Insight into the Binding Energetics of Targeted Reversible Covalent Inhibitors of the SARS-CoV-2 Main Protease

Ernest Awoonor-Williams *

† Department of Chemistry, Memorial University of Newfoundland, St. John's, NL, A1B 3X9, Canada.

* Corresponding author. 

Absolute binding Free Energy, QM/MM, SARS-CoV-2, Mpro or 3CLpro, Reversible Covalent Inhibitors

ABSTRACT: The main protease (Mpro) of the SARS-CoV-2 virus is an attractive therapeutic target for developing antivirals to combat COVID-19. Mpro is essential for the replication cycle of the SARS-CoV-2 virus, so inhibiting Mpro blocks a vital piece of the cell replication machinery of the virus. A promising strategy to disrupt the viral replication cycle is to design inhibitors that bind to the active site cysteine (Cys145) of the Mpro. Cysteine targeted covalent inhibitors are gaining traction in drug discovery owing to the benefits of improved potency and extended drug-target engagement. An interesting aspect of these inhibitors is that they can be chemically tuned to form a covalent, but reversible bond, with their targets of interest. Several small-molecule cysteine-targeting covalent inhibitors of the Mpro have been discovered—some of which are currently undergoing evaluation in early phase human clinical trials. Understanding the binding energetics of these inhibitors could provide new insights to facilitate the design of potential drug candidates against COVID-19. Motivated by this, we employed rigorous absolute binding free energy calculations and hybrid quantum mechanical/molecular mechanical (QM/MM) calculations to estimate the energetics of binding of some promising reversible covalent inhibitors of the Mpro. We find that the inclusion of enhanced sampling techniques such as replica-exchange algorithm in binding free energy calculations can improve the convergence of predicted non-covalent binding free energy estimates of inhibitors binding to the Mpro target. In addition, our results indicate that binding free energy calculations coupled with multiscale simulations can be a useful approach to employ in ranking covalent inhibitors to their targets. This approach may be valuable in prioritizing and refining covalent inhibitor compounds for lead discovery efforts against COVID-19 and future coronavirus infections.

Introduction

Despite the recent success of multiple vaccines in the prevention of COVID-19, there is still an urgent need to develop targeted therapeutics for the treatment of COVID-19. This is imperative given the emergence of rapidly spreading new variants of the SARS-CoV-2 virus—causative agent of COVID-19, which threatens to compromise vaccine efficacy. To date, the nucleoside analog Veklury (remdesivir) remains the only antiviral drug approved by the U.S. Food and Drug Administration for treatment of patients with COVID-19 requiring hospitalization—although no meaningful mortality benefit to patients has been shown. This presents a serious unmet medical need in COVID-19 treatment given the significant global impact of the disease.

Among the group of SARS-CoV-2 protein targets, the main protease (Mpro)—also known as 3-chymotrypsin-like protease (3CLpro), has emerged as an attractive target for the development of antivirals to combat COVID-19.6,6 Mpro is a highly conserved protein that is essential for the viral replication cycle. Mpro has unique substrate specificity and cleaves the viral polyprotein at 11 conserved sites for viral replication and transcription.7 There are no human homologs of the Mpro with the same specificity, so inhibiting Mpro offers reduced risk for off-target reactivity and represents a suitable therapeutic strategy in the development of pan-coronavirus drugs.

Several small-molecule inhibitors of the SARS-CoV-2 Mpro have been discovered,12–17 with a few progressing into early phase human clinical trials.18–21 The fundamental aim of these inhibitors is to disrupt the viral life cycle. One therapeutic strategy is to design inhibitors with electrophilic functional groups (a.k.a., “warhead”) that bind to the nucleophilic active site cysteine (Cys145) of the Mpro target. Cys145 of the Mpro is essential for catalytic activity, so blockage or modification of this residue is detrimental to the virus. This strategy of targeted covalent modification of druggable sites in disease targets is gaining traction in drug discovery,22–25 as evidenced by the successful number of approved drugs employing this approach (e.g., ibrutinib).26 Importantly, covalent inhibitors offer superior potencies, prolonged duration of therapeutic action, and high target selectivity than their non-covalent drug binding counterparts. Also, since such covalent binding events can be chemically tuned to be reversible,27 concerns regarding the risk of potential off-target reactivity and toxicity can be mitigated while maintaining potency and efficacy.

Understanding the binding energetics and interactions of inhibitors with their targets offers insight at the molecular level to facilitate drug design and discovery. Over 1000 high-resolution crystallographic structures of SARS-CoV-2 proteins have been determined,28 which has enabled structure-guided drug design efforts towards the development of potent compounds as therapeutic leads in the fight against COVID-19. In silico methods such as virtual screening, molecular dynamics, and binding free energy calculations have provided structural insight into the binding interactions, molecular
mechanism, and conformational dynamics of inhibitors binding to the M\textsuperscript{pro} target to inform COVID-19 drug discovery.\textsuperscript{31} In our own work, we have showed that alchemical absolute binding free energy (ABFE) calculations coupled with multiscale quantum chemical calculations can predict the binding energies of covalent inhibitors to the M\textsuperscript{pro} target reasonably well.\textsuperscript{32}

Building upon our previous work, we employ rigorous binding free energy calculations to explore the binding energetics of reversible covalent inhibitors of the SARS-CoV-2 M\textsuperscript{pro} target. Reversible covalent inhibitors form covalent but reversible bonds with their target enzyme (Figure 1(a)) and are considered safer alternatives to conventional irreversible covalent inhibitors due to less risk for idiosyncratic toxicities. The inhibitors studied include 11a,\textsuperscript{11} GC373,\textsuperscript{14} MI-30,\textsuperscript{16} and PF-00835231 (‘PF-231’).\textsuperscript{17} Figure 1(b). These inhibitors contain either an aldehyde or ketone functional group as electrophilic warhead and can reversibly bind to the catalytic Cys145 of the M\textsuperscript{pro}.\textsuperscript{14,17} Forming thiohemiacetal/thiohemiketal tetrahedral complexes. To calculate the absolute binding free energies of the inhibitors to the M\textsuperscript{pro} target, we employed alchemical free energy calculations and ONIOM QM/MM calculations on model protein-ligand complexes of the system. The alchemical ABFE calculations were performed using both non-enhanced and enhanced replica-exchange sampling techniques within the equilibrium free energy framework.\textsuperscript{33} Our results indicate that enhanced free energy sampling methods via the Hamiltonian replica-exchange algorithm (H-REMD)\textsuperscript{34} improved the convergence and standard deviation of the binding free energy predictions. These results provide valuable insight into the binding energetics of reversible covalent inhibitors to the M\textsuperscript{pro} target and highlights an important case where enhanced sampling techniques can improve the convergence of computed free energy estimates. This has the potential to assist in the rational design and evaluation of novel chemical entities, as well as support decisions in lead optimization campaigns towards identifying potent antiviral leads for current and future coronavirus diseases.

![Figure 1. (a) Classical representation of the mechanism of reversible covalent modification of an enzyme by inhibitor highlighting the kinetic inhibition parameters: \( K_i \) and \( k_{\text{cat},\text{act}} \). (b) Examples of covalent inhibitors of SARS-CoV-2 M\textsuperscript{pro} under development for treatment of COVID-19. The warhead of the inhibitors is indicated in red, with the reactive center highlighted by an asterisk.](image-url)

### Methods

#### System Setup:
Model structures of the protein–ligand complexes were downloaded from the Protein Data Bank.\textsuperscript{35} The PDB accession codes for the complexes studied include 6LZE,\textsuperscript{36} 7D1M,\textsuperscript{37} 7D31,\textsuperscript{38} and 6XHM,\textsuperscript{39} for 11a, GC373, MI-30, and PF-231, respectively. The covalent bond between the warhead of the inhibitor and active-site cysteine (Cys145) of the protein were broken prior to running the molecular dynamics (MD) simulations. The M\textsuperscript{pro} amino acid residues were assigned protonation states corresponding to their default values at neutral pH. We have previously evaluated methods for the calculation of the pK\textsubscript{a} of titratable cysteine residues in proteins\textsuperscript{16} and have performed MD simulations to estimate the protonation state of the catalytic Cys145/His41 dyad of the M\textsuperscript{pro}.\textsuperscript{32} Based on our results and the work of others,\textsuperscript{37–39} we treated the catalytic Cys145/His41 dyad of the M\textsuperscript{pro} as neutral residues. We also modeled His41 residue in two tautomeric states (i.e., N-\( \delta \) and N-\( \epsilon \) protonated states) to explore the preferred protonation state. The protonation state of histidines in SARS-CoV-2 M\textsuperscript{pro}, in particular His41, has been shown to impact ligand binding and protein structure dynamics.\textsuperscript{40} The GROMACS MD software\textsuperscript{41} (version 2020.4) was used to perform the calculations. Model structures of the protein–ligand complexes were prepared in explicit solvent using the TIP3P\textsuperscript{42} water model. The AMBER99SB-ILDNP\textsuperscript{43} and GAFF\textsuperscript{44} force fields were used to generate parameters for the protein and ligands, respectively. Ligand charges were parameterized using AM1-BCC.\textsuperscript{45,46} The model systems were solvated in a dodecahedral box with a maximum cutoff distance of 12 Å from the edge of the box. The simulation cell was charge neutralized with an appropriate number of Na\textsuperscript{+} ions. Following this, energy minimization and 2-ns equilibration runs (using both NVT and NpT ensembles) were performed on the system. This was subsequently followed by 10-ns simulation of the complex. The MD simulations were performed to determine the preferred protonation state of His41 in the model complexes prior to running the binding free energy simulations.

#### Absolute Binding Free Energy Calculation:
We performed ABFE calculations to compute the non-covalent binding free energy contribution of the inhibitors to the M\textsuperscript{pro} target. ABFE calculations, although computationally expensive, provide a theoretically rigorous framework\textsuperscript{37} for estimating ligand binding energies and have been shown to yield results that agree well with experimental binding energies.\textsuperscript{47–49} For performing the ABFE calculations, we followed the protocol reported in our recent publication\textsuperscript{30} on estimating the total binding energy of peptidomimetic inhibitors of the SARS-CoV-2 M\textsuperscript{pro}. Two sets of simulations for each model system were performed for the binding energy calculations, consisting of the ligand in bulk solution and in the protein-binding site. A set of restraints defined by one distance, two angles, and three dihedral harmonic potentials were imposed on the bound ligand to keep it in the protein-binding site following the decoupling steps of the free energy calculation (see Supporting Information for details). For the ligand in bulk solution, the restraint term following the decoupling event is computed analytically using the method proposed by Boresch et al.\textsuperscript{50} Similar to our previous protocol,\textsuperscript{32} 31 \( \lambda \) windows were used for the ligand in bulk solution and 42 \( \lambda \) windows were used for the protein–ligand simulation. For each \( \lambda \) window, a series of simulations consisting of 100,000-step steepest descent energy minimization, 1ns of NVT equilibration and 1 ns of NPT equilibration were performed. A timestep of 2 fs was used for the simulations with reference temperature set to 298.15 K. The LINCS constraint algorithm was used to
constrain H-bonds. The Particle Mesh Ewald (PME) algorithm\textsuperscript{4,55} was used to treat long-range electrostatic interactions and a cut-off distance of 12 Å was used for short-range electrostatics.

Following the equilibration simulations, a series of 12 ns production runs were performed for each \( \lambda \) window under isothermal-isobaric ensemble conditions using the Parrinello–Rahman pressure coupling scheme.\textsuperscript{46} The simulations were performed with and without enhanced sampling using the Hamiltonian replica-exchange MD (H-REMD) approach. All simulations were performed in triplicate to estimate the uncertainty in the calculations and to check for convergence in the results. For each protein–ligand model system, an aggregate total of \textasciitilde 5.5 \( \mu \)s in sampling time was achieved. The results of the simulations were analyzed using the alchemical_analysis.py python tool,\textsuperscript{57} and were calculated using multistate Bennet acceptance ratio (MBAR)\textsuperscript{58} method after discarding the first 2 ns from each \( \lambda \) window as equilibration. The reported binding free energy results are the averages and standard deviation of the three independent repeats.

**ONIOM(QM:MM) Methodology:** To estimate the energy leading to the formation of the covalent bond between the target Cys145 of the M\textsuperscript{Mm} and the reactive center of the inhibitors, we performed QM/MM calculations on the complexes using the ONIOM\textsuperscript{49} approach implemented in Gaussian 16.\textsuperscript{60} Explicit solvent water molecules were excluded from the model system. For the QM/MM ONIOM model, we partitioned the system into two layers: (1) a High-level QM region consisting of the full inhibitor structure, Cys145 and His41 side chains (2) Low-level molecular mechanical region which accounts for the remaining residues of the system. On average, our high-level QM region consisted of 74–82 atoms, while the low-level MM region contained 4599–4649 atoms. We treated the high-level region using the M06-2X functional with the def2-TZVP basis set—which performs well for modeling the thermochemistry\textsuperscript{41} and reaction energies\textsuperscript{42,63} of chemical reactions. The low-level region on the other hand, was treated with the amber molecular mechanical force field. Missing ligand atom parameters were obtained from the generalized Amber force field (GAFF).\textsuperscript{64} Hydrogen link atoms were used to cap the QM–MM boundary region. Figure 2 shows a representative snapshot of the atoms comprising the QM region for a selected model system.

The ONIOM model structures were geometry optimized using the M06-2X/def2-TZVP:AMBER method within an electronic embedding scheme. We also performed frequency analyses on the optimized geometries using the same level of theory to verify that the optimized structures were minima and to calculate the Gibbs energy.

**Results and Discussion**

**His41 Protonation State**

The active-site of the SARS-CoV-2 M\textsuperscript{Mm} consists of catalytic dyad residues (His41/Cys145) which are responsible for the proteolytic activity and function of the protease.\textsuperscript{65} Among the residues involved in the enzymatic activity of the M\textsuperscript{Mm}, His41 is suggested to act as a proton carrier of the thiol hydrogen of Cys145—enabling nucleophilic attack on the reactive center of the inhibitor by the activated Cys145 thiolate formed. Recent work by Gumbart and coworkers have shown that the protonation state of His41 of the M\textsuperscript{Mm} is sensitive to the presence of a ligand,\textsuperscript{60} which can influence the structural stability and dynamics of the M\textsuperscript{Mm} complex. Therefore, for each of the protein–ligand complexes studied, we investigated the preferred protonation state for His41 to inform our binding energy simulations.

Prior to running the binding free energy calculations, we performed 10-ns MD simulations on the model complexes to determine the preferred protonation state of His41. The protonation states of His41 in the model structures were assigned either as HID (N\textsubscript{a} protonated) or HIE (N\textsubscript{e} protonated), according to amber force field naming convention. All the other histidines were left unaltered in their default protonation states as chosen by the pdb2gmx program of GROMACS MD engine (see Supporting Information). The difference in structure and binding interactions of the inhibitors with the M\textsuperscript{Mm} target in the structures were analyzed upon detailed visual inspection of their contacts. Also, the root-mean-square deviation, or RMSD, between select corresponding atoms (\( \alpha \)-carbon of the protein and heavy atoms of the ligand) in the simulated complexes and the crystallographic starting structures were measured after alignement of the structures. Our results indicate that the inhibitor-bound complexes of 11a, GC373, and MI-30 favour the HID states of His41, while the PF-231 inhibitor-bound complex favours the HIE state of His41. Figure 3 depicts snapshots of the structural environment of the catalytic dyad configurations in the M\textsuperscript{Mm} target complexes taken from the equilibrated MD simulation.

The distance between the carbonyl carbon of the inhibitors and the cysteine S atom is around 3 Å for all the equilibrated inhibitor-bound M\textsuperscript{Mm} complexes (Figure 3). For both 11a and GC373 M\textsuperscript{Mm} complexes, the thiol side chain of Cys145 forms a direct hydrogen bonding interaction with the N\textsubscript{a} of His41 (Fig.3; Top). On the other hand, for MI-30 and PF-231 M\textsuperscript{Mm} complexes, we observe a hydrogen bond between Cys145 thiol side chain and a nearby water molecule (Fig.3; Bottom). In fact, studies have shown that both His41 and conserved catalytic water molecules in the protein active site play an active role in the reaction mechanism of the M\textsuperscript{Mm}.\textsuperscript{38,56–70} As a minor caveat, only the protonation states of His41 was explored in our model complexes due to its significant role in the binding mechanism.\textsuperscript{68} Binding free energy simulations that can model the dynamic variation in residue protonation and tautomeric states will be useful to explore the influence of other titratable residues— although not the focus of this study.
Non-covalent Binding Free Energy Contribution

The non-covalent binding free energy contribution of the inhibitors to the M\textsuperscript{exo} target was computed using alchemical absolute binding free energy calculations. Alchemical ABFE calculations provide a direct way of computing ligand binding energies within a rigorous statistical thermodynamic framework.\textsuperscript{46} This approach employs a non-physical thermodynamic cycle to compute inhibitor binding free energies and considers the conformational entropy and dynamics of both the protein target and inhibitor upon binding.\textsuperscript{75,76,72} ABFE calculations can yield highly accurate binding affinities for ligands that agree very well with experiment.\textsuperscript{76-80,84,85}\textsuperscript{35} Recently, we employed this approach to estimate the binding affinity of the peptidomimetic inhibitors N3 and α-ketoamide of the SARS-CoV-2 M\textsuperscript{exo}.\textsuperscript{32}

In this work, we calculated the absolute binding free energies of the inhibitors to the M\textsuperscript{exo} target using both non-enhanced and enhanced sampling techniques within the equilibrium ABFE framework. For the enhanced ABFE simulations, a Hamiltonian replica exchange molecular dynamics (H-REMD) algorithm\textsuperscript{34} was employed to facilitate sampling convergence in the free energy estimates. Table 1 provides a summary of the binding free energy results with and without H-REMD for the complexes studied in this work.

The results show that for each inhibitor–M\textsuperscript{exo} pair, both equilibrium approaches yield relatively similar results for the binding free energy estimates. However, the uncertainty in the computed binding free energy estimates is larger for the non-enhanced equilibrium ABFE approach (no H-REMD) relative to the enhanced sampling approach (H-REMD). This result is consistent with previous studies that have showed that the inclusion of the replica exchange algorithm in alchemical free energy simulations greatly improves the conformational sampling and convergence of free energy estimates.\textsuperscript{73,74}

Our binding free energy results indicate that the non-covalent binding free energy contributions of the inhibitors to the M\textsuperscript{exo} target is greatest for the PF-231 inhibitor (Table 1). This is followed by the MI-30 inhibitor, which has a predicted binding energy of ~7.5 kcal mol\textsuperscript{-1} to the M\textsuperscript{exo}. The inhibitors 11a and GC373 were predicted to have lower binding affinities in comparison to PF-231 and MI-30. Importantly, the predicted binding free energy results we report represent the non-covalent interaction energies between the inhibitors and the target protein. We were unable to find experimental inhibition constants (K\textsubscript{i}) for all the inhibitors to directly compare with our predicted binding free energy results. However, comparison to experimental IC\textsubscript{50} values for the inhibitors suggest that our predicted non-covalent binding free energy results correspond well with the relative potency rankings of the inhibitors to the M\textsuperscript{exo} (Figure 4).—although we note that IC\textsubscript{50}'s in general are not suitable descriptors for estimating the biochemical potency of covalent inhibitors.\textsuperscript{75-77} This is because IC\textsubscript{50}'s for covalent inhibitors will vary depending on the incubation time of the assay and do not provide any information about the affinity or chemical reactivity of a covalent inhibitor.

Table 1. Summary of the non-covalent binding free energy ($\Delta G_{\text{non-covalent}}$) results of the inhibitor-bound M\textsuperscript{exo} complexes studied in this work. All energies are reported in kcal mol\textsuperscript{-1}.

| Complex     | $\Delta G_{\text{non-covalent}}$ (no H-REMD) | $\Delta G_{\text{non-covalent}}$ (H-REMD) |
|-------------|--------------------------------------------|------------------------------------------|
| M\textsuperscript{exo}:11a | $-4.8 \pm 0.8$ | $-5.4 \pm 0.5$ |
| M\textsuperscript{exo}:GC373 | $-4.8 \pm 0.9$ | $-5.1 \pm 0.3$ |
| M\textsuperscript{exo}:MI-30 | $-7.6 \pm 0.7$ | $-7.5 \pm 0.2$ |
| M\textsuperscript{exo}:PF-231 | $-8.2 \pm 0.8$ | $-8.2 \pm 0.2$ |

Figure 4. Relative potency rankings of the inhibitors 11a, GC373, MI-30, and PF-231 based on experimental IC\textsubscript{50} results. The reported IC\textsubscript{50} values for 11a, GC373, and MI-30 are 0.053 ± 0.005 μM,\textsuperscript{79} 0.40 ± 0.05 μM,\textsuperscript{79} and 17.2 ± 0.6 nM,\textsuperscript{48} respectively. In the case of the PF-
231 inhibitor, no IC₅₀ value was reported so the Kᵢ of 0.27 ± 0.1 nM was used for the comparison.

Figure 5 shows a representation of the interaction diagram of the inhibitors in the active site of the Mₚᵣo. The inhibitors form extensive interactions in the target binding site by engaging in multiple hydrogen bonding interactions with key amino acid residues of the Mₚᵣo. An example we highlight is Glu166, whose backbone amino acid groups and carboxylate side chain establishes multiple hydrogen bonding contacts with functional groups of the inhibitors (Figure 5, right panel). Other hydrogen bonding residues include Gly143, His164, Cys145, Gln189, and Ser144. In addition to these hydrogen bond interactions, there are van der Waals and hydrophobic interactions between the Mₚᵣo residues and the inhibitors. Together, these interactions between the target protein and inhibitors contribute to the binding free energies of the inhibitors to the Mₚᵣo.

Covalent Binding Free Energy Contribution

Covalent inhibitors typically follow a two-step kinetic mechanism characterized by an equilibrium dissociation constant (Kᵢ) for the non-covalent E-I complex and an inactivation rate constant (kₐₑₓ) to form the covalent E–I complex, Figure 1(a). Reversible covalent inhibitors differ from irreversible inhibitors in that the step leading to the formation of the final covalent E–I complex is not permanent. Modeling the binding of covalent inhibitors with their targets requires characterization of both the non-covalent and covalent binding free energy steps. Multiscale QM/MM calculations are useful in this regard since they allow chemical processes to be modeled—particularly, the chemical bond formation step between the target residue in the protein and an inhibitor.

To estimate the covalent binding free energy contribution of the Mₚᵣo inhibitors studied, we performed ONIOM (QM:MM) calculations on the model systems. Using this approach, we calculated the relative Gibbs energy difference between the equilibrium non-covalent complex and the covalent complex. Our results for the covalent binding free energy of the inhibitors to the Mₚᵣo target are summarized in Table 2.

Table 2. Summary of the covalent binding free energy (ΔG°_covalent) results of the inhibitor bound Mₚᵣo complexes studied in this work.

| Complex             | ΔG°_covalent (kcal mol⁻¹) |
|---------------------|--------------------------|
| Mₚᵣo–11a            | −6.1                     |
| Mₚᵣo–GC373          | −8.2                     |
| Mₚᵣo–MI-30          | −9.8                     |
| Mₚᵣo–PF-231         | −6.8                     |

The results indicate that the reaction leading to the formation of the covalent bond between Cys145 of the Mₚᵣo and the inhibitors is moderately exergonic. For irreversible Michael acceptor inhibitors of the Mₚᵣo, covalent binding reaction energies as high as −28 kcal/mol have been predicted. We found that among the complexes studied, the reaction free energy for adduct formation is greatest for the MI-30 inhibitor bound complex (ΔG°_covalent = −10.5 kcal/mol). The covalent binding free energy calculated for the other adducts are −6.0 kcal/mol, −8.3 kcal/mol and −6.8 kcal/mol, for the inhibitors 11a, GC373, and PF-231, respectively (Table 2). These results suggest that the inhibitors 11a and PF-231 engage more reversibly towards the Mₚᵣo target than GC-373 and MI-30 inhibitors. The relative magnitude of the covalent binding free energy predicted for these inhibitors is in line with their reversible characteristics as observed by experiment. In addition, our predicted covalent binding energy results for the Mₚᵣo–PF231 and Mₚᵣo–11a inhibitor bound adducts are in agreement with recent multiscale QM/MM studies that investigated the mechanism of covalent inhibition of the Mₚᵣo target by these compounds. Interestingly, we observe different covalent binding energies for the inhibitors 11a, GC373, and MI-30 which share a common warhead (i.e., an aldehyde functional group).

This suggests that there are a range of factors beyond warhead type that can influence the binding energetics of reversible covalent inhibitors to their targets. Knowledge of both covalent and functional groups of the inhibitors are highlighted. Only amino acid residues within 3.5 Å of the inhibitors are displayed in the figure.
noncovalent binding free energy contributions is essential in determining the overall binding energetics of covalent inhibitors to their enzyme targets.

Conclusion

In this study, we performed rigorous absolute binding free energy simulations and ONIOM QM/MM calculations to estimate—from a thermodynamical standpoint—the energetics of binding of some promising reversible covalent inhibitors of the SARS-CoV-2 Mpro. The inhibitors studied include 11a, GC373, MI30, and PF-231. These inhibitors contain either an aldehyde or ketone moiety as warhead and form a covalent bond with the active site Cys145 of the Mpro. Cys145 is essential for the catalytic activity of the Mpro, so modification of this residue by a targeted covalent inhibitor will have deleterious effects on the function of the Mpro and the virus as a whole.

Our computational approach involves the calculation of both the covalent and non-covalent binding free energy contributions of inhibitors to the Mpro target. From an in silico perspective, understanding these two binding contributions is important for the efficient optimization and design of potent compounds. For estimating the noncovalent binding free energy contributions of the inhibitors to the Mpro target, we employed equilibrium binding absolute free energy simulations using both non-enhanced and enhanced sampling methods. Our results indicate that enhanced sampling methods via the replica exchange algorithm yield binding free energy estimates that are better converged and have a lower standard deviation than the non-enhanced methods. Hydrogen bonding and dispersion forces are significant contributors to the non-covalent binding free energy of the inhibitors to the Mpro target. Our calculation of the covalent binding free energy contribution of the inhibitors to the Mpro suggest that the reaction is moderately exergonic, which is in line with the reversible character of these inhibitors as reported by experiment.

We observe different covalent binding energies for Mpro inhibitors compounds sharing a common aldehyde warhead group, suggesting that multiple factors other than warhead type influence the binding energetics of reversible covalent inhibitors.

Finally, we note that the present work has some caveats that needs pointing out. We only explored different protonation states of the catalytic His41 in our simulations due to its critical role in the molecular mechanism. But in principle, all histidine protonation states should have been explored and considered. Nevertheless, we find that binding free energy estimates of the inhibitors to the Mpro target are impacted when different protonation/tautomeric states are chosen for the active site His41 residue. We caution selection of appropriate residue protonation and tautomeric states (in particular, for His) for ionizable residues when performing binding free energy studies of the Mpro target. Methods that can account for the dynamic variation in titratable residue and ligand protonation states at different solution pH will be valuable in this regard, enabling complex dynamic transitions in protein–ligand binding events to be modelled more accurately.

ACKNOWLEDGMENT

Computational resources were provided by Compute Canada consortia. Additional computational resources were provided by the Center for Health Informatics and Analytics (CHIA) of the Faculty of Medicine at Memorial University of Newfoundland. E.A-W thanks Professor Christopher Rowley for providing valuable feedback on this work.

ABBREVIATIONS

ABFE, absolute binding free energy; COVID-19, coronavirus disease 2019; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; Mpro or 3CLpro, main protease or 3-chymotrypsin-like protease; MD, molecular dynamics; QM/MM, quantum mechanics/molecular mechanics; ONIOM, Our own N-layered Integrated Molecular Orbital and molecular Mechanics.

REFERENCES

(1) Mascara, J. R.; Graham, B. S.; Fauci, A. S. SARS-CoV-2 Viral Variants—Tackling a Moving Target. JAMA 2021, 325 (17), 1621.
(2) Zhou, W.; Wang, W. Fast-Spreading SARS-CoV-2 Variants: Challenges to and New Design Strategies of COVID-19 Vaccines. Signal Transduct. Target. Ther. 2021, 6 (1), 1–6.
(3) Repurposed Antiviral Drugs for COVID-19 — Interim WHO Solidarity Trial Results. N. Engl. J. Med. 2021, 384 (6), 497–511.
(4) Gil, C.; Ginex, T.; Maestro, I.; Nozal, V.; Barrado-Gil, L.; Cuestra-Gelio, M. A.; Urquiza, J.; Ramirez, D.; Alonso, C.; Campillo, N. E.; et al. COVID-19: Drug Targets and Potential Mechanisms. J. Med. Chem. 2020, 63 (21), 12359–12386.
(5) Anand, K.; John, Z.; Parvesh, W.; Mesters, J. R.; Hilgenfeld, R. Coronavirus Main Proteinase (3CLpro) Structure: Basis for Design of Anti-SARS Drugs. Science (80). 2003, 300 (5626), 1763–1767.
(6) Ulrich, S.; Nitsche, C. The SARS-CoV-2 Main Protease as Drug Target. Bioorganic Med. Chem. Lett. 2020, 30 (17), 127377.
(7) Thiel, V.; Ivanov, K. A.; Putics, Á.; Hertzig, T.; Schelle, B.; Bayer, S.; Weißbrich, B.; Snijder, E. J.; Rabenaub, H.; Doerr, H. W.; et al. Mechanisms and Enzymes Involved in SARS Coronavirus Genome Expression. J. Gen. Virol. 2003, 84 (9), 2303–2315.
(8) Banerjee, R.; Perera, L.; Tillekeratne, L. M. V. Potential SARS-CoV-2 Main Protease Inhibitors. Drug Discov. Today 2021, 26 (3), 804–816.
(9) Yang, H.; Yang, J. A Review of the Latest Research on Mpro Targeting SARS-CoV Inhibitors. J. Med. Chem. 2021.
(10) 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 (7811), 289–293.
(11) Zhang, L.; Lin, D.; Sun, X.; Curth, U.; Drost, C.; Saueringer, L.; Becker, S.; Rox, K.; Hilgenfeld, R. Crystal Structure of SARS-CoV-2 Main Protease Provides a Basis for Design of Improved a Ketoamide Inhibitors. Science (80). 2020, 369 (6489), 409–412.
(12) Dai, W.; Zhang, B.; Jiang, X.-M.; Su, H.; Li, J.; Zhao, Y.; Xie, X.; Jin, Z.; Peng, Z.; Liu, F.; et al. Structure-Based Design of Antiviral Drug Candidates Targeting the SARS-CoV-2 Main Protease. Science (80). 2020, 369 (6497), 1331–1335.
(13) Ma, C.; Sacco, M. D.; Hurst, B.; Townsend, J. A.; Hu, Y.; Seeto, T.; Zhang, X.; Tarbet, B.; Marty, M. D.; Chen, Y.; et al. Boceprevir, GC-3736, and Calpain Inhibitors II, XII Inhibit SARS-CoV-2 Viral Replication by Targeting the Viral Main Protease. Cell Res. 2020, No. May, 1–15.
(14) 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.; et al. Feline Coronavirus Drug Inhibits the Main Protease of SARS-CoV-2 and Blocks Virus Replication. Nat. Commun. 2020, 11 (1), 4282.
(15) Fu, L.; Ye, F.; Feng, Y.; Yu, F.; Wang, Q.; Wu, Y.; Zhao, C.; Sun, H.; Huang, B.; Niu, P.; et al. Both Boceprevir and GC-376 Efficiently Inhibit SARS-CoV-2 by Targeting Its Main Protease. Nat. Commun. 2020, 11 (1), 1–8.
(16) Qiao, J.; Li, Y.-S.; Zeng, R.; Liu, F.-L.; Luo, R.; Huang, C.; Wang, Y.-F.; Zhang, J.; Quan, B.; Shen, C.; et al. SARS-CoV-2 M pro Inhibitors with Antiviral Activity in a Transgenic Mice Model. Science (80). 2021, 371 (6536), 1374–1378.
(17) Hoffman, R. L.; Kania, R. S.; Brothers, M. A.; Davies, J. F.; Ferre, R. A.; Gajiwala, K. S.; He, M.; Hogan, R. J.; Koizminski, K.; Li, L. Y.; et al. Discovery of Ketone-Based Covalent Inhibitors of Coronavirus 3CL Proteases for the Potential Therapeutic Treatment of COVID-19. J. Med. Chem. 2020, 63 (31), 12725–12747.
Clinical Candidate for the Potential Treatment of COVID-19. bioRxiv Prepr. Serv. Biol. 2020.

Bethany, H. Pfizer Unveils Its Oral SARS-CoV-2 Inhibitor. Chemical and Engineering News April 7, 2021.

Vandyck, K.; Deval, J. Considerations for the Discovery and Development of 3-Chymotrypsin-like Cysteine Protease Inhibitors Targeting SARS-CoV-2 Infection. Curr. Opin. Virol. 2021, 49, 36–40.

Pitzer Initiates Phase 1 Study of Novel Oral Antiviral Therapeutic Agent Against SARS-CoV-2 https://www.pfizer.com/news/press-release/press-release-detail/pitzer-initiates-phase-1-study-novel-oral-antiviral (accessed Apr 5, 2021).

Potashman, M. H.; Duggan, M. E. Covalent Modifiers: An Orthogonal Approach to Drug Design. J. Med. Chem. 2009, 52 (5), 1231–1246.

Singh, J.; Petter, R. C.; Baillie, T. A.; Whitty, A. The Resurgence of Covalent Drugs. Nat. Rev. Drug Discov. 2011, 10 (4), 307–317.

Bauer, R. A. Covalent Inhibitors in Drug Discovery: From Accidental Discoveries to Avoided Liabilities and Designed Therapies. Drug Discov. Today 2015, 20 (9), 1061–1073.

Baillie, T. A. Targeted Covalent Inhibitors for Drug Design. Angew. Chemie - Int. Ed. 2016, 55 (43), 13408–13421.

Awoonor-Williams, E.; Rowley, C. N. How Reactive Are Druggable Cysteines in Protein Kinases? J. Chem. Inf. Model. 2018, 58 (9), 1935–1946.

Vita, E. De 10 Years into the Resurgence of Covalent Drugs. Future Med. Chem. 2021, 13 (2), 193–210.

Sutanto, F.; Konstantinidou, M.; Dömling, A. Covalent Inhibitors: A Rational Approach to Drug Discovery. RSC Med. Chem. 2020, 11 (8), 876–884.

Serafinova, I. M.; Pufall, M. A.; Krishnan, S.; Duda, K.; Cohen, M. S.; Maglathlin, R. L.; McFarland, J. M.; Miller, R. M.; Frodin, M.; Tanton, J. Reversible Targeting of Noncatalytic Cysteines with Chemically Tuned Electrophiles. Nat. Chem. Biol. 2012, 8 (5), 471–476.

Cox, R. M.; Plemer, R. K. The Impact of High-Resolution Structural Data on Stemming the COVID-19 Pandemic. Curr. Opin. Virol. 2021, 59, 127–138.

Muratore, E. N.; Amaro, R.; Andrade, C. H.; Brown, N.; Ekins, S.; Fourches, D.; Isayev, O.; Kosakov, D.; Medina-Franco, J. L.; Merz, K. M.; et al. A Critical Overview of Computational Approaches Employed for COVID-19 Drug Discovery. Chem. Soc. Rev. 2021.

Awoonor-Williams, E.; Abu-Saleh, A. A.-A. A. Covalent and Non-Covalent Binding Free Energy Calculations for Peptidomimetic Inhibitors of SARS-CoV-2 Main Protease. Phys. Chem. Chem. Phys. 2021.

Aldeghi, M.; Heifetz, A.; Bodkin, M. J.; Knapp, S.; Biggin, P. C. Accurate Calculation of the Absolute Free Energy of Binding for Drug Molecules. Chem. Sci. 2016, 7 (1), 207–218.

Fukunishi, H.; Watanabe, O.; Takada, S. On the Hamiltonian Replica Exchange Method for Efficient Sampling of Biomolecular Systems: Application to Protein Structure Prediction. J. Phys. Chem. 2002, 116 (20), 9058.

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 (1), 235–242.

Aldeghi, M.; Heifetz, A.; Bodkin, M. J.; Knapp, S.; Biggin, P. C. Accurate Calculation of the Absolute Free Energy of Binding for Drug Molecules. J. Phys. Chem. Lett. 2012, 3 (9), 4662–4673.

Huang, C.; Wei, P.; Fan, K.; Liu, Y.; Lai, L. 3C-like Proteinase from SARS Coronavirus Catalyzes Substrate Hydrolysis by a General Base Mechanism. Biochemistry 2004, 43 (15), 4568–4574.

Świderek, K.; Moliner, V. Revealing the Molecular Mechanisms of Proteolysis of SARS-CoV-2 M Pro by QM/MM Computational Methods. Chem. Sci. 2020, 11 (39), 10626–10630.

Ramos-Guzmán, C. A.; Ruiz-Pernía, J. J.; Tufón, I. Unraveling the Calculation of the PIC of Cysteine Residues in Proteins. J. Chem. Theory Comput. 2016, 12 (9), 4662–4673.

Huang, C.; Wei, P.; Fan, K.; Liu, Y.; Lai, L. 3C-like Proteinase from SARS Coronavirus Catalyzes Substrate Hydrolysis by a General Base Mechanism. Biochemistry 2004, 43 (15), 4568–4574.

Aldeghi, M.; Heifetz, A.; Bodkin, M. J.; Knapp, S.; Biggin, P. C. Accurate Calculation of the Absolute Free Energy of Binding for Drug Molecules. J. Phys. Chem. Lett. 2012, 3 (9), 4662–4673.

Huang, C.; Wei, P.; Fan, K.; Liu, Y.; Lai, L. 3C-like Proteinase from SARS Coronavirus Catalyzes Substrate Hydrolysis by a General Base Mechanism. Biochemistry 2004, 43 (15), 4568–4574.

Świderek, K.; Moliner, V. Revealing the Molecular Mechanisms of Proteolysis of SARS-CoV-2 M Pro by QM/MM Computational Methods. Chem. Sci. 2020, 11 (39), 10626–10630.

Ramos-Guzmán, C. A.; Ruiz-Pernía, J. J.; Tufón, I. Unraveling the Calculation of the PIC of Cysteine Residues in Proteins. J. Chem. Theory Comput. 2016, 12 (9), 4662–4673.

Huang, C.; Wei, P.; Fan, K.; Liu, Y.; Lai, L. 3C-like Proteinase from SARS Coronavirus Catalyzes Substrate Hydrolysis by a General Base Mechanism. Biochemistry 2004, 43 (15), 4568–4574.

Świderek, K.; Moliner, V. Revealing the Molecular Mechanisms of Proteolysis of SARS-CoV-2 M Pro by QM/MM Computational Methods. Chem. Sci. 2020, 11 (39), 10626–10630.

Ramos-Guzmán, C. A.; Ruiz-Pernía, J. J.; Tufón, I. Unraveling the Calculation of the PIC of Cysteine Residues in Proteins. J. Chem. Theory Comput. 2016, 12 (9), 4662–4673.

Huang, C.; Wei, P.; Fan, K.; Liu, Y.; Lai, L. 3C-like Proteinase from SARS Coronavirus Catalyzes Substrate Hydrolysis by a General Base Mechanism. Biochemistry 2004, 43 (15), 4568–4574.

Świderek, K.; Moliner, V. Revealing the Molecular Mechanisms of Proteolysis of SARS-CoV-2 M Pro by QM/MM Computational Methods. Chem. Sci. 2020, 11 (39), 10626–10630.

Ramos-Guzmán, C. A.; Ruiz-Pernía, J. J.; Tufón, I. Unraveling the Calculation of the PIC of Cysteine Residues in Proteins. J. Chem. Theory Comput. 2016, 12 (9), 4662–4673.
Awoonor-Williams, E.; Rowley, C. N. The Hydration Structure of Methylthiolate from QM/MM Molecular Dynamics. *J. Chem. Phys.* 2018, 149(4), 045103.

Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* 2004, 25 (9), 1157–1174.

Ferreira, J. C.; Fadl, S.; Villanueva, A. J.; Rabeh, W. M. Catalytic Dyad Residues His41 and Cys145 Impact the Catalytic Activity and Overall Conformational Fold of the Main SARS-CoV-2 Protease 3-Chymotrypsin-Like Protease. *Front. Chem.* 2021, 9 (June), 1–11.

Paasche, A.; Zipper, A.; Schäfer, S.; Ziebuhr, J.; Schirmeister, T.; Ramos-Murgía, C. A.; Ruiz-Tuñón, I. Multiscale Simulations of SARS-CoV-2 3CL Propeptide Inhibition with Aldehyde Derivatives. Role of Protein and Inhibitor Conformational Changes in the Reaction Mechanism. *ACS Catal.* 2021, 11 (7), 4157–4168.

Kneller, D. W.; Phillips, G.; O’Neill, H. M.; Jedrzejczak, R.; Stols, L.; Langan, P.; Joachimiak, A.; Coates, L.; Kovalevsky, A. Structural Plasticity of SARS-CoV-2 3CL Mpro Active Site Cavity Revealed by Room Temperature X-Ray Crystallography. *Nat. Commun.* 2020, 11 (1), 7–12.

González, F. V.; Arafet, K.; Serrano-Bruton, J. M. A Perspective on the Kinetics of Covalent and Irreversible Inhibition. *SLAS Discov. Adv. Sci. Drug Discov.* 2017, 22 (1), 3–20.

Kuzmic, P. A Two-Point IC50 Method for Evaluating the Biochemical Potency of Irreversible Enzyme Inhibitors. *bioRxiv* 2020, No. July.

Awoonor-Williams, E.; Kennedy, J.; Rowley, C. N. Measuring and Predicting Warhead and Residue Reactivity. In *Annual Reports in Medicinal Chemistry*, 2021; pp 203–227.

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* (80-.)) 2020, 368 (6497), 1331–1335.

Awoonor-Williams, E.; Walsh, A. G.; Rowley, C. N. Modeling Covalent-Modifier Drugs. *Biochim. Biophys. Acta - Proteins Proteomics* 2017, 1865 (11), 1664–1675.

Awoonor-Williams, E.; Rowley, C. Calculating the Full Free Energy Profile for Covalent Modification of a Druggable Cysteine in Bruton’s Tyrosine Kinase. *ChemRat* 2020.

Awoonor-Williams, E. Modelling Covalent Modification of Cysteine Residues in Proteins, Memorial University, 2020.

Mihalovits, L. M.; Ferency, G. G.; Keserü, G. M. Affinity and Selectivity Assessment of Covalent Inhibitors by Free Energy Calculations. *J. Chem. Inf. Model.* 2020, 60 (12), 6579–6594.

Mihalovits, L. M.; Ferency, G. G.; Keserü, G. M. The Role of Quantum Chemistry in Covalent Inhibitor Design. *Int. J. Quantum Chem.* 2021, No. June, 1–17.

Arafet, K.; Serrano-Aparicio, N.; Lodola, A.; Mulholland, A. J.; González, F. V.; Świderski, K.; Moliner, V. Mechanism of Inhibition of SARS-CoV-2 M pro by N3 Peptidyl Michael Acceptor Explained by QM/MM Simulations and Design of New Derivatives with Tunable Chemical Reactivity. *Chem. Sci.* 2021.

Ramos-Guzmán, Carlos A.; Ruiz-Pernia, J. Javier; Tunon, I. Inhibition Mechanism of SARS-CoV-2 Main Protease with Ketone-Based Inhibitors Unveiled by Multiscale Simulations. *ChemRxiv* 2021, 1–22.

Zhang, H.; Jiang, W.; Chatterjee, P.; Luo, Y. Ranking Reversible Covalent Drugs: From Free Energy Perturbation to Fragment Docking. *J. Chem. Inf. Model.* 2019, 59 (5), 2093–2102.