hand, we do not obtain a high energy barrier on the reaction pathway—our values do not exceed 5 kcal/mol (Fig. 11), in contrast to the values of 14-18 kcal/mol in Refs. 25-27. We cannot confirm the hypothesis of the Cys145-His41-Asp187 catalytic triad put forward in Ref. 26. Qualitatively, our reaction mechanism is close to the one described in Refs. 25,26, but assumes the formation of the reaction intermediate INT (Figs. 10, 11) with the features resembling those of TI in the reaction of M^{Pro} with X77A (Fig. 6).

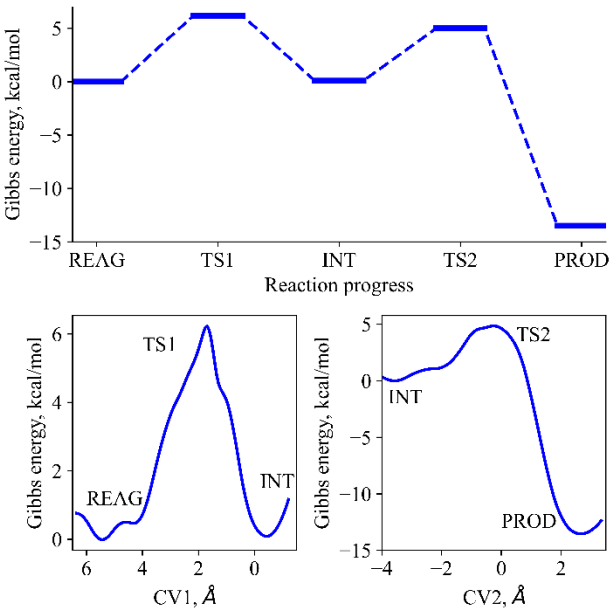


Figure 11. The computed Gibbs free energy profiles for the M^{Pro}–nirmatrelvir reaction. The upper panel shows the diagram combining the results at the two reaction steps illustrated in the bottom panels. The collective variables are as follows: CV1=d(NE-HS)-d(SG-HS)-d(SG-C), CV2=d(NE-HS)-d(HS-OWat1)+d(OWat1-HWat1)-d(HWat1-OWat2)+d(OWat2-HWat2)-d(N-HWat2).

To conclude this Section, we note that the COVID-19 pandemic motivated numerous studies of the SARS-related enzymes, which considerably expanded the understanding of the enzyme catalysis. Our study contributes to these efforts. We modeled reactions of four compounds and show that these compounds are capable of bindding chemically to the catalytic cysteine residue of M^{Pro} and, therefore, can serve as irreversible inhibitors of this enzyme. The simulations revealed three distinct reaction mechanisms. We recognize that these three mechanisms do not exhaust all possible scenarios—other documented examples of the M^{Pro} inhibition include the Michael addition to the unsaturated carbon-carbon bond[34,35] and the reactions with ketones.[24,32]

Our simulations contribute to the ongoing efforts to find more effective drugs to fight COVID-19. We show that the employed computational protocols are sufficiently reliable and produce the results consistent with the already known information: the computed energy profiles for carmofur and nirmatrelvir show that the corresponding reactions with M^{Pro} are efficient with respect to energy barriers and reaction energies. Therefore, we expect that the compounds designed computationally in our work, X77A and X77C, and characterized at the same level of theory are promising drug candidates for blocking M^{Pro}.

CONCLUSION

The results of our QM/MM modeling of chemical reactions of the catalytic Cys145 amino acid residue of the SARS-CoV-2 main protease with four compounds,

carmofur, nirmatrelvir, X77A, X77C, show that these species can form stable covalent adducts with M^{Pro}, and the activation barriers are sufficiently low for the reactions to be efficient. The results for carmofur and nirmatrelvir are consistent with the experimental findings, and the success of the simulations provides a sound basis for a prediction of the two novel potential inhibitors, X77A and X77C, proposed in this work. From the fundamental perspective, this study illustrates that the formation of covalent adducts follow three distinct reaction mechanisms of the irreversible inhibition of cysteine proteases.

Supporting Information.
Details about model systems preparation, QM/MM setup and free energy calculations, input parameters, list of structures deposited to the MolSSI COVID-19 hub. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Authors

* Anna I. Krylov (krylov@usc.edu) and Alexander V. Nemukhin (anemukhin@yahoo.com)

Funding Sources

This work was supported by the U.S. National Science Foundation (No. CHE-1856342 to AIK) and by the Russian Science Foundation (grant No. 19-73-20032 to BLG, IVP, and AVN). We used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562 (XSEDE resource: Comet; allocation ID: TG-CHE200131). We acknowledge resources of MolSSI. BLG, IVP, MGK, and AVN acknowledge the use of HPC computing resources at the Lomonosov Moscow State University.

Notes

The authors declare the following competing financial interest(s): A.I.K. is the president and a part-owner of Q-Chem, Inc.

ACKNOWLEDGMENT

We thank S.V. Lushchekina for help with the molecular docking calculations.

While finalizing this manuscript, one of the authors (AIK) contracted COVID-19 and was treated by the currently available drug Paxlovid, which contains nirmatrelvir (PF-07321332) as an active ingredient targeting M^{pro}.

REFERENCES

(1) Warshel, A.; Levitt, M. Theoretical Studies of Enzymatic Reactions: Dielectric Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103*, 227.

(2) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198–1229.

(3) Liu, M.; Wang, Y.; Chen, Y.; Field, M. J.; Gao, J. QM/MM through the 1990s: The First Twenty Years of Method Development and Applications. *Isr. J. Chem.* **2014**, *54* (8–9), 1250–1263.

(4) van der Kamp, M. W.; Mulholland, A. J. Combined Quantum Mechanics/Molecular Mechanics (QM/MM) Methods in Computational Enzymology. *Biochemistry* **2013**, *52*, 2708–2728.

(5) Świderek, K.; Tuñón, I.; Moliner, V. Predicting Enzymatic Reactivity: From Theory to Design. *WIREs Comput. Mol. Sci.* **2014**, *4* (5), 407–421.

(6) Kulik, H. J.; Zhang, J.; Klinman, J. P.; Martínez, T. J. How Large Should the QM Region Be in QM/MM Calculations? The Case of Catechol O-Methyltransferase. *J. Phys. Chem. B* **2016**, *120* (44), 11381–11394.

(7) Roßbach, S.; Ochsenfeld, C. Influence of Coupling and Embedding Schemes on QM Size Convergence in QM/MM Approaches for the Example of a Proton Transfer in DNA. *J. Chem. Theory Comput.* **2017**, *13*, 1102–1107.

(8) Sousa, S. F.; Ribeiro, A. J. M.; Neves, R. P. P.; Brás, N. F.; Cerqueira, N. M. F. S. A.; Fernandes, P. A.; Ramos, M. J. Application of Quantum Mechanics/Molecular Mechanics Methods in the Study of Enzymatic Reaction Mechanisms. *WIREs Comput. Mol. Sci.* **2017**, *7* (2), e1281.

(9) Cui, Q.; Pal, T.; Xie, L. Biomolecular QM/MM Simulations: What Are Some of the "Burning Issues"? *J. Phys. Chem. B* **2021**, *125* (3), 689–702.

(10) Gao, K.; Wang, R.; Chen, J.; Cheng, L.; Frishcosy, J.; Huzumi, Y.; Qiu, Y.; Schluckbier, T.; Wei, X.; Wei, G. Methodology-Centered Review of Molecular Modeling, Simulation, and Prediction of SARS-CoV-2. *Chem. Rev.* **2022**, *122* (13), 11287–11368.

(11) Bai, C.; Wang, J.; Chen, G.; Zhang, H.; An, K.; Xu, P.; Du, Y.; D. Ye, R.; Saha, A.; Zhang, A.; Warshel, A. Predicting Mutational Effects on Receptor Binding of the Spike Protein of SARS-CoV-2 Variants. *J. Am. Chem. Soc.* **2021**, *143* (42), 17646–17654. https://doi.org/10.1021/jacs.1c07965.

(12) Zhou, J.; Saha, A.; Huang, Z.; Warshel, A. Fast and Effective Prediction of the Absolute Binding Free Energies of Covalent Inhibitors of SARS-CoV-2 Main Protease and 20S Proteasome. *J. Am. Chem. Soc.* **2022**, *144* (17), 7568–7572. https://doi.org/10.1021/jacs.2c00853.

(13) Wang, J.; Liang, B.; Chen, Y.; Fuk-Woo Chan, J.; Yuan, S.; Ye, H.; Nie, L.; Zhou, J.; Wu, Y.; Wu, M.; Huang, L. S.; An, J.; Warshel, A.; Yuen, K. Y.; Ciechanover, A.; Huang, Z.; Xu, Y. A New Class of α-Ketoamide Derivatives with Potent Anticancer and Anti-SARS-CoV-2 Activities. *Eur. J. Med. Chem.* **2021**, *215*, 113267. https://doi.org/10.1016/J.EJMECH.2021.113267.

(14) Anand, K.; Ziebuhr, J.; Wadhwani, P.; Mesters, J. R.; Hilgenfeld, R. Coronavirus Main Proteinase (3CLpro) Structure: Basis for Design of anti-SARS Drugs. *Science* **2003**, *300*, 1763–1767.

(15) Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; Duan, Y.; Yu, J.; Wang, L.; Yang, K.; Liu, F.; Jiang, R.; Yang, X.; You, T.; Liu, X.; Yang, X.; Bai; Liu, H.; Liu, X.; Guddat, L. W.; Xu, W.; Xiao, G.; Qin, C.; Shi, Z.; Jiang, H.; Rao, Z.; Yang, H. Structure of Mpro from SARS-

CoV-2 and Discovery of Its Inhibitors. *Nature* **2020**, *582* (7811).

(16) 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* (6489), 409–412.

(17) Cannalire, R.; Cerchia, C.; Beccari, A. R.; Di Leva, F. S.; Summa, V. Targeting SARS-CoV-2 Proteases and Polymerase for COVID-19 Treatment: State of the Art and Future Opportunities. *J. Med. Chem.* **2022**, *65* (4), 2716–2746.

(18) Ullrich, S.; Sasi, V. M.; Mahawaththa, M. C.; Ekanayake, K. B.; Morewood, R.; George, J.; Shuttleworth, L.; Zhang, X.; Whitefield, C.; Otting, G.; Jackson, C.; Nitsche, C. Challenges of Short Substrate Analogues as SARS-CoV-2 Main Protease Inhibitors. *Bioorg. & Med. Chem. Lett.* **2021**, *50*, 128333.

(19) Li, Q.; Kang, C. Progress in Developing Inhibitors of SARS-CoV-2 3C-Like Protease. *Microorganisms* **2020**, *8*, 1250.

(20) Dai, W.; Zhang, B.; Jiang, X.; Su, H.; Li, J.; Zhao, Y.; Xie, X.; Peng, Z. J. J.; Liu, F.; Li, C.; Li, Y.; Bai, F.; Wang, H.; Cheng, X.; Cen, X.; Hu, S.; Yang, X.; Wang, J.; Liu, X.; Xiao, G.; Jiang, H.; Rao, Z.; Zhang, L.; Xu, Y.; Yang, H.; Liu, H. Structure-Based Design of Antiviral Drug Candidates Targeting the SARS-CoV-2 Main Protease. *Science* **2020**, *368*, 1331–1335.

(21) Jin, Z.; Zhao, Y.; Sun, Y.; Zhang, B.; Wang, H.; Wu, Y.; Zhu, Y.; Zhu, C.; Hu, T.; Du, X.; Duan, Y.; Yu, J.; Yang, X.; Yang, X.; Yang, K.; Liu, X.; Guddat, L. W.; Xiao, G.; Zhang, L.; Yang, H.; Rao, Z. Structural Basis for the Inhibition of SARS-CoV-2 Main Protease by Antineoplastic Drug Carmofur. *Nat. Struct. & Mol. Bio.* **2020**, *27*, 529–532.

(22) Ma, S.; Devi-Kesavan, L. S.; Gao, J. Molecular Dynamics Simulations of the Catalytic Pathway of a Cysteine Protease: A Combined QM/MM Study of Human Cathepsin K. *J. Am. Chem. Soc.* **2007**, *129* (44), 13633–13645.

(23) Taranto, A. G.; Carvalho, P.; Avery, M. A. QM/QM Studies for Michael Reaction in Coronavirus Main Protease (3CL Pro). *J. Mol. Graph. Model.* **2008**, *27* (3), 275–285.

(24) Ramos-Guzmán, C. A.; Ruiz-Pernía, J.J. ; Tuñón, I. Inhibition Mechanism of SARS-CoV-2 Main Protease with Ketone-Based Inhibitors Unveiled by Multiscale Simulations: Insights for Improved Designs. *Angew. Chem., Int. Ed.* **2021**, *60* (49), 25933–25941.

(25) Ramos-Guzmán, C. A.; Ruiz-Pernía, J. J.; Tuñón I. Computational Simulations on the Binding and Reactivity of a Nitrile Inhibitor of the SARS-CoV-2 Main Protease. *Chem. Comm.* **2021**, *57* (72), 9096–9099.

(26) Ramos-Guzmán, C. A.; Andjelkovic, M.; Zinovjev, K; Ruiz-Pernía, J. J.; Tuñón I. Impact of SARS-Cov-2 3CL Protease Mutations on Nirmatrelvir Inhibitory Efficiency. Computational Insights into Potential Resistance Mechanisms. *Chem. Sci.*, **2023**, DOI:10.1039/D2SC06584C.

(27) Ngo, S. T.; Nguyen, T. H.; Tung, N. T.; Mai, B. K. Insights into the Binding and Covalent Inhibition Mechanism of PF-07321332 to SARS-CoV-2 Mpro. *RSC Adv.* **2022**, *12* (6), 3729–3737. https://doi.org/10.1039/D1RA08752E.

(28) Oanca, G.; Asadi, M.; Saha, A.; Ramachandran, B.; Warshel, A. Exploring the Catalytic Reaction of Cysteine Proteases. *J. Phys. Chem. B* **2020**, *124* (50), 11349–11356.

(29) Wei, D.; Huang, X.; Liu, J.; Tang, M.; Zhan, C. Reaction Pathway and Free Energy Profile for Papain-Catalyzed Hydrolysis of N-Acetyl-Phe-Gly 4-Nitroanilide. *Biochemistry* **2013**, *52* (30), 5145–5154.

(30) Świderek, K.; Moliner, V. Revealing the Molecular Mechanisms of Proteolysis of SARS-CoV-2 Mpro by QM/MM Computational Methods. *Chem. Sci.* **2020**, *11*, 10626–10630.

(31) Ramos-Guzmán, C. A.; Ruiz-Pernía, J. J.; Tuñón I.. Unraveling the SARS-CoV-2 Main Protease Mechanism Using Multiscale Methods. *ACS Cat.* **2020**, *10* (21), 12544–12554.

(32) Mondal, D.; Warshel, A. Exploring the Mechanism of Covalent Inhibition: Simulating the Binding Free Energy of α-Ketoamide Inhibitors of the Main Protease of SARS-CoV-2. *Biochemistry* **2020**, *48*, 4601–4608.

(33) Arafet, K.; Serrano-Aparicio, N.; Lodola, A.; Mulholland, A. J.; González, F. V; Swiderek, K.; Moliner, V. Mechanism of Inhibition of SARS-CoV-2 Mpro by N3 Peptidyl Michael Acceptor Explained by QM/MM Simulations and Design of New Derivatives with Tunable Chemical Reactivity. *Chem. Sci.* **2021**, *12*, 1433–1444.

(34) Ramos-Guzmán, C. A.; Ruiz-Pernía, J. J.; Tuñón I. A Microscopic Description of SARS-CoV-2 Main Protease Inhibition with Michael Acceptors. Strategies for Improving Inhibitor Design. *Chem. Sci.* **2021**, *12* (10), 3489–3496.

(35) Ramos-Guzmán, C. A.; Ruiz-Pernía, J. J.; Tuñón I. Multiscale Simulations of SARS-CoV-2 3CL Protease Inhibition with Aldehyde Derivatives. Role of Protein and Inhibitor Conformational Changes in the Reaction Mechanism. *ACS Cat.* **2021**, *11* (7), 4157–4168.

(36) Otto, H.-H.; Schirmeister, T. Cysteine Proteases and Their Inhibitors. *Chem. Rev.* **1997**, *97* (1), 133–172.

(37) Powers, J. C.; Asgian, J. L.; Ekici, Ö. D.; James, K. E. Irreversible Inhibitors of Serine, Cysteine, and Threonine Proteases. *Chem. Rev.* **2002**, *102* (12), 4639–4750.

(38) Arafet, K.; Ferrer, S.; Moliner, V. First Quantum Mechanics/Molecular Mechanics Studies of the Inhibition Mechanism of Cruzain by Peptidyl Halomethyl Ketones. *Biochemistry* **2015**, *54*, 3381–3391.

(39) Arafet, K.; Ferrer, S.; Moliner, V. Computational Study of the Catalytic Mechanism of the Cruzain Cysteine Protease. *ACS Cat.* **2017**, *7* (2), 1207–1215.

(40) Arafet, K.; Ferrer, S.; González, F. V; Moliner, V. Quantum Mechanics/Molecular Mechanics Studies of the Mechanism of Cysteine Protease Inhibition by Peptidyl-2,3-Epoxyketones. *Phys. Chem. Chem. Phys.* **2017**, *19* (20), 12740–12748.

(41) Arafet, K.; Świderek, K.; Moliner, V. Computational Study of the Michaelis Complex Formation and the Effect on the Reaction Mechanism of Cruzain Cysteine Protease. *ACS Omega* **2018**, *3* (12), 18613–18622.

(42) Silva, J. R. A.; Cianni, L.; Araujo, D.; Batista, P. H. J.; Vita, D.; Rosini, F.; Leitao, A.; Lameira, J.; Montanari, C. A. Assessment of the Cruzain Cysteine Protease Reversible and Irreversible Covalent Inhibition Mechanism. *J. Chem. Inf. Mod.* **2020**, *60*, 880–889.

(43) Dos Santos, A. M.; Cianni, L.; De Vita, D.; Rosini, F.; Leitão, A.; Laughton, C. A.; Lameira, J.; Montanari, C. A. Experimental Study and Computational Modelling of Cruzain Cysteine Protease Inhibition by Dipeptidyl Nitriles. *Phys. Chem. Chem. Phys.* **2018**, *20* (37), 24317–24328.

(44) da Costa, C. H. S.; Bonatto, V.; Dos~Santos, A. M.; Lameira, J.; Leitaob, A.; Montanari, C. A. Evaluating QM/MM Free Energy Surfaces for Ranking Cysteine Protease Covalent Inhibitors. *J. Chem. Inf. Mod.* **2020**, *60*, 880–889.

(45) Klein, P.; Johe, P.; Wagner, A.; Jung, S.; Kühlborn, J.; Barthels, F.; Tenzer, S.; Distler, U.; Waigel, W.; Engels, B.; Hellmich, U. A.; Opatz, T.; Schirmeister, T. New Cysteine Protease Inhibitors: Electrophilic (Het)Arenes and Unexpected Prodrug Identification for the Trypanosoma Protease Rhodesain. *Molecules* **2020**, *25*, 1451.

(46) Vuong, W.; Khan, M. B.; Fischer, C.; Arutyunova, E.; Lamer, T.; Shields, J.; Saffran, H. A.; McKay, R. T.; van Belkum, M. J.; Joyce, M. A.; Young, H. S.; Tyrrell, D. L.; Vederas, J. C.; Lemieux, M. J. Feline Coronavirus Drug Inhibits the Main Protease of SARS-CoV-2 and Blocks Virus Replication. *Nat. Comm.* **2020**, *11*, 4282.

(47) Sakamoto, J.; Hamada, C.; Rahman, M.; Kodaira, S.; Ito, K.; Nakazato, H.; Ohashi, Y.; Yasutomi, M. An Individual Patient Data Meta-Analysis of Adjuvant Therapy with Carmofur in Patients with Curatively Resected Colon Cancer. *Jap. J. Clin. Onc.* **2005**, *35*, 536–544. https://doi.org/10.1093/jjco/hyi147.

(48) Owen, D. R.; Allerton, C. M. N.; Anderson, A. S.; Aschenbrenner, L.; A, M.; Berritt, S.; Boras, B.; Cardin, R. D.; Carlo, A.; Coffman, K. J.; Dantonio, A.; Di, L.; Eng, H.; Ferre, R.; Gajiwala, K. S.; Gibson, S. A.; Greasley, S. E.; Hurst, B. L.; Kadar, E. P.; Kalgutkar, A. S.; Lee, J. C.; Lee, J.; Liu, W.; Mason, S. W.; Noell, S.; Novak, J. J.; Obach, R. S.; Ogilvie, K.; Patel, N. C.; Pettersson, M.; Rai, D. K.; Reese, M. R.; Sammons, M. F.; Sathish, J. G.; Singh, R. S. P.; Steppan, C. M.; Stewart, A. E.; Tuttle, J. B.; Updyke, L.; Verhoest, P. R.; Wei, L.; Yang, Q.; Zhu, Y. An Oral SARS-CoV-2 Mpro Inhibitor Clinical Candidate for the Treatment of COVID-19. *Science* **2021**, *374*, 1586–1593.

(49) Ahmad, B.; Batool, M.; Ain, Q. U.; Kim, M. S.; Choi, S. Exploring the Binding Mechanism of PF-07321332 SARS-CoV-2 Protease Inhibitor through Molecular Dynamics and Binding Free Energy Simulations. *Int. J. Mol. Sci.* **2021**, *22*, 9124.

(50) Andrianov, A. M.; Kornoushenko, Y. V; Karpenko, A. D.; Bosko, I. P.; Tuzikov, A. V. Computational Discovery of Small Drug-like Compounds as Potential Inhibitors of SARS-CoV-2 Main Protease. *J. Biomol. Struct. Dyn.* **2021**, *39* (15), 5779–5791.

(51) Welborn, V. V. Beyond Structural Analysis of Molecular Enzyme-Inhibitor Interactions. *El. Struct.* **2022**, *4*, 14006.

(52) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.

(53) Richter, F.; Blomberg, R.; Khare, S. D.; Kiss, G.; Kuzin, A. P.; Smith, A. J. T.; Gallaher, J.; Pianowski, Z.; Helgeson, R. C.; Grjasnow, A.; Xiao, R.; Seetharaman, J.; Su, M.; Vorobiev, S.; Lew, S.; Forouhar, F.; Kornhaber, G. J.; Hunt, J. F.; Montelione, G. T.; Tong, L.; Houk, K. N.; Hilvert, D.; Baker, D. Computational Design of Catalytic Dyads and Oxyanion Holes for Ester Hydrolysis. *J. Am. Chem. Soc.* **2012**, *134* (39), 16197–16206.

(54) Paasche, A.; Schirmeister, T.; Engels, B. Benchmark Study for the Cysteine–Histidine Proton Transfer Reaction in a Protein Environment: Gas Phase, COSMO, QM/MM Approaches. *J. Chem. Theory Comput.* **2013**, *9* (3), 1765–1777.

(55) Paasche, A.; Zipper, A.; Schäfer, S.; Ziebuhr, J.; Schirmeister, T.; Engels, B. Evidence for Substrate Binding-Induced Zwitterion Formation in the Catalytic Cys-His Dyad of the SARS-CoV Main Protease. *Biochemistry* **2014**, *53* (37), 5930–5946.

(56) Elsässer, B.; Zauner, F. B.; Messner, J.; Soh, W. T.; Dall, E.; Brandstetter, H. Distinct Roles of Catalytic Cysteine and Histidine in the Protease and Ligase Mechanisms of Human Legumain As Revealed by DFT-Based QM/MM Simulations. *ACS Catal* **2017**, *7* (9), 5585–5593. https://doi.org/10.1021/acscatal.7b01505.

(57) Zanetti-Polzi, L.; Smith, M. D.; Chipot, C.; Gumbart, J. C.; Lynch, D. L.; Pavlova, A.; Smith, J. C.; Daidone, I. Tuning Proton Transfer Thermodynamics in SARS-CoV-2 Main Protease: Implications for Catalysis and Inhibitor Design. *J. Phys. Chem. Lett.* **2021**, *12* (17), 4195–4202.

(58) Aprá, E.; Bylaska, E. J.; de Jong, W. A.; Govind, N.; Kowalski, K.; Straatsma, T. P.; Valiev, M.; van Dam, H. J. J.; Alexeev, Y.; Anchell, J.; Anisimov, V.; Aquino, F. W.; Atta-Fynn, R.; Autschbach, J.; Bauman, N. P.; Becca, J. C.; Bernholdt, D. E.; Bhaskaran-Nair, K.; Bogatko, S.; Borowski, P.; Boschen, J.; Brabec, J.; Bruner, A.; Cauët, E.; Chen, Y.; Chuev, G. N.; Cramer, C. J.; Daily, J.; Deegan, M. J. O.; Dunning, T. H.; Dupuis, M.; Dyall, K. G.; Fann, G. I.; Fischer, S. A.; Fonari, A.; Früchtl, H.; Gagliardi, L.; Garza, J.; Gawande, N.; Ghosh, S.; Glaesemann, K.; Götz, A. W.; Hammond, J.; Helms, V.; Hermes, E. D.; Hirao, K.; Hirata, S.; Jacquelin, M.; Jensen, L.; Johnson, B. G.; Jónsson, H.; Kendall, R. A.; Klemm, M.; Kobayashi, R.; Konkov, V.; Krishnamoorthy, S.; Krishnan, M.; Lin, Z.; Lins, R. D.; Littlefield, R. J.; Logsdail, A. J.; Lopata, K.; Ma, W.; Marenich, A. V; del Campo, J. M.; Mejia-Rodriguez, D.; Moore, J. E.; Mullin, J. M.; Nakajima, T.; Nascimento, D. R.; Nichols, J. A.; Nichols, P. J.; Nieplocha, J.; Otero-de-la-Roza, A.; Palmer, B.; Panyala, A.; Pirojsirikul, T.; Peng, B.; Peverati, R.; Pittner, J.; Pollack, L.; Richard, R. M.; Sadayappan, P.; Schatz, G. C.; Shelton, W. A.; Silverstein, D. W.; Smith, D. M. A.; Soares, T. A.; Song, D.; Swart, M.; Taylor, H. L.; Thomas, G. S.; Tipparaju, V.; Truhlar, D. G.; Tsemekhman, K.; Van Voorhis, T.; Vázquez-Mayagoitia, Á.; Verma, P.; Villa, O.; Vishnu, A.; Vogiatzis, K. D.; Wang, D.; Weare, J. H.; Williamson, M. J.; Windus, T. L.; Woliński, K.; Wong, A. T.; Wu, Q.; Yang, C.; Yu, Q.; Zacharias, M.; Zhang, Z.; Zhao, Y.; Harrison, R. J. NWChem: Past, Present, and Future. *J. Chem. Phys.* **2020**, *152* (18), 184102. https://doi.org/10.1063/5.0004997.

(59) Krylov, A. I.; Gill, P. M. W. Q-Chem: An Engine for Innovation. *WIREs Comput. Mol. Sci.* **2013**, *3*, 317–326.

(60) Epifanovsky, E.; Gilbert, T. B.; Feng, X.; Lee, J.; Mao, Y.; Mardirossian, N.; Pokhilko, P.; White, A. F.; Coons, M. P.; Dempwolff, A. L.; Gan, Z.; Hait, D.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Kussmann, J.; Lange, A. W.; Lao, K. U.; Levine, D. S.; Liu, J.; McKenzie, S. C.; Morrison, A. F.; Nanda, K. D.; Plasser, F.; Rehn, D. R.; Vidal, M. L.; You, Z.-Q.; Zhu, Y.; Alam, B.; Albrecht, B. J.; Aldossary, A.; Alguire, E.; Andersen, J. H.; Athavale, V.; Barton, D.; Begam, K.; Behn, A.; Bellonzi, N.; Bernard, Y. A.; Berquist, E. J.; Burton, H. G. A.; Carreras, A.; Carter-Fenk, K.; Chakraborty, R.; Chien, A. D.; Closser, K. D.; Cofer-Shabica, V.; Dasgupta, S.; de Wergifosse, M.; Deng, J.; Diedenhofen, M.; Do, H.; Ehlert, S.; Fang, P.-T.; Fatehi, S.; Feng, Q.; Friedhoff, T.; Gayvert, J.; Ge,

Q.; Gidofalvi, G.; Goldey, M.; Gomes, J.; González-Espinoza, C. E.; Gulania, S.; Gunina, A. O.; Hanson-Heine, M. W. D.; Harbach, P. H. P.; Hauser, A.; Herbst, M. F.; Hernández Vera, M.; Hodecker, M.; Holden, Z. C.; Houck, S.; Huang, X.; Hui, K.; Huynh, B. C.; Ivanov, M.; Jász, A.; Ji, H.; Jiang, H.; Kaduk, B.; Kähler, S.; Khistyaev, K.; Kim, J.; Kis, G.; Klunzinger, P.; Koczor-Benda, Z.; Koh, J. H.; Kosenkov, D.; Koulias, L.; Kowalczyk, T.; Krauter, C. M.; Kue, K.; Kunitsa, A.; Kus, T.; Ladjánszki, I.; Landau, A.; Lawler, K. V; Lefrancois, D.; Lehtola, S.; Li, R. R.; Li, Y.-P.; Liang, J.; Liebenthal, M.; Lin, H.-H.; Lin, Y.-S.; Liu, F.; Liu, K.-Y.; Loipersberger, M.; Luenser, A.; Manjanath, A.; Manohar, P.; Mansoor, E.; Manzer, S. F.; Mao, S.-P.; Marenich, A. V; Markovich, T.; Mason, S.; Maurer, S. A.; McLaughlin, P. F.; Menger, M. F. S. J.; Mewes, J.-M.; Mewes, S. A.; Morgante, P.; Mullinax, J. W.; Oosterbaan, K. J.; Paran, G.; Paul, A. C.; Paul, S. K.; Pavošević, F.; Pei, Z.; Prager, S.; Proynov, E. I.; Rák, A.; Ramos-Cordoba, E.; Rana, B.; Rask, A. E.; Rettig, A.; Richard, R. M.; Rob, F.; Rossomme, E.; Scheele, T.; Scheurer, M.; Schneider, M.; Sergueev, N.; Sharada, S. M.; Skomorowski, W.; Small, D. W.; Stein, C. J.; Su, Y.-C.; Sundstrom, E. J.; Tao, Z.; Thirman, J.; Tornai, G. J.; Tsuchimochi, T.; Tubman, N. M.; Veccham, S. P.; Vydrov, O.; Wenzel, J.; Witte, J.; Yamada, A.; Yao, K.; Yeganeh, S.; Yost, S. R.; Zech, A.; Zhang, I. Y.; Zhang, X.; Zhang, Y.; Zuev, D.; Aspuru-Guzik, A.; Bell, A. T.; Besley, N. A.; Bravaya, K. B.; Brooks, B. R.; Casanova, D.; Chai, J.-D.; Coriani, S.; Cramer, C. J.; Cserey, G.; DePrince, A. E.; DiStasio, R. A.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; Goddard, W. A.; Hammes-Schiffer, S.; Head-Gordon, T.; Hehre, W. J.; Hsu, C.-P.; Jagau, T.-C.; Jung, Y.; Klamt, A.; Kong, J.; Lambrecht, D. S.; Liang, W.; Mayhall, N. J.; McCurdy, C. W.; Neaton, J. B.; Ochsenfeld, C.; Parkhill, J. A.; Peverati, R.; Rassolov, V. A.; Shao, Y.; Slipchenko, L. V; Stauch, T.; Steele, R. P.; Subotnik, J. E.; Thom, A. J. W.; Tkatchenko, A.; Truhlar, D. G.; Van Voorhis, T.; Wesolowski, T. A.; Whaley, K. B.; Woodcock, H. L.; Zimmerman, P. M.; Faraji, S.; Gill, P. M. W.; Head-Gordon, M.; Herbert, J. M.; Krylov, A. I. Software for the Frontiers of Quantum Chemistry: An Overview of Developments in the Q-Chem 5 Package. *J. Chem. Phys.* **2021**, *155* (8), 84801. https://doi.org/10.1063/5.0055522.

(61) Giudetti, G.; Polyakov, I.; Grigorenko, B. L.; Faraji, S.; Nemukhin, A. V; Krylov, A. I. How Reproducible Are QM/MM Simulations? Lessons from Computational Studies of the Covalent Inhibition of the SARS-CoV-2 Main Protease by Carmofur. *J. Chem. Theory Comput.* **2022**, *18*, 5056-5067.

(62) Adamo, C.; Barone, V. Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0 Model. *J. Chem. Phys.* **1999**, *110*, 6158–6170.

(63) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32* (7), 1456–1465. https://doi.org/https://doi.org/10.1002/jcc.21759.

(64) Mardirossian, N.; Head-Gordon, M. Thirty Years of Density Functional Theory in Computational Chemistry: An Overview and Extensive Assessment of 200 Density Functionals. *Mol. Phys.* **2017**, *115* (19), 2315–2372. https://doi.org/10.1080/00268976.2017.1333644.

(65) Ponder, J. W.; Case, D. A. Force Fields for Protein Simulations. *Adv. Protein Chem.* **2003**, *66*, 27–85. https://doi.org/10.1016/S0065-3233(03)66002-X.

(66) Lee, T. S.; Cerutti, D. S.; Mermelstein, D.; Lin, C.; Legrand, S.; Giese, T. J.; Roitberg, A.; Case, D. A.; Walker, R. C.; York, D. M. GPU-Accelerated Molecular Dynamics and Free Energy Methods in Amber18: Performance Enhancements and New Features. *J. Chem. Inf. Model.* **2018**, *58* (10), 2043–2050. https://doi.org/10.1021/acs.jcim.8b00462.

(67) Kästner, J.; Thiel, W. Bridging the Gap between Thermodynamic Integration and Umbrella Sampling Provides a Novel Analysis Method: "Umbrella Integration." *J. Chem. Phys.* **2005**, *123* (14), 144104.

(68) Fiorin, G.; Klein, M. L.; Hénin, J. Using Collective Variables to Drive Molecular Dynamics Simulations. *Mol. Phys.* **2013**, *111* (22–23), 3345–3362.

(69) Melo, M. C. R.; Bernardi, R. C.; Rudack, T.; Scheurer, M.; Riplinger, C.; Phillips, J. C.; Maia, J. D. C.; Rocha, G. B.; Ribeiro, J. V; Stone, J. E.; Neese, F.; Schulten, K.; Luthey-Schulten, Z. NAMD Goes Quantum: An Integrative Suite for Hybrid Simulations. *Nat. Methods* **2018**, *15* (5), 351–354.

(70) Phillips, J. C.; Hardy, D. J.; c. Maia, J. D.; Stone, J. E.; Ribeiro, J. V; Bernardi, R. C.; Buch, R.; Fiorin, G.; Hénin, J.; Jiang, W.; McGreevy, R.; Melo, M. C. R.; Radak, B. K.; Skeel, R. D.; Singharoy, A.; Wang, Y.; Roux, B.; Aksimentiev, A.; Luthey-Schulten, Z.; Kalé, L. V; Schulten, K.; Chipot, C.; Tajkhorshid, E. Scalable Molecular Dynamics on CPU and GPU Architectures with NAMD. *J. Chem. Phys.* **2020**, *153* (4), 44130.

(71) Khrenova, M. G.; Tsirelson, V. G.; Nemukhin, A. V. Dynamical Properties of Enzyme–Substrate Complexes Disclose Substrate Specificity of the SARS-CoV-2 Main Protease as Characterized by the Electron Density Descriptors. *Phys. Chem. Chem. Phys.* **2020**, *22*, 19069–19079.

(72) Khrenova, M. G.; Grigorenko, B. L.; Nemukhin, A. V. Molecular Modeling Reveals the Mechanism of Ran-RanGAP-Catalyzed Guanosine Triphosphate Hydrolysis without an Arginine Finger. *ACS Cat.* **2021**, *11* (15), 8985–8998.

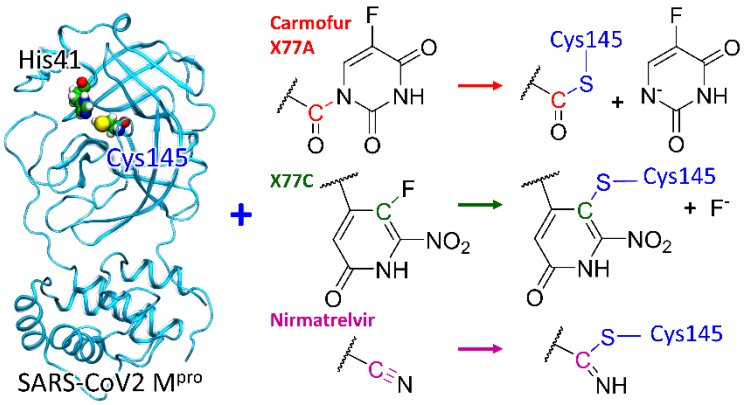
(73) Khrenova, M. G.; Kulakova, A. M.; Nemukhin, A. V. Light-Induced Change of Arginine Conformation Modulates the Rate of Adenosine Triphosphate to Cyclic Adenosine Monophosphate Conversion in the Optogenetic System Containing Photoactivated Adenylyl Cyclase. *J. Chem. Inf. Mod.* **2021**, *61* (3), 1215–1225.

(74) Seritan, S.; Bannwarth, C.; Fales, B. S.; Hohenstein, E. G.; Isborn, C. M.; Kokkila-Schumacher, S. I. L.; Li, X.; Liu, F.; Luehr, N.; Snyder Jr., J. W.; Song, C.; Titov, A. V; Ufimtsev, I. S.; Wang, L.; Martínez, T. J. TeraChem: A Graphical Processing Unit-Accelerated Electronic Structure Package for Large-Scale Ab Initio Molecular Dynamics. *WIREs Comput. Mol. Sci.* **2021**, *11* (2), e1494.

(75) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ, ψ and Side-Chain χ_1 and χ_2 Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8* (9), 3257–3273. https://doi.org/10.1021/ct300400x.

(76) Lennox, A. J. J. Meisenheimer Complexes in S_NAr Reactions: Intermediates or Transition States? *Angew. Chem., Int. Ed.* **2018**, *57*, 2–5.

(77) Senger, N. A.; Bo, B.; Cheng, Q.; Keeffe, J. R.; Gronert, S.; Wu, W. The Element Effect Revisited: Factors Determining Leaving Group Ability in Activated Nucleophilic Aromatic Substitution Reactions. *J. Org. Chem.* **2012**, *77* (21), 9535–9540.



TOC graphics