JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

Open Access

pubs.acs.org/jcim

Article

Computational Study of the Inhibition of RgpB Gingipain, a Promising Target for the Treatment of Alzheimer's Disease

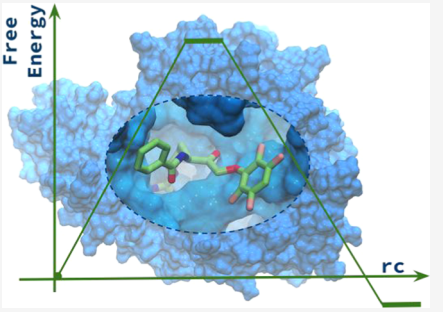
Santiago Movilla, Sergio Martí, Maite Roca,* and Vicent Moliner*

Cite This: *J. Chem. Inf. Model.* 2023, 63, 950−958

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Alzheimer's disease represents one of the most ambitious challenges for biomedical sciences due to the growing number of cases worldwide in the elderly population and the lack of efficient treatments. One of the recent attempts to develop a treatment points to the cysteine protease RgpB as a promising drug target. In this attempt, several small-molecule covalent inhibitors of this enzyme have been proposed. Here, we report a computational study at the atomic level of the inhibition mechanism of the most promising reported compounds. Molecular dynamics simulations were performed on six of them, and their binding energies in the active site of the protein were computed. Contact maps and interaction energies were decomposed by residues to disclose those key interactions with the enzyme. Finally, quantum mechanics/molecular mechanics (QM/MM) molecular dynamics (MD) simulations were performed to evaluate the reaction mechanism by which these drug candidates lead to covalently bound complexes, inhibiting the RgpB protease. The results provide a guide for future re-design of prospective and efficient inhibitors for the treatment of Alzheimer's disease.



1. INTRODUCTION

Gingipains are a group of enzymes secreted by the pathogenic bacterium *Porphyromonas gingivalis*.[1] Although their activity is mainly related to oropharyngeal problems, there are reports that also associate them with other health disorders.[2−5] Some of these disorders are of particular medical interest, such as heart conditions and Alzheimer's disease.[2,6−11] The increasing number of neurodegenerative diseases worldwide and the lack of an appropriate medical treatment explains why this is the greatest current challenge for biomedical sciences.

Gingipains are a small family of cysteine protease enzymes which catalyze the cleavage of peptidic bonds in several protein substrates.[12] In this family, we can find two types of gingipains, Lys-gingipains (Kgp), and Arg-gingipains (Rgp), classified according to the residue that they recognize at the P1 position, lysine and arginine, respectively, to perform the cleavage.[12] Without extra marked preferences for other amino acids at positions beyond P1, gingipains catalyze the proteolysis of a large number of proteins and peptides.[13,14] This makes them potentially harmful to the host integrity. Indeed, a recent study strongly linked one of the Rgp, RgpB, to the progression of Alzheimer's disease.[6] This study also showed that RgpB inhibition resulted in a strong arrest of disease progression in mice. With this study, RgpB was positioned as an important pharmacological target to understand its function and selectively inhibit it.

The proteolysis reaction mechanism of RgpB was recently studied by computational methods and reported in detail by our group.[15] In this study, we highlight several critical features to understand the catalytic mechanism of this cysteine protease and proceed to its inhibition. In general, and as is common in cysteine proteases, the mechanism is divided into an acylation step and a deacylation step.[16−26] According to our results,[15] the deprotonation of the nucleophilic cysteine of RgpB in the acylation stage is performed by the substrate itself in response to the steric impossibility of the catalytic histidine to perform this process. Similar mechanisms for activation have been reported in enzymes whose active site distribution coincides with that of RgpB,[18] where the substrate is positioned between the Cys/His catalytic dyad. After the deprotonation of the catalytic Cys, the sulfur atom of Cys attacks the carbonyl group of the substrate to obtain the acylenzyme. The deacylation stage is carried out in a single step where a water molecule attacks the carbonyl group of the substrate and one proton of the water is transferred to the sulfur atom of the Cys residue assisted by the catalytic His.

The interest in the inhibition of gingipains has been growing since their potential implications in the treatment of neurodegenerative diseases have been demonstrated.[27] In fact, some potential drug candidates, which have been patented (International Patent Application PCT/US2015/054050 and PCT/US2016/061197), are in advanced stages of medical trials. Interestingly, these compounds were reported to be irreversible

Received: September 26, 2022
Published: January 17, 2023

https://doi.org/10.1021/acs.jcim.2c01198
J. Chem. Inf. Model. 2023, 63, 950−958

inhibitors after the corresponding kinetic tests had been carried out.[6] Given the pharmacological potential of these molecules, the main goal of the present study is to investigate how these irreversible inhibitors can stop the catalytic activity of RgpB as a protease.

Structurally, these inhibitors have the side chain pattern of arginine but without the nitrogen of the P1−P1′ peptide bond. The absence of this nitrogen makes it impossible to carry out the hydrolysis reaction, breaking the peptide bond, thus inhibiting the enzyme. Interestingly, they also lack the reactive warheads commonly used for the covalent inhibition of cysteine proteases.[20,28−30] This fact limits the reactive possibilities of these compounds against the action of RgpB gingipain. In our previously reported study,[15] it was proposed that, in the absence of a peptidic nitrogen, the activation of catalytic cysteine was carried out by the proton transfer from the sulfur atom of Cys to the oxygen atom of the carbonyl group of the substrate.[15] Based on these results, we proposed that this mechanism could be exploited for the design of potential irreversible drugs.

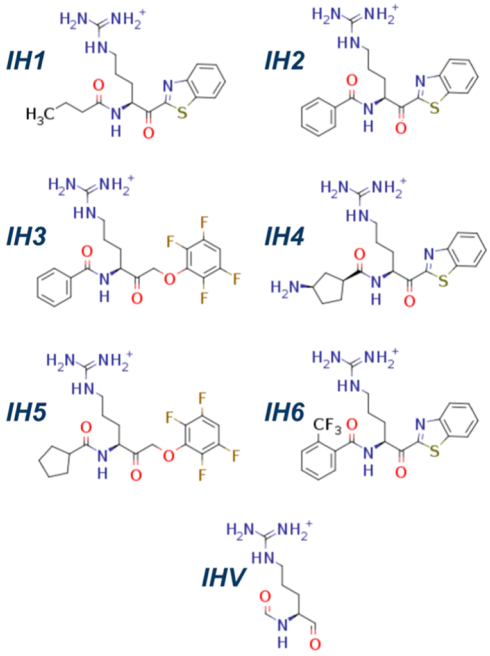
This background and in view of the limited reactive nature of the already irreversible patented inhibitors, a detailed understanding of their inhibition mechanism is necessary to exploit their potential in the future. Namely, in the case of covalent inhibitors, the inhibitory potency derives from the synergy between the noncovalent interactions with the enzyme and the kinetics/thermodynamics of the reactive process.[29] Therefore, an atomic-level study of the inhibition mechanism must incorporate both the study of the inhibitor−enzyme interactions and the study of the covalent binding reaction mechanism. Here, we present a comprehensive computational study that provides insights into important atomic details in the reaction mechanism of inhibition, binding affinity, and the interaction between the enzyme and a representative group of inhibitors. Classical molecular dynamics (MD) simulations, alchemical transformations,[31] molecular mechanics Poisson−Boltzmann surface area (MM/PBSA)[32] calculations supplemented with interaction entropies, and MD simulations with quantum mechanics molecular mechanics (QM/MM) potentials were carried out to get a detailed description of the inhibitor−enzyme binding step and the chemical steps of the covalent inhibitor−enzyme bond formation. The obtained results throw light on the design of better inhibitors for Alzheimer's disease treatment.

2. METHODS

2.1. System Set Up. The initial coordinates of the system were obtained from the previously equilibrated model[15] of RgpB in complex with a short peptide. This model was reported in the reactivity study of RgpB performed by our group. The peptide was replaced by six active inhibitors, (Scheme 1) reported in International Patent Application PCT/US2016/061197, which covered the largest extent of the chemical space. The protonation state of all titratable residues was verified using the PROPKA3 server[33] at reference pH 7.5,[6,13,14] no atypical protonation states were observed. Each of the systems was solvated with a cubic box of TIP3P water[34] molecules with a minimum distance of 15 Å between any solute atom and the edge of the box. A total number of 18 Na^+ ions were added until the systems were neutralized.

2.2. Classical Molecular Dynamics (MD). Protein and water molecules were described using the AMBERff14SB[35] and TIP3P[34] force field parameters, respectively. Monovalent ions were treated using the parameters proposed by Joung−Cheatham.[36] For inhibitors, restrained electrostatic potential

Scheme 1. Chemical Structures of the Studied Irreversible Inhibitors of RgpB (IH1−IH6) and Sketch of the Virtual Inhibitor (IHV) Used as the Common Point for the Alchemical Transformations



(RESP) charges[37] were obtained from HF/6-31G* calculations on optimized geometries with the B3LYP functional[38,39] with the 6-31G* basis set, according to previously reported protocols.[40] The remaining force field parameters for the inhibitors were assigned from the small-molecules generalized Amber force field (GAFF).[40] The topologies of each system were obtained using the tLEaP package from Ambertools20.[41]

All MD classical simulations were carried out in several steps: (i) solvent molecules, ions, and hydrogens were minimized using 2500 steps with the conjugate gradient algorithm; (ii) 200 ps of MD simulations of the solvent molecules and monovalent ions were carried out, with the positions of the backbone atoms restrained; (iii) two energy minimizations, one with the protein backbone restrained and another fully unrestrained were done for each system; (iv) the whole systems were heated gradually in three consecutive NPT simulations from 100 to 310 K, with a constant pressure of 1 bar; and finally, (v) 250 ns of NPT MD simulations were performed at 310 K and 1 bar. All parameters of the classical simulations were replicated from the previously reported study.[15] A 10 Å cutoff was set up for short-range nonbonded interactions while a Particle Mesh Ewald (PME)[42,43] model was used for the long-range interactions. Langevin dynamics thermostat[44,45] was applied with a 3 ps^{-1} collision frequency. For all equilibration simulations, the SHAKE algorithm[46,47] was used to constrain light atoms, and the velocity Verlet algorithm[48] was used to propagate the system.

2.3. MM/PBSA and Interaction Entropies. To estimate the binding free energy (ΔG_{bind}) of the inhibitors, calculations of MM/PBSA[32] supplemented with corrections to the solute entropy using the methodology proposed by Duan et. al.[49] were performed. For these calculations, frames were taken every 500 ps from each classical MD simulation. MM/PBSA calculations were performed with an implicit salt concentration of 150 mM (to reproduce experimental conditions)[6,13] and using

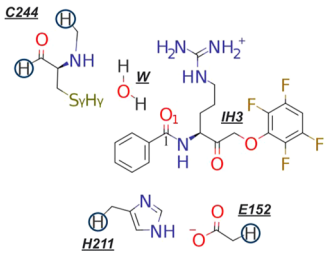
https://doi.org/10.1021/acs.jcim.2c01198
J. Chem. Inf. Model. 2023, 63, 950−958

MMPBSA.py as implemented in Amber20.[41] To obtain the binding free energy (ΔG_{bind}), we applied the interaction entropy approach[49] to compute the entropic contribution term ($-T\Delta S$) at 310 K.

2.4. Alchemical Calculations. To obtain accurate relative values of the binding free energy ($\Delta\Delta G_{bind}$), we performed alchemical transformations[31] from the inhibitors to a virtual ligand (IHV) connecting them all (Scheme 1). For this purpose, the 3-step Amber Thermodynamic Integration protocol ("decharge−LJ−recharge" protocol)[50] was used. Each transformation was carried out in ten windows equally distributed throughout the λ range (0−1). In each window, 500 ps of CPU equilibration (pmemd.MPI) and 5 ns of sampling with GPU algorithms (pmemd.cuda) were performed.[51,52] Only the atoms appearing/disappearing during the transformation were included in the soft-core region. All the simulations were performed at the NVT ensemble at 310 K, starting from the volume-equilibrated structures of the classical MD trajectories. For data analysis, the first 5% of the simulation time (250 ps) of each window was discarded.

2.5. Potential of the Mean Force (PMF). The six studied inhibitors present the same reactive carbonyl warhead (C1:IH3−O1:IH3, Scheme 2). Then, the inhibitor that showed the most favorable binding energy was selected for studying the reaction mechanism of the formation of the inhibitor−enzyme covalent complex by generating the free energy surfaces (FES) in terms of potential of mean force (PMFs). For the PMF calculations, two-dimensional (2D) potential energy surfaces (PES) were first computed through sequential minimizations along selected collective variables (CV) that best describe the reactions. A conjugate gradient algorithm was employed for the minimizations using a gradient tolerance of 0.1 kcal mol^{-1} as a convergence criterion. Later, the corresponding FES were generated at 310 K taking the structures obtained along the PES as starting points. Each window had a relaxation time of 5 ps and a sampling time of 25 ps using a time step of 0.5 fs in the NVT ensemble. Temperature control was performed using Langevin dynamics with a 3 ps^{-1} collision frequency.[44,45] The umbrella sampling method[53] was used to restrain the reaction coordinates. The force constant used for each window was 580 kcal mol^{-1} Å^{-2}, and the window width was 0.05 Å for those collective variables corresponding to the antisymmetric combination of two distances and 0.1 Å for those distances. The number and the width of the windows selected ensure a correct overlapping of windows. The umbrella integration method, implemented in the QM3 suite,[54] was used to reweight the biased sampling dynamics and to generate the PMFs along selected coordinates. The quantum mechanical (QM) region used for the generation of the PMFs is shown in Scheme 2. The method used to describe the QM region was PM6.[55] Following the same protocol of our previous study on this system,[15] a structure close to the transition state region was selected and the transition state structure was located at the PM6/MM level and verified by tracing down the intrinsic reaction coordinate path (IRC). The obtained minima were optimized. Single point energy calculations were carried out on the stationary point structures at the PBE-D3(BJ)[56−60]/MM level with the 6-311+G** basis set, to correct the PM6/MM potential energy. The thermal contributions calculated by the statistical methods at the PM6/MM level were thus preserved. To verify the results, all the critical points were reoptimized at the PBE-D3(BJ)/MM level using the QM3 suite.[54]

Scheme 2. Schematic Representation of the Region Treated Quantum Mechanically to Explore the Inhibitory Mechanism of Binding between IH3 and RgpB[a]



[a]The link atoms are shown as hydrogens inside circles. The water molecule (W) was included in the QM region when exploring the mechanism with the participation of a water molecule.

3. RESULTS AND DISCUSSION

3.1. Binding Affinities (ΔG_{bind}) and Inhibitor−Enzyme Interaction Profiles. To study the noncovalent interaction profiles between inhibitors and RgpB enzyme, 250 ns of classical MD simulations were run per inhibitor in the inhibitor−enzyme noncovalent reactant complex. None of the systems, in the presence of the corresponding inhibitor, exhibited significant changes in the protein structure (all backbone RMSDs <2 Å on average, Figure S1).

Starting from the equilibrated structures of the classical MD simulations, we initially proceeded to estimate the binding free energies for each of the compounds. For this purpose, we computed the binding affinity energies by MM/PBSA[32] calculations supplemented with a correction term to the entropic contributions based on the interaction entropies proposed by Duan et al.[49] (Table 1). Qualitatively speaking, compounds IH3 and IH4 proved to produce the most stable inhibitor−enzyme complexes. It should also be noted that IH3 is

Table 1. Binding Free Energies for the Studied Inhibitors Obtained over the Classical MD Simulations Using MM/PBSA and Interaction Entropies ($\Delta G_{bind\text{-}MM/PBSA}$) and Binding Free Energies Relative to the Virtual Inhibitor (IHV) Computed by Means of Alchemical Transformations ($\Delta G_{bind\text{-}TI}$)[a]

Inhibitor	$\Delta H_{bind\text{-}MM/PBSA}$	$-T\Delta S_{bind\text{-}MM/PBSA}$	$\Delta G_{bind\text{-}MM/PBSA}$	$\Delta\Delta G_{bind\text{-}TI}$
IH1	−57.3	37.0	−20.3	0.20
IH2	−54.2	50.6	−3.6	5.42
IH3	−69.4	27.3	−42.1	0.55
IH4	−67.2	28.2	−39.0	1.99
IH5	−63.3	27.8	−35.5	2.55
IH6	−60.9	33.1	−27.8	3.31

[a]All values are in kcal mol^{-1}.

https://doi.org/10.1021/acs.jcim.2c01198
J. Chem. Inf. Model. 2023, 63, 950−958

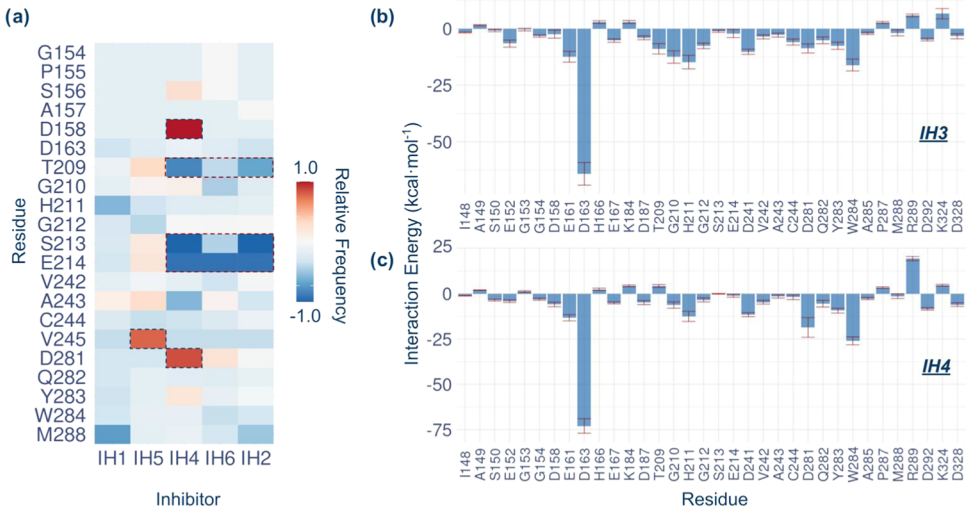


Figure 1. (a) Relative frequency of contacts between residues of RgpB and the inhibitors. IH3 was used as a reference and a contact was counted if the distance between atoms was <4.0 Å. The cells corresponding to the interactions that differ the most from the IH3 inhibitor are highlighted with dashed-line borders. Panels (b, c) show averaged interaction energies (electrostatic plus Lennard−Jones) between residues of RgpB and IH3 and IH4, respectively, over 1 ns MD simulations at the PM3/MM level.

the one that shows the most favorable binding energy in both entropic and enthalpic terms. On the other hand, inhibitor IH2 presented a lower binding enthalpy and a higher $-T\Delta S$ term. We emphasize that the values computed by means of MM/PBSA and interaction entropies have no quantitative meaning and are purely used to qualitatively analyze the enthalpic and entropic contributions of the binding processes.[61] Thus, keeping in mind the inherent uncertainty associated with the MM/PBSA method, alchemical transformations[31] were employed to compute the differences between the binding free energies of every inhibitor and the virtual inhibitor IHV ($\Delta\Delta G_{\text{bind-TI}}$) to obtain more precise and quantitatively meaningful inhibitor−enzyme affinities.

The results of the alchemical transformations show a similar trend to the one obtained from MM/PBSA calculations, with the only exception of IH1, previously ranked as the second weakest, is repositioned in first place with practically the same $\Delta\Delta G_{\text{bind-TI}}$ as IH3. The same as the MM/PBSA calculations,[32] IH4 ranks after IH3 as one of the most potent candidates. Likewise, IH2 is the inhibitor with the lowest affinity to the enzyme with a difference of 5.2 kcal mol^{-1} with respect to IH1. Henceforth, IH3 can be considered as the reference inhibitor given the agreement between the binding free energies estimated by both methods.

A contact frequency map allowed us to analyze the differences in the interaction patterns/profiles of each inhibitor, which can be complemented with the analysis of the averaged inhibitor−enzyme interaction energies decomposed by the residue (Figure 1). Figure 1a shows the relative contact frequencies with respect to the inhibitor with the highest affinity, IH3. Thus, positive values represent a higher contact frequency than in IH3 while negative values represent a lower contact frequency than in IH3. In general, most of the inhibitors present a similar interaction profile with the enzyme, which is in agreement with the differences observed between the binding free energies ($\Delta\Delta G_{\text{bind-TI}}$) computed by alchemical transformations. However, it can be observed from the contact map that three inhibitors (IH2, IH4, and IH6) show a significantly lower contact frequency with residues Ser213 and Glu214 (see Figure 1a). A visual inspection revealed that for the case of inhibitors IH3 and IH5 this interaction corresponds to a hydrogen bond between the HN:Glu214 and the F1 atom of the inhibitor. In the case of the interaction with the Ser213 residue does not form a hydrogen bond due to the angle of the atoms involved (angle around 59 ± 16°) (Figure 2). Thus, for those inhibitors that

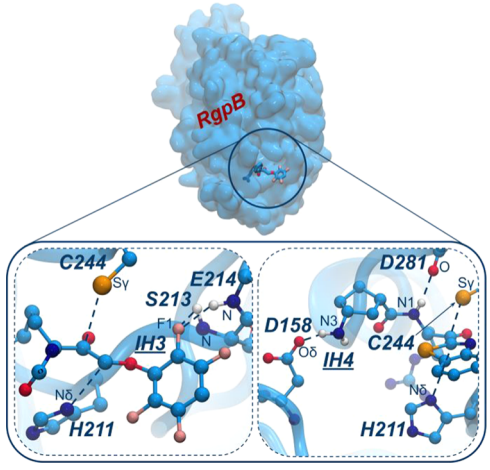


Figure 2. Top panel: surface representation of the whole protein and location of the binding pocket. IH3 inhibitor is represented as sticks. Bottom panels: Schematic representation of the key interactions between IH3 (left) or IH4 (right) and the binding pocket of RgpB.

instead of the tetrafluorophenoxy substituent have the benzothiazole ring, the interactions with Ser213 and Glu214 are not present. However, IH1 manages to interact with HN:Glu214 via the nitrogen atom of the bicyclo ring. Although no preferences in selectivity of the groups around the arginine moiety have been reported in RgpB that affect the catalytic capacity, it has shown a higher affinity for hydrophobic substituents, mainly aromatic.[13] However, these results suggest that hydrogen bonding groups are necessary to stabilize these aromatic substituents within the active site and can be optimized to achieve better affinity to the enzyme.

Another interaction worth highlighting is the one observed between compound IH4 and residues Asp158 and Asp281 (see

https://doi.org/10.1021/acs.jcim.2c01198
J. Chem. Inf. Model. 2023, 63, 950−958

Figures 1a and 2). Namely, IH4 is the only one of the compounds studied that presents a hydrogen bond donor at the nitrogen substituent group of the arginine. Consequently, HN3:IH4 manages to interact with Oδ:Asp158 (Figure 2). As a result, this inhibitor is repositioned in such a way that it manages to form an extra hydrogen bond interaction between the HN1:IH4 and O:Asp281 atoms. Although this interaction would be possible in all inhibitors, it was only observed in IH4, suggesting that it responds to the conformation adopted due to the interaction with the Asp158 residue. Other differences in the contact map, such as those observed in residues Thr209 or Val245, are less specific and are the consequence of a particular physical proximity during the simulations.

Finally, to energetically characterize the favorable observed contacts between the inhibitors and the RgpB residues, the interaction energies were calculated (Figures 1b,c and S3−S6). We can observe the determinative role played by residues Asp163, Trp284, and His211 in arginine binding, a conclusion that was predicted from structural analysis and our previous report on the proteolysis reaction catalyzed by RgpB.[15] This result is not surprising given the high selectivity of RgpB to arginine residues. On the other hand, it can be observed that the interaction with residues Ser213 and Glu214 is always favorable. In the same way, the interaction between IH4 and residues Asp158 and Asp281 shows a lower energy, being remarkable for the one with Asp281. In contrast, residues Lys184, Arg289, and Lys324 show an unfavorable interaction energy with all compounds. Although this may be a starting point for future optimizations, it derives from the spatial proximity (without contact) between the charges of these residues and the positive charge of the guanidinium group of the arginine moiety of the inhibitors.

To sum up, our results suggest that the optimization should be focused on compounds presenting hydrogen bond acceptor groups capable of interacting with HN:Glu214 and HN:Ser213. Moreover, hydrogen bond donor groups interacting with Asp158 and Asp281 seem to enhance the affinity with the enzyme. On the other hand, the guanidinium group present in all inhibitors and in the natural substrate should be preserved as has been shown to be essential for recognition by the RgpB binding pocket.

3.2. Covalent Binding Chemical Step. To study the reaction mechanism by which these inhibitors covalently bind to the enzyme, QM/MM MD simulations were performed to generate the full free energy landscapes of the most plausible mechanisms. Given the similarity between the inhibitors, the reactivity study was carried out using only the IH3 inhibitor, which showed the most favorable binding energy.

Initially, and analogously to the proteolysis reaction catalyzed by the enzyme in the presence of wild-type substrate, Cys244 is protonated and requires activation to attack the carbonyl group in the reactant state (R in Figures 3a,b and 4b). We also considered the possibility that the mechanism proceeded from the deprotonated Sγ:Cys244 form (mechanism $R_{S(-)} \rightarrow P_{S(-)}$ in Figure 3c). However, several attempts (with and without restraints) to obtain stable reactive structures were unsuccessful. In all simulated cases, the negative charge on Cys244 resulted in deformations of the system to chemically unviable structures.

The family of inhibitors reported does not possess a clear reactive moiety, such as a Michael acceptor or an epoxide group. Instead, they only possess the carbonyl oxygen atom O1:IH3 as a possible acceptor/activator for the cysteine residue. This possibility was previously hypothesized[15] in light of the small

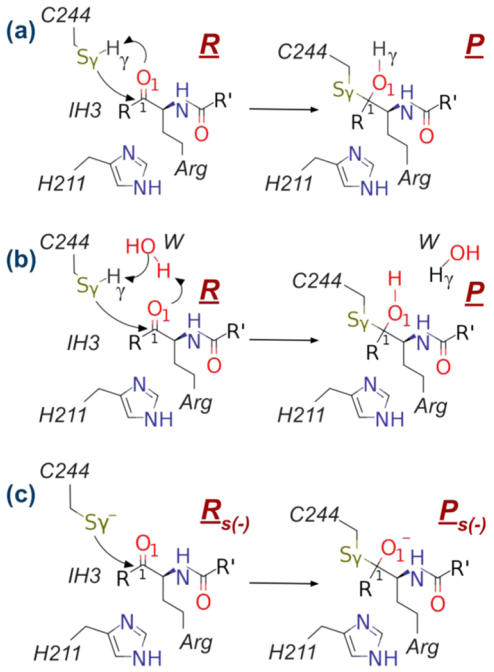


Figure 3. Schematic representation of the considered mechanisms of covalent binding between IH3 and RgpB. Sketch of the reaction mechanism in which transfer of the Hγ:Cys244 to the O1 atom of the inhibitor occurs (a) directly or (b) mediated by a water molecule. (c) Representation of the reaction mechanism starting from the deprotonated Sγ:C244 form. As shown in Scheme 1 for IH3, R and R′ correspond to phenyl and 2,3,5,6-tetrafluorophenoxy substituents, respectively.

difference between the free energy barriers shown by the wild-type being activated by either the peptide nitrogen (more favorable mechanism by ∼1.1 kcal mol^{-1}) or carbonyl oxygen of the substrate. Other possible bases were evaluated by calculating the population of hydrogen bonds between Hγ:Cys244 and other possible proton acceptors. However, the interaction with O1:IH3 was shown to populate more than 50% of the time, while other possible bases, such as Asp281, interacted less than 1% with it, making them poor base candidates. With the carbonyl group as the only plausible activator of Cys244, the mechanisms are limited to whether the attack of Sγ:Cys244 on C1:IH3 and the transfer of Hγ:Cys244 from Sγ:Cys244 to O1:IH3 occur directly or mediated by a water molecule (Figure 3a,b).

As revealed by the computed FESs (Figure S7), both mechanisms, the one in which the Hγ:Cys244 is directly transferred to the O1:IH3 atom (mechanism a in Figure 3) and the one mediated by a water molecule (mechanism b in Figure 3), proceed through a single concerted step. In the case of the direct mechanism, that was explored using the Sγ:Cys244−C1:IH3 distance and the difference between distances [Sγ:Cys244−Hγ:Cys244] − [O1:IH3−Hγ:Cys244] as collective variables to describe the process, the activation free energy associated with the transition state (TS, see Figure 4a) is 22.8 kcal mol^{-1} with respect to the reactants. A value that is very close to that previously computed for the natural substrate, 23.5 kcal mol^{-1}.[15,15]However, in the case of the transition state mediated by a water molecule (TS_W), that was explored using the distance Sγ:Cys244−C1:IH3 and the difference between distances [Sγ:Cys244−Hγ:Cys244] − [O:W−Hγ:Cys244] as collective variables to describe the process, the corresponding activation

https://doi.org/10.1021/acs.jcim.2c01198
J. Chem. Inf. Model. 2023, 63, 950−958

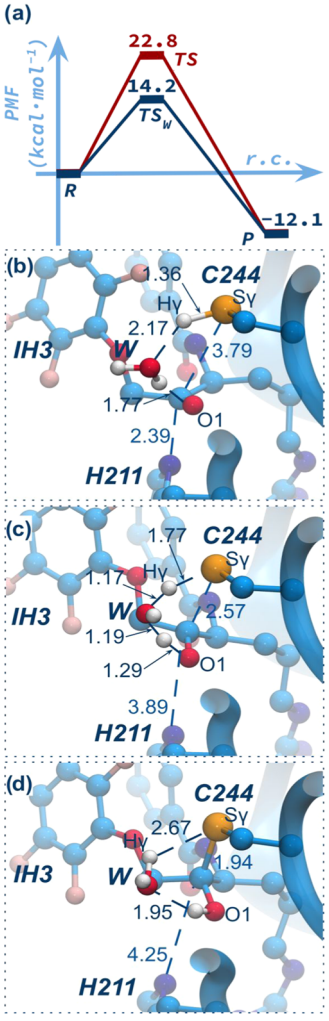


Figure 4. (a) Free energy profiles computed for the inhibitory covalent binding mechanisms of RgpB by IH3 at the PBE + D3(BJ):PM6/MM. The direct proton transfer mechanism from Cys244 to O1 atom of IH3 is depicted in red line while the proton transfer mechanism mediated by a water molecule is shown in blue line. PBE + D3(BJ)/MM optimized structures of the critical points (b) R, (c) TS_W, and (d) P along the most likely covalent binding mechanism of IH3 inhibiting RgpB catalytic activity. Selected distances are in Å. FESs are shown in Figure S7.

energy is 14.2 kcal mol^{-1}. This result suggests that this mechanism is the most viable for the covalent binding reaction with a significantly lower free energy barrier compared to that of the wild-type substrate. The obtained product (P) is located at −12.1 kcal mol^{-1} with respect to the reactants, thus giving an irreversible exergonic reaction. To verify the mechanism obtained at the PM6/MM level, R, TS_W, and P were fully optimized at a higher level of theory, PBE + D3(BJ)/MM. The high-level optimized structures, presented in Figure 4, match the reaction pathway predicted at a lower level, verifying the mechanism deduced from PM6/MM FES.

As mentioned above, the mechanism starting from the deprotonated Cys244 (mechanism c in Figure 3) was not explored because the active site was deformed along the classical MD simulations and no appropriately reactive structures were obtained.

In light of these results, it is remarkable to highlight the ability of RgpB to react without the presence of highly reactive groups. Common reactive groups such as Michael acceptors or epoxides deal with selectivity issues due to their reactivity with other undesired targets. The unusual carbonyl warhead enables to exploit the design of covalent inhibitors, with high potency, without having to face problems of selectivity. These features qualify RgpB as a pharmacological target that promises effective treatments by the use of these kinds of inhibitors without side effects for the treatment of Alzheimer's disease.

4. CONCLUSIONS

Alzheimer's disease is one of the most studied medical challenges today. Gingipains proteases, including RgpB, have positioned themselves as potential drug targets for the development of treatments for the disease. Herein, we unveil atomic-level details of the noncovalent binding processes for a set of RgpB gingipain inhibitors, potential candidates for the treatment of Alzheimer's disease. Six compounds, presenting the side chain pattern of arginine but without the nitrogen of the P1−P1′ peptide bond, were chosen to cover the widest breadth of the chemical space previously proposed and patented,[6] some of them are in advanced stages of clinical testing. Initially, noncovalent inhibitor−enzyme complexes were simulated by classical molecular dynamics. Over the MD sampling, binding free energies ($\Delta\Delta G_{bind\text{-}TI}$) were computed with alchemical transformation calculations.

Based on a relative contact map analysis between the inhibitors and the enzyme, we concluded that IH3, IH1, and IH5 present similar patterns of interactions, showing only minor differences between the contacts established with the protein residues along the sampled trajectories. Meanwhile, among IH2, IH4, and IH6, some of them rendered lower binding energies, and do not interact with key residues Ser213 and Glu214. However, IH4 exhibits two hydrogen bonding interactions that no other shows with Asp158 and Asp281 residues. These interactions as well as those with Ser213 and Glu214 were found to have a favorable interaction energy whenever present. The analyses of the noncovalent complexes suggest that IH1 and IH3 would be the most promising candidates for further refinements, according to their high binding affinities to the active site of RgpB.

The reaction mechanism of the covalent bond formation between the inhibitors and the enzyme was computed by QM/MM MD simulations on IH3, which showed the best binding profile in the noncovalent complex. The most plausible reaction mechanism proceeds through a single concerted step, in which activation/deprotonation of Cys244 is carried out by the O1:IH3 atom mediated by a water molecule (W). At the same time, the Sγ:Cys244 atom attacks the C1:IH3 of the inhibitor to give way to a stable product (−12.1 kcal mol^{-1}). The reaction proceeds through an activation barrier (14.2 kcal mol^{-1}) significantly lower than that reported previously, computed in our laboratory for the wild-type substrate (23.5 kcal mol^{-1}).[15] Thus, RgpB has a remarkable ability to react without the presence of highly reactive groups, enabling to exploit the redesign of covalent inhibitors, recognized for their potency, without predicting problems of selectivity. These results qualify RgpB as a pharmacological target that promises effective treatments by the use of these kinds of inhibitors without side effects. In particular, the inhibitors should contain a hydrogen bond donor and acceptor groups to be able to interact with Glu214, Ser213, Asp158, and Asp281, and conserving the reported guanidinium group and warhead to provide potency and selectivity. The interactions reported here and the mechanistic details represent a key starting point for future re-

https://doi.org/10.1021/acs.jcim.2c01198
J. Chem. Inf. Model. 2023, 63, 950−958

design of prospective and efficient inhibitors for the treatment of Alzheimer's disease.

■ ASSOCIATED CONTENT

Data Availability Statement

Amber14 Package can be purchased from ambermd.org/GetAmber.php. QM3 Suite is freely available via a public GitHub repository github.com/sergio-marti/qm3. AmberTools17 can be obtained from ambermd.org/AmberTools.php. Gaussian 09 D01 can be purchased from gaussian.com. fDYNAMO v2.2 can be freely downloaded from www.pdynamo.org/downloads. Force field parameters for IH3 and IHV are reported in the Supporting Information together with a complete IH3 inhibitor/protein/ions/solvent box PDB file. The coordinates of the QM region for the critical points located at the PBE + D3(BJ)/MM level are also reported in the Supporting Information. Examples of the inputs for classical MD, alchemical transformations and MMPBSA calculations are openly available on the AMBER website.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.2c01198.

Root-mean-square deviation (RMSD) analyses for classical MD simulations; root-mean-square fluctuation (RMSF) analyses for classical MD simulations; averaged interaction energies between the inhibitors and some protein residues for IH1, IH2, IH5, and IH6; free energy surfaces for the covalent binding inhibition mechanisms of the RgpB gingipain with IH3; force field parameters for IH3 and IHV; cartesian coordinates for R, TS_W, and P structures localized at the PBE + D3(BJ)/MM level for the covalent binding mechanism with the IH3 inhibitor (PDF)

A full structure of the solvated protein in complex with the IH3 inhibitor (PDB)

■ AUTHOR INFORMATION

Corresponding Authors

Maite Roca − *BioComp Group, Institute of Advanced Materials (INAM), Universitat Jaume I, 12071 Castellón, Spain;* orcid.org/0000-0003-0937-4722; Email: mroca@uji.es

Vicent Moliner − *BioComp Group, Institute of Advanced Materials (INAM), Universitat Jaume I, 12071 Castellón, Spain;* orcid.org/0000-0002-3665-3391; Email: moliner@uji.es

Authors

Santiago Movilla − *BioComp Group, Institute of Advanced Materials (INAM), Universitat Jaume I, 12071 Castellón, Spain;* orcid.org/0000-0002-6936-2806

Sergio Martí − *BioComp Group, Institute of Advanced Materials (INAM), Universitat Jaume I, 12071 Castellón, Spain;* orcid.org/0000-0002-1087-7143

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.2c01198

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We would like to thank the Spanish Ministerio de Ciencia e Innovación (grant PGC2021-23332OB-C21), the Generalitat Valenciana (PROMETEO, grant ref CIPROM/2021/079), and Universitat Jaume I (grants UJI-2020-03 and UJI-2019-43). S. Movilla thanks the Generalitat Valenciana for a Grisolia Ph.D. grant (GRISOLIAP/2019/064). Authors acknowledge computational resources from the Servei d'Informàtica of Universitat Jaume I.

■ REFERENCES

(1) Darveau, R. P.; Hajishengallis, G.; Curtis, M. A. *Porphyromonas gingivalis* as a Potential Community Activist for Disease. *J. Dent. Res.* **2012**, *91*, 816−820.

(2) Mahendra, J.; Mahendra, L.; Kurian, V. M.; Jaishankar, K.; Mythilli, R. Prevalence of Periodontal Pathogens in Coronary Atherosclerotic Plaque of Patients Undergoing Coronary Artery Bypass Graft Surgery. *J. Maxillofac. Oral Surg.* **2009**, *8*, 108−113.

(3) Katz, J.; Chegini, N.; Shiverick, K. T.; Lamont, R. J. Localization of *P. gingivalis* in Preterm Delivery Placenta. *J. Dent. Res.* **2009**, *88*, 575−578.

(4) Ishikawa, M.; Yoshida, K.; Okamura, H.; Ochiai, K.; Takamura, H.; Fujiwara, N.; Ozaki, K. Oral *Porphyromonas gingivalis* Translocates to the Liver and Regulates Hepatic Glycogen Synthesis through the Akt/GSK-3β Signaling Pathway. *Biochim. Biophys. Acta, Mol. Basis Dis.* **2013**, *1832*, 2035−2043.

(5) Mougeot, J.-L. C.; Stevens, C. B.; Paster, B. J.; Brennan, M. T.; Lockhart, P. B.; Mougeot, F. K. B. *Porphyromonas gingivalis* Is the Most Abundant Species Detected in Coronary and Femoral Arteries. *J. Oral Microbiol.* **2017**, *9*, No. 1281562.

(6) Dominy, S. S.; Lynch, C.; Ermini, F.; Benedyk, M.; Marczyk, A.; Konradi, A.; Nguyen, M.; Haditsch, U.; Raha, D.; Griffin, C.; Holsinger, L. J.; Arastu-Kapur, S.; Kaba, S.; Lee, A.; Ryder, M. I.; Potempa, B.; Mydel, P.; Hellvard, A.; Adamowicz, K.; Hasturk, H.; Walker, G. D.; Reynolds, E. C.; Faull, R. L. M.; Curtis, M. A.; Dragunow, M.; Potempa, J. *Porphyromonas gingivalis* in Alzheimer's Disease Brains: Evidence for Disease Causation and Treatment with Small-Molecule Inhibitors. *Sci. Adv.* **2019**, *5*, No. eaau3333.

(7) Kaye, E. K.; Valencia, A.; Baba, N.; Spiro, A.; Dietrich, T.; Garcia, R. I. Tooth Loss and Periodontal Disease Predict Poor Cognitive Function in Older Men. *J. Am. Geriatr. Soc.* **2010**, *58*, 713−718.

(8) Gatz, M.; Mortimer, J. A.; Fratiglioni, L.; Johansson, B.; Berg, S.; Reynolds, C. A.; Pedersen, N. L. Potentially Modifiable Risk Factors for Dementia in Identical Twins. *Alzheimer's Dementia* **2006**, *2*, 110−117.

(9) Stein, P. S.; Desrosiers, M.; Donegan, S. J.; Yepes, J. F.; Kryscio, R. J. Tooth Loss, Dementia and Neuropathology in the Nun Study. *J. Am. Dent. Assoc.* **2007**, *138*, 1314−1322.

(10) Kamer, A. R.; Pirraglia, E.; Tsui, W.; Rusinek, H.; Vallabhajosula, S.; Mosconi, L.; Yi, L.; McHugh, P.; Craig, R. G.; Svetcov, S.; Linker, R.; Shi, C.; Glodzik, L.; Williams, S.; Corby, P.; Saxena, D.; de Leon, M. J. Periodontal Disease Associates with Higher Brain Amyloid Load in Normal Elderly. *Neurobiol. Aging* **2015**, *36*, 627−633.

(11) Noble, J. M.; Borrell, L. N.; Papapanou, P. N.; Elkind, M. S. V.; Scarmeas, N.; Wright, C. B. Periodontitis Is Associated with Cognitive Impairment among Older Adults: Analysis of NHANES-III. *J. Neurol., Neurosurg. Psychiatry* **2009**, *80*, 1206−1211.

(12) Mayrand, D.; Holt, S. C. Biology of Asaccharolytic Black-Pigmented Bacteroides Species. *Microbiol. Rev.* **1988**, *52*, 134−152.

(13) Ally, N.; Whisstock, J. C.; Sleprawska-Lupa, M.; Potempa, J.; Le Bonniec, B. F.; Travis, J.; Pike, R. N. Characterization of the Specificity of Arginine-Specific Gingipains from *Porphyromonas gingivalis* Reveals Active Site Differences between Different Forms of the Enzymes. *Biochemistry* **2003**, *42*, 11693−11700.

(14) Eichinger, A.; Beisel, H. G.; Jacob, U.; Huber, R.; Medrano, F. J.; Banbula, A.; Potempa, J.; Travis, J.; Bode, W. Crystal Structure of Gingipain R: An Arg-Specific Bacterial Cysteine Proteinase with a Caspase-like Fold. *EMBO J.* **1999**, *18*, 5453−5462.

(15) Movilla, S.; Martí, S.; Roca, M.; Moliner, V. Unrevealing the Proteolytic Activity of RgpB Gingipain from Computational Simulations. *J. Chem. Inf. Model.* **2021**, *61*, 4582−4593.

(16) Wei, D.; Huang, X.; Liu, J.; Tang, M.; Zhan, C. G. Reaction Pathway and Free Energy Profile for Papain-Catalyzed Hydrolysis of N-Acetyl-Phe-Gly 4-Nitroanilide. *Biochemistry* **2013**, *52*, 5145−5154.
(17) Świderek, K.; Moliner, V. Revealing the Molecular Mechanisms of Proteolysis of SARS-CoV-2 Mproby QM/MM Computational Methods. *Chem. Sci.* **2020**, *11*, 10626−10630.
(18) 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*, 5585−5593.
(19) Ramos-Guzmán, C. A.; Zinovjev, K.; Tuñón, I. Modeling Caspase-1 Inhibition: Implications for Catalytic Mechanism and Drug Design. *Eur. J. Med. Chem.* **2019**, *169*, 159−167.
(20) 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*, 12740−12748.
(21) 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*, 18613−18622.
(22) Sulpizi, M.; Rothlisberger, U.; Carloni, P. Molecular Dynamics Studies of Caspase-3. *Biophys. J.* **2003**, *84*, 2207−2215.
(23) Miscione, G. P.; Calvaresi, M.; Bottoni, A. Computational Evidence for the Catalytic Mechanism of Caspase-7. A DFT Investigation. *J. Phys. Chem. B* **2010**, *114*, 4637−4645.
(24) Angelides, K. J.; Fink, A. L. Mechanism of Thiol Protease Catalysis: Detection and Stabilization of a Tetrahedral Intermediate in Papain Catalysis. *Biochemistry* **1979**, *18*, 2363−2369.
(25) Štrajbl, M.; Florián, J.; Warshel, A. Ab Initio Evaluation of the Free Energy Surfaces for the General Base/Acid Catalyzed Thiolysis of Formamide and the Hydrolysis of Methyl Thiolformate: A Reference Solution Reaction for Studies of Cysteine Proteases. *J. Phys. Chem. B* **2001**, *105*, 4471−4484.
(26) Arafet, K.; Ferrer, S.; Moliner, V. Computational Study of the Catalytic Mechanism of the Cruzain Cysteine Protease. *ACS Catal.* **2017**, *7*, 1207−1215.
(27) Parra-Torres, V.; Melgar-Rodríguez, S.; Muñoz-Manríquez, C.; Sanhueza, B.; Cafferata, E. A.; Paula-Lima, A. C.; Díaz-Zúñiga, J. Periodontal Bacteria in the Brain—Implication for Alzheimer's Disease: A Systematic Review. *Oral Dis.* **2021**, *29*, 21−28.
(28) Martí, S.; Arafet, K.; Lodola, A.; Mulholland, A. J.; Świderek, K.; Moliner, V. Impact of Warhead Modulations on the Covalent Inhibition of SARS-CoV-2 M pro Explored by QM/MM Simulations. *ACS Catal.* **2022**, *12*, 698−708.
(29) Gehringer, M.; Laufer, S. A. Emerging and Re-Emerging Warheads for Targeted Covalent Inhibitors: Applications in Medicinal Chemistry and Chemical Biology. *J. Med. Chem.* **2019**, *62*, 5673−5724.
(30) Arafet, K.; Serrano-Aparicio, N.; Lodola, A.; Mulholland, A. J.; González, F. V.; Świderek, 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**, *12*, 1433−1444.
(31) Jorgensen, W. L.; Ravimohan, C. Monte Carlo Simulation of Differences in Free Energies of Hydration. *J. Chem. Phys.* **1985**, *83*, 3050−3054.
(32) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc. Chem. Res.* **2000**, *33*, 889−897.
(33) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical p K a Predictions. *J. Chem. Theory Comput.* **2011**, *7*, 525−537.
(34) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(35) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696−3713.
(36) Joung, I. S.; Cheatham, T. E. Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *J. Phys. Chem. B* **2008**, *112*, 9020−9041.
(37) Woods, R.; Chappelle, R. Restrained Electrostatic Potential Atomic Partial Charges for Condensed-Phase Simulations of Carbohydrates. *J. Mol. Struct.: THEOCHEM* **2000**, *527*, 149−156.
(38) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098−3100.
(39) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37*, 785−789.
(40) 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*, 1157−1174.
(41) Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. An Overview of the Amber Biomolecular Simulation Package. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2013**, *3*, 198−210.
(42) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(43) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N · log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(44) Uberuaga, B. P.; Anghel, M.; Voter, A. F. Synchronization of Trajectories in Canonical Molecular-Dynamics Simulations: Observation, Explanation, and Exploitation. *J. Chem. Phys.* **2004**, *120*, 6363−6374.
(45) Sindhikara, D. J.; Kim, S.; Voter, A. F.; Roitberg, A. E. Bad Seeds Sprout Perilous Dynamics: Stochastic Thermostat Induced Trajectory Synchronization in Biomolecules. *J. Chem. Theory Comput.* **2009**, *5*, 1624−1631.
(46) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.
(47) Kräutler, V.; van Gunsteren, W. F.; Hünenberger, P. H. A Fast SHAKE Algorithm to Solve Distance Constraint Equations for Small Molecules in Molecular Dynamics Simulations. *J. Comput. Chem.* **2001**, *22*, 501−508.
(48) Verlet, L. Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.* **1967**, *159*, 98−103.
(49) Duan, L.; Liu, X.; Zhang, J. Z. H. Interaction Entropy: A New Paradigm for Highly Efficient and Reliable Computation of Protein−Ligand Binding Free Energy. *J. Am. Chem. Soc.* **2016**, *138*, 5722−5728.
(50) Loeffler, H. H.; Bosisio, S.; Matos, G. D. R.; Suh, D.; Roux, B.; Mobley, D. L.; Michel, J. Reproducibility of Free Energy Calculations across Different Molecular Simulation Software Packages. *J. Chem. Theory Comput.* **2018**, *14*, 5567−5582.
(51) Götz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. *J. Chem. Theory Comput.* **2012**, *8*, 1542−1555.
(52) Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878−3888.
(53) Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23*, 187−199.
(54) Martí, S. QMCube (QM3): An All-Purpose Suite for Multiscale QM/MM Calculations. *J. Comput. Chem.* **2021**, *42*, 447−457.

https://doi.org/10.1021/acs.jcim.2c01198
J. Chem. Inf. Model. 2023, 63, 950−958

(55) Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods V: Modification of NDDO Approximations and Application to 70 Elements. *J. Mol. Model.* **2007**, *13*, 1173−1213.
(56) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.
(57) Ditchfield, R.; Hehre, W. J.; Pople, J. A. Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1971**, *54*, 724−728.
(58) Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self—Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian—Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972**, *56*, 2257−2261.
(59) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. Efficient Diffuse Function-Augmented Basis Sets for Anion Calculations. III. The 3-21+G Basis Set for First-Row Elements, Li-F. *J. Comput. Chem.* **1983**, *4*, 294−301.
(60) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32*, 1456−1465.
(61) Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 1. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. *J. Chem. Inf. Model.* **2011**, *51*, 69−82.

https://doi.org/10.1021/acs.jcim.2c01198
J. Chem. Inf. Model. 2023, 63, 950−958