First QM/MM Studies of Inhibition Mechanism of Cruzain by Peptidyl Halomethyl Ketones.

Funding Source Statement: We thank the FEDER Spanish *Ministerio de Economía y Competitividad* for project CTQ2012-36253-C03-01; Generalitat Valenciana for *PrometeoII/2014/022* project; and Universitat Jaume I for project P11B2014-26.

Kemel Arafet, Silvia Ferrer and Vicent Moliner**

Departament de Química Física i Analítica, Universitat Jaume I, 12071 Castelló, Spain

Corresponding authors:

S. Ferrer: sferrer@uji.es

V. Moliner: moliner@uji.es

Phone Number: +34964728084

Abbreviations

1D: monodimensional
2D: bidimensional
E64: N-[N-[1-hydroxycarboxyethyl-carbony]leucylamino-butyl]-guanidine
FP2: falcipain-2
HL: high-level
LL: low-level
MD: molecular dynamics
OPLS-AA: optimized potential for liquid simulations all-atom force field
PClK: benzoil-tyrosine-alanine-chloromethyl ketone
PDB: protein data bank
PES: potential energy surface
FES: free energy surface
PFK: benzoil-tyrosine-alanine-fluoro-methyl ketone
PHK: peptidyl halomethyl ketones
PMF: Potential of Mean Force
QM/MM: Quantum Mechanics / Molecular Mechanics
QM/MM-FE: Quantum Mechanical/Molecular Mechanical-Free Energy
RC: reaction coordinate
RMSD: root-mean-square-deviation
TIP3P: transferable intermolecular potential 3P
TH: thiohemiketal
RH: stable intermediate
THH: protonated thiohemiketal
IH: protonated intermediate
I: unprotonated intermediate
TMSI: three-membered sulfonium intermediate
WHAM: weighted histogram analysis method

Abstract

Cruzain is a primary cysteine protease expressed by the protozoan parasite *Trypanosoma cruzi* during infection of Chagas disease and thus, the development of inhibitors of this protein is a promising target for designing an effective therapy against the disease. In this paper, the inhibition mechanism of cruzain by two different irreversible peptidyl halomethyl ketones (PHK) inhibitors has been studied by means of hybrid QM/MM molecular dynamics (MD) simulations to obtain a complete representation of the possible free energy reaction paths. These have been traced on Free energy surfaces (FES) in terms of Potential of Mean Force computed at AM1d/MM and DFT/MM level of theory. An analysis of the possible reaction mechanisms of the inhibition process has been performed showing that the nucleophilic attack of an active site cysteine, Cys25, on a carbon atom of the inhibitor and the cleavage of the halogen-carbon bond take place in a single step. PClK appears much more favorable than PFK from the kinetic point of view. This result would be in agreement with experimental studies in other papain-like enzymes. A deeper analysis of the results suggest that the origin of the differences between PClK and PFK can be on the different stabilizing interactions established between the inhibitors and the residues of the active site of the protein. Any attempt to explore the viability of the inhibition process through a step-wise mechanism involving the formation of a thiohemiketal intermediate and a three-membered sulfonium intermediate have been unsuccessful. Nevertheless, a mechanism through a protonated thiohemiketal, with participation of His159 as proton donor, appears to be feasible despite showing higher free energy barriers. Our results suggest that PClK can be used as starting point to develop a proper inhibitor of cruzain.

Keywords

Chagas disease, PHK inhibitors, QM/MM, Molecular Dynamics, PMF

Introduction

Chagas disease (American trypanosomiasis) is an illness caused by the protozoan parasite *Trypanosoma cruzi (T. cruzi)*, and was discovered in 1909 by the Brazilian doctor Carlos Chagas (1879-1934).[1] It is estimated that the illness affects about 7 million to 8 million people living mainly in endemic Latin American countries.[2] The movement of Chagas disease to areas previously considered non-endemic, resulting from increasing population mobility between Latin America and the rest of the world, represents a serious public health challenge.[3] There is not any available vaccine to prevent Chagas disease and the two available drugs for treatment, benznidazole and nifurtimox, are not only toxic and with important contra-indications (pregnancy, renal or hepatic failure, psychiatric and neuronal disorders), but also ineffective for the chronic stage of the disease.[4,5] Clearly, there is an urgent need for developing an effective therapy against Chagas disease. One approach consists of developing inhibitors of cruzain, the primary cysteine protease expressed by *T. cruzi* during infection.[6-8]

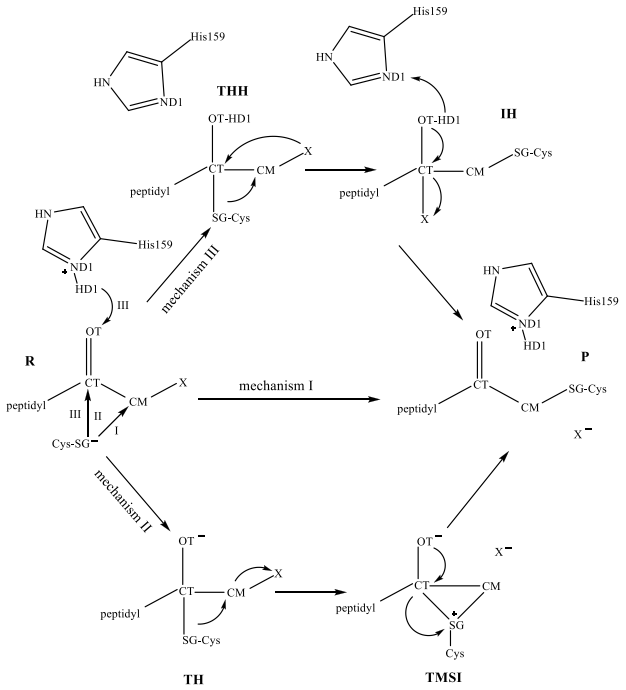
Cruzain, that belongs to the family of cysteine proteases, was initially discovered from the parasite cell-free extracts and subsequently heterologously expressed in Escherichia coli.[9,10] This protease is expressed in all life cycle stages of the parasite and is involved in nutrition and fight against host defense mechanisms.[11-14] Addition of a cruzain inhibitor to cultures of mammalian cells exposed to trypomastigotes or to mammalian cells already infected with *T. cruzi* amastigotes blocks replication and differentiation of the parasite, thus interrupting the parasite life cycle.[15-21]

Several groups have demonstrated that irreversible inhibition of cruzain by small molecules eradicates infection of the parasite in cell culture and animal models.[17,22-25] Irreversible inhibitors that contain an electrophilic functional group, such as vinyl sulfones, fluoro methyl ketones, aziridines or nitriles, covalently bind to cruzain via nucleophilic attack of the active site cysteine,[26] showing good inhibition activity.[27,28] In fact, to date, only irreversible inhibitors of cruzain have successfully cured parasitic infection,[23] implying that tight binding to the enzyme may be essential. In this sense, Shoellmann and Shaw developed in 1962 the first peptidyl chloromethyl ketones, L-1-tosyl-amido-2-phenylethyl chloromethyl ketone, as specific inhibitors for the serine protease chymotrypsin.[29] The major disadvantage of this inhibitor was their lack of

selectivity due the great chemical reactivity of the chloroketone functional group. The development of chloromethyl ketone inhibitors leds to the investigation of analogous inhibitor structures with different halo leaving groups replacing the chlorine atom. Bromomethyl and iodomethyl ketones have been synthesized and are typically more reactive but less stable in aqueous solutions. The first peptide fluoromethyl ketones inhibitors were reported in the literature by Rasnick in 1985[30] and by Shaw's group in 1986.[31] These peptide fluoromethyl ketones were shown to be highly reactive and selective irreversible inhibitors for cysteine proteases. They are poor irreversible inactivators for serine proteases and are not reactive toward bionucleophiles. Inhibition of cruzain activity with fluoromethyl ketone-based inhibitors seems to be correlated with the loss of feasibility of parasites in both *ex-vivo* tissue culture and *in vivo* mouse models.[19,32,33]

Information derived from the study of the molecular mechanism of hydrolysis catalyzed by cysteine proteases could be used as starting point to explore the inhibition mechanism at atomistic level. Early studies revealed a participation of the residues Cys25 and His159 (cruzain numbering) from the active site of theses proteases.[34-36] The catalytic cysteine mediates protein hydrolysis via a nucleophilic attack on the carbonyl carbon of a susceptible peptide bond. The imidazole group of the histidine polarizes the SH group of the cysteine and enables deprotonation even at neutral to weakly pH, and a highly nucleophilic thiolate/imidazolium ion pair is thereby produce.[37] The ion pair mechanism explains the unusually high reactivity of cysteine proteases toward electrophilic reagents in comparison to the nucleophilicity of the sulfur of cysteine or glutathione, especially in slightly acidic environments.[38]

Regarding the study of the inhibition mechanism of cysteine proteases, only few irreversible inhibitors have been studied theoretically.[39-46] The study of diazomethyl ketones based on gas phase MNDO calculations proposed a mechanism in which Cys25 residue binds to the inner diazo nitrogen thus forming an intermediate that evolves to a stable beta-thioketone and molecular nitrogen.[44] More recently, DFT and semiempirical methods have also been used to study the inhibition process by nitrile-containing compounds in gas phase.[41] Classical MD simulations and exploration of potential energy surfaces (PES) obtained with QM/MM potentials have been used to study the inhibition of cysteine proteases by epoxides and aziridine-based compounds.[42]



Scheme 1. Reaction mechanisms of inhibition of cysteine proteases by peptidyl halomethyl ketones (X: F, Cl). The mechanisms I and II are proposed by Powers and co-workers.[47] TH refers to thiohemiketal intermediate and TMSI to the three-membered sulfonium intermediate. The mechanism III is an alternative mechanism that takes place through a thiohemiketal protonated intermediate, THH.

Powers and co-workers proposed two possible mechanisms (mechanism I and II) of irreversible inhibition by peptidyl halomethyl ketones (PHK) based on different crystal structures of cysteine proteases (see Scheme 1).[47] The mechanism I, as shown in Scheme 1, is the direct displacement of the halide group by the thiolate anion. The mechanism II involves formation of a tetrahedral intermediate named thiohemiketal (TH) and a three-membered sulfonium intermediate (TMSI) that rearranges to give the final thioether adduct. The formation of the TH is equivalent to the presence of a tetrahedral intermediate in serine proteases mechanism.[27] Nevertheless, mechanism II has not been supported by theoretical studies. Kollman and co-workers, employing classical molecular mechanics simulations and semiempirical quantum mechanics with model systems to study the cysteine protease catalysis, showed that the attack of S^- on a carbonyl carbon does not involve a stable anionic tetrahedral structure.[48,49] Later, Suhai and co-workers in a QM/MM study of the active site of free papain and of the NMA-Papain complex did not obtain a stable tetrahedral NMA-papain complex.[50] Gao and Byun, in a combined QM/MM study of the nucleophilic addition reaction of methanethiolate and N-methylacetamide, concluded that there is no stable tetrahedral

intermediate in going from the reactant to the tetrahedral adduct.[51] Moreover, Warshel and co-workers in an *ab initio* study of general base/acid catalyzed thiolysis of formamide and the hydrolysis of methyl thiolformate shown that the anionic tetrahedral intermediate for the acylation reaction was found to be unstable in aqueous solution and to collapse immediately into the neutral form, which is the only intermediate on the reaction pathway.[52] Based on DFT methods with model compounds, Kolandaivel and co-workers explored the PESs of the inhibition of haloketones[46] and diketones[43] in gas phase. According to their results, the thiohemiketal protonated intermediate would be stable but, in general, less stable than the reactant.[43,46] More recently, Zhan and co-workers in a pseudobond first-principles QM/MM-FE study of the reaction pathway for papain-catalyzed hydrolysis of N-acetyl-Phe-Gly 4-nitroline, concluded that in the acylation step any path that include the formation of the TH do not exist.[53] Nevertheless, to date, the inhibition mechanism of cysteine protease by PHKs has not yet been studied by computational tools including the protein environment effects.

The principal aim of this paper is to gain insights into the mechanism whereby PHK inhibitors bound to a cysteine protease, in particular to cruzain. In a previous theoretical study in our laboratory, the mechanism inhibition of a protease involved in the degradation of human hemoglobin, namely falcipain-2 (FP2), by the epoxysuccinate N-[N-[1-hydroxycarboxyethyl-carbony]leucylamino-butyl]-guanidine, E64, was carried out.[39] The results indicated that the irreversible attack of a conserved cysteine to the E64 can take place on both carbon atoms of its epoxy ring since both processes present similar barriers. We herein present a theoretical study of the inhibition mechanism of cruzain by two PHK, benzoil-tyrosine-alanine-fluoromethyl ketone (Bz-Tyr-Ala-CH_2F, or PFK) and benzoil-tyrosine-alanine-chloromethyl ketone (Bz-Tyr-Ala-CH_2Cl, or PClK). The main difference between these two inhibitors and the previously studied E64 is the presence of the electrophilic functional group, an halogen atom in PHKs and an epoxide ring plus a carboxylic group in the E64. The studied proteins belong at the same family of cysteine proteases (clan CA) but are responsible of different human diseases, FP2 is associated to malaria and cruzain to Chagas disease. Hybrid Quantum Mechanical/Molecular Mechanical (QM/MM)[54-56] potentials have been employed to explore the PESs, and to run molecular dynamics (MD) simulations that allow obtaining the free energy profiles of the reaction in terms the Potential of Mean Force (PMF). The reaction mechanism, with their corresponding free energy barriers, averaged geometries

and interactions between the inhibitor and the protein have been the subject of a deep analysis in order to understand the molecular mechanism, paving the way to the design new inhibitors that allow an effective therapy against Chagas disease.

Computational Model

The initial coordinates were taken from the X-ray crystal structure of cruzain from *Trypanosoma cruzy* bound to Bz-Tyr-Ala-CH_2F with PDB entry 1AIM[32] and resolution 2.0 Å. The structure consists in one chain of 215 amino acids (cruzain), the inhibitor without the fluor atom and 46 water molecules (see Figure 1a).

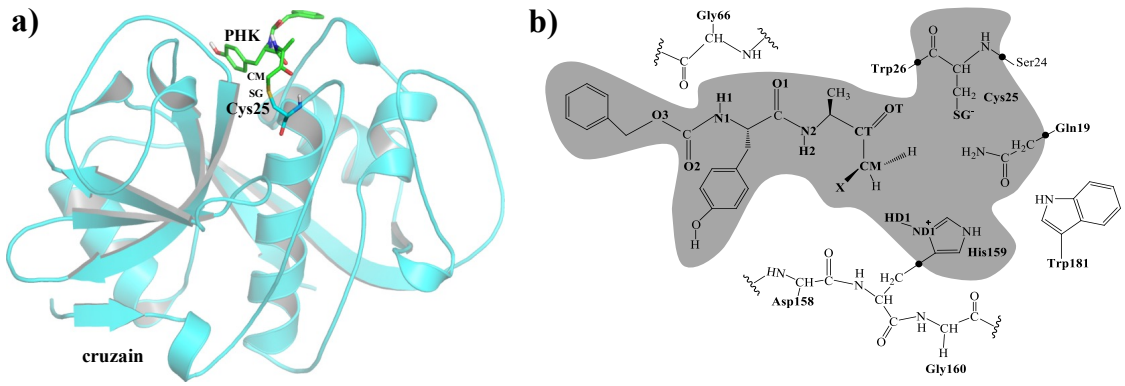


Figure 1. a) Monomer 1AIM structure showing the crystal structure of cruzain from *Trypanosoma cruzy* bound to Bz-Tyr-Ala-CH_2F (PHK). The inhibitor (CM) is covalently bound to the enzyme through the SG atom from residue Cys25. The inhibitor and Cys25 are represented by thick sticks. d) Details of the active site corresponding to the studied model. Grey region corresponds to the QM subset of atoms that includes the inhibitor Bz-Tyr-Ala-CH_2X (X: F, Cl), Cys25, the side chain of Gln19, and the imidazole ring of His159. The link atoms between the QM and the MM regions are indicated as black dots.

Cruzain is located in the intracellular vesicles during the dormant stage of the parasite at a slighly acidic pH. During the intracellular amastigote stage of the *T. Cruzi* life cycle, cruzain is present on the surface of the parasite where the pH of the host cytoplasm is almost neutral (pH 7.4).[32] Thus, hydrogen atoms were added at this physiological pH using fDYNAMO library,[57] according to the pKa values of the residues of 1AIM structure calculated within the empirical PROPKA 3.1 program of Jensen et al.[58-60] Fluor and chlorine atoms were added using Pymol[61] to create the two systems under study: Bz-Tyr-Ala-CH_2F and Bz-Tyr-Ala-CH_2Cl. A total of 12 counter ions (Na^+) were placed into optimal electrostatic positions around both systems, in order to obtain electro neutrality. Finally, the systems were placed in a 79.5 Å side cubic box of water molecules.

As depicted in Figure 1b, the inhibitor, Cys25, the side-chain of Gln19, and the imidazole ring of His159 (81 atoms) were described by means of the AM1d semiempirical Hamiltonian.[62] Gln19 was treated quantum mechanically to explore its active role in mechanism II, since formation of a bond or charge transfer from OT atom of inhibitor must be considered as a possibility. His159 was treated quantum mechanically to explore mechanism III. The rest of the protein and water molecules were described by OPLS-AA[63] and TIP3P[64] force fields, respectively. To saturate the valence of the QM/MM frontier we used the link atoms procedure.[55,65] Because of the size of the full system, all residues further than 25 Å from the CA1 atom of the inhibitor were kept frozen during the simulations (43169 atoms from a total of 50126). A force switch function with a cutoff distance in the range of 14.5 to 18 Å was applied to treat the non-bonding interactions. All the QM/MM calculations were carried out using fDYNAMO library.[57]

The two systems were relaxed by means of 1 ns of QM/MM MD at 300K using the NVT ensemble and the Langevin-Verlet integrator. Analysis of the time evolution of the root-mean-square-deviation (RMSD) of backbone atoms of the protein and the inhibitor for the Bz-Tyr-Ala-CH_2F and the Bz-Tyr-Ala-CH_2Cl systems confirms that the systems were equilibrated (see Figure S1 of Supporting Information). Structures obtained after relaxation were used to generate hybrid AM1d/MM PESs. Stationary structures (including reactants, products, intermediates and transition state structures) were located and characterized guided by means of a micro-macro iterations scheme.[66]

Once the PESs were explored, PMFs generated as a function of a distinguished reaction coordinate (RC) were obtained using the weighted histogram analysis method (WHAM) combined with the umbrella sampling approach.[67,68] The RC is defined depending on the chemical step under study. In general, for the monodimensional PMFs (1D-PMF) the value of the force constant used for the harmonic umbrella sampling was 2500 $kJ \cdot mol^{-1} Å^{-2}$ and the simulation windows consisted in 20 ps of equilibration and 40 ps of production, with a time step of 1 fs. Subsequently, 200 ps of AM1d/MM MD simulations of the stationary points were performed to analyze the main geometrical parameters as an average. For the study of mechanism I (see scheme 1), a 1D-PMF AM1d/MM was obtained using the antisymmetric combination of the bond-breaking and bond-forming distances, RC = (d(X-CM)-d(SG-CM)), where X means F or Cl. This required series of 92 and 85 simulation windows for Bz-Tyr-Ala-CH_2-F and Bz-Tyr-

Ala-CH_2-Cl systems respectively. For the mechanism II (see scheme 1), a 1D-PMF AM1d/MM was computed to study the formation of the TH from the reactant state using the antisymmetric combination of the bond-breaking and bond-forming distances, RC = (d(SG-CT)-d(CT-OT). This required series of 61 and 66 simulation windows for Bz-Tyr-Ala-CH_2-F and Bz-Tyr-Ala-CH_2-Cl systems, respectively. A 1D-PMF AM1d/MM was obtained for the formation of the TMSI from the product state using the SG-CT bond-breaking distance as RC. This required series of 41 and 37 simulation windows for Bz-Tyr-Ala-CH_2-F and Bz-Tyr-Ala-CH_2-Cl systems respectively. For the study of mechanism III (see scheme 1), a two-dimensional AM1d/MM PMF (2D-PMF) was computed for the formation of the THH intermediate from the reactant state using the ND1-HD1 bond-breaking distance as RC1 and SG-CT bond-forming distance as RC2. This required series of 1092 and 1288 simulation windows for Bz-Tyr-Ala-CH_2-F and Bz-Tyr-Ala-CH_2-Cl systems respectively. For the formation of IH intermediate from THH intermediate, an AM1d/MM 2D-PMF was obtained using the SG-CM bond-forming distance as RC1 and X-CM (X: F, Cl) bond-breaking distance as RC2. This required series of 651 simulation windows for both systems. Finally, a 2D-PMF AM1d/MM was obtained for the formation of the product state from IH intermediate using the ND1-HD1 bond-forming distance as RC1 and X-CT (X: F, Cl) bond-breaking distance as RC2. This required series of 1288 and 2310 simulation windows for Bz-Tyr-Ala-CH_2-F and Bz-Tyr-Ala-CH_2-Cl systems respectively.

Because of the large number of structures that must be evaluated during free energy calculations, QM/MM calculations are usually restricted to the use of semiempirical Hamiltonians. In order to correct the low-level AM1d energy function used in the 1D-PMFs, an interpolated correction scheme developed in our laboratory has been applied.[69] In this correction scheme, based on a method proposed by Truhlar et al. for dynamical calculations of gas phase chemical reactions,[70] the new energy function employed in the simulations is defined as:

$$E = E_{LL/MM} + S\left[\Delta E_{LL}^{HL}(\zeta)\right] \quad (1)$$

where *S* denotes a spline function, whose argument $\Delta E_{LL}^{HL}(\zeta)$ is a correction term taken as the difference between single-point calculations of the QM subsystem using a high-level (HL) method, and the low-level (LL) one. The correction term is expressed as a function of the distinguished reaction coordinate ζ (RC). When corrections were done

on 2D-PMF, the correction term is then expressed as a function of the two coordinates ζ_1, ζ_2 to generate the higher level 2D-PMF. The HL calculations were carry out by means of the M06-2X functional[71] with the 6-31+G(d,p) basis set[72] following Truhlar and co-workers suggestions.[71,73] The 6-311+G(2df,2p) basis set[74,75] was also checked for the F system (mechanism I), since this is the one specifically recommended for this element. Test on the use of both basis sets rendered almost quantitatively equivalent results (see Figure S2 of Supporting Information) and consequently the 6-31+G(d,p) basis set was used for all calculations. These calculations were carried out using the *Gaussian09* program.[76]

Results and Discussion

The first proposed mechanism of inhibition of cruzain by the two studied PHKs mechanism I in Scheme 1,[47] has been initially explored by computing the PESs. Once the stationary points were located on the PESs, the mechanisms of both systems were clearly identify as single step mechanisms, where the attack of Cys25 on carbon CM takes place concomitantly with the bond-breaking of the halogen with CM.

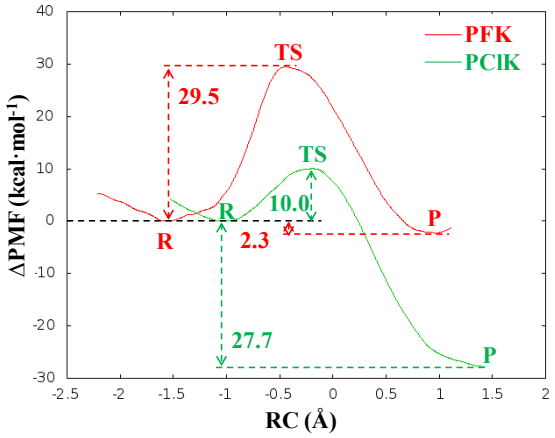


Figure 2. Free energy profiles of the cruzain inhibition with PHKs for mechanism I, computed at M06-2X/6-31+G(d,p)/MM level. RC corresponds to d(X-CM)-d(SG-CM), with X: F (red line), Cl (green line).

In the next step, the free energy profiles have been calculated in terms of 1D PMFs at AM1d/MM level (see Figure S3 of Supporting Information) and corrected at M06-2X/6-31+G(d,p)/MM level (Figure 2). The first observation is the dramatic influence that increasing the level of the QM region description has in the energy barriers and the reaction energies. This result justifies the use of high level Hamiltonians to properly

describe the energetics of this chemical reaction when studied by means of QM/MM methods. Consequently, the analysis of energies will be done on M06-2X/MM results. As observed in Figure 2, the free energy barrier for the inhibition of cruzain with PClK (ca. 10 kcal·mol^{-1}) is much lower than the one obtained with PFK (ca. 29.5 kcal·mol^{-1}). This result would describe a higher chemical reactivity of the peptide inhibitor with cloro than with fluor atoms. Despite no experimental kinetic data are available for the inhibition of cruzain with PHKs, experimental studies in Cathepsin B (papain-like enzymes known as clan CA)[77] concluded the higher reactivity of a PClK in comparison with a PFK,[27,30,31] in agreement with our results.

Analysis of the reaction energies obtained for both inhibitors shows that the process with PClK appears to be much more exergonic than the reaction with the PFK. Nevertheless, the reaction is irreversible in any case because the leaving anion will diffuse into the surrounding water shell due to the concentration gradients. Quantitative differences on reaction free energies obtained from the PMFs can be due to the limits of the distinguished reaction coordinate that, obviously, do not cover the full diffusion process of the anion into the bulk solvent. Interestingly, and as can be confirmed by the averaged values of the key distances reported in Table 1, the transition state in the PClK system appears at a more advanced value of the RC (-0.09 Å) than in the reaction of PFK (-0.48 Å). The fact that a more advanced TS was not associated with a higher activation barrier is due to the differences observed in the reactants complex: the minimum obtained in the PFK is located at a much negative value of the RC (-1.60 Å) than in the PClK (-1.03 Å). Interestingly, the change of the RC from reactants complex to transition state in both systems is very similar ($\Delta RC_{(R,TS)}$ = 0.94 Å and 1.12 Å for PClK and PFK, respectively). Then, the differences on the free energies of activation can not be completely discussed in these terms, and further analysis on average structures of stationary points of the free energy surfaces has to be carried out. Representative snapshots of reactants, transition state and products of both inhibitory reactions are presented in Figure 3. As observed in the figure, and quantitatively confirm from the geometrical data reported in Table 1, more hydrogen bond interactions between the inhibitor and the residues of the active site are detected in the TS of PClK (OT-H_{Cys25}, O1-H_{Trp26}, H1-O_{Gly66}) than in the TS of PFK (H2-O_{Gly66}) which is in agreement with the lower barrier obtained for the former reaction. The larger amount of favorable inhibitor-protein interactions obtained in products of the PClK than in the

products of PFK inhibitor would also support a higher stabilization of the PClK-protein complex (see Table 1). The halogen atoms in products appear to be stabilized mainly by interactions with water molecules, in both reactions. It can be observed that fluor ion interacts with a hydrogen H2 atom of the inhibitor and two water molecules. In the case of PClK, chlorine ion interacts, mainly, with His159 and two water molecules, but also presents weak interactions with Gln19 and Trp181. Thus, the thioether adduct (products) presents stronger interactions than reactants, supporting a description as an irreversible inhibition.

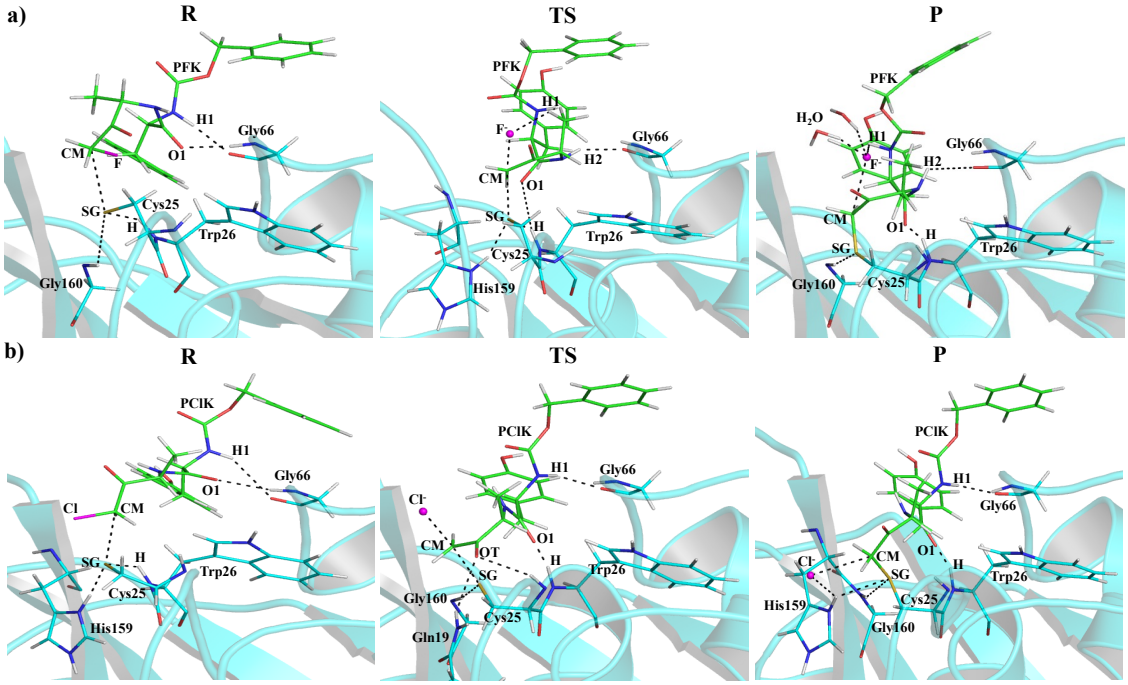


Figure 3. Representative snapshots of the key states of the reaction mechanism I of cruzain inhibition by (a) PFK and (b) PClK.

Table 1. Averaged distances for key states located along the inhibition of cruzain by PHK, through the direct mechanism I. Results were obtained from 200 ps of AM1d/MM MD simulations on the stationary points taking from the M06-2X/6-31+G(d,p)/MM free energy profiles (in Å). Interatomic distances associated to the RCs were constrained during the simulations.

a) PFK.

d (Å)	R	TS	P
SG-CM	3.15 ± 0.03	2.52 ± 0.03	1.79 ± 0.03
F-CM	1.55 ± 0.02	2.04 ± 0.03	2.70 ± 0.02
F-H1	4.58 ± 0.11	3.03 ± 0.14	2.55 ± 0.12
F-H2	3.18 ± 0.13	3.41 ± 0.15	6.24 ± 0.91
F-H (H_2O)	4.67 ± 0.99	3.81 ± 1.30	1.52 ± 0.09
F-H (H_2O)	4.23 ± 0.90	3.82 ± 0.74	1.75 ± 0.51
SG-H (Trp26)	1.54 ± 0.14	1.55 ± 0.20	3.15 ± 0.22
SG-H (Gly160)	3.10 ± 0.65	3.56 ± 0.81	2.63 ± 0.24
SG-HD1 (His159)	5.00 ± 0.38	3.17 ± 0.99	4.12 ± 0.28
OT-H (Cys25)	4.74 ± 0.40	5.46 ± 1.13	3.07 ± 0.35
OT-H (Gln19)	6.12 ± 0.82	7.75 ± 1.05	2.93 ± 0.47
O1-H (Trp26)	3.89 ± 0.29	3.94 ± 0.41	1.59 ± 0.23
O1-H (Gly66)	2.22 ± 0.19	5.55 ± 0.43	3.80 ± 0.40
N2-H (Trp26)	5.44 ± 0.24	4.09 ± 0.44	3.49 ± 0.34
N1-H (Gly66)	3.49 ± 0.31	5.23 ± 0.28	2.68 ± 0.40
H2-O (Gly66)	5.37 ± 0.25	2.18 ± 0.31	2.19 ± 0.67
H1-O (Gly66)	2.04 ± 0.24	3.58 ± 0.26	5.60 ± 0.54

b) PClK.

d (Å)	R	TS	P
SG-CM	2.80 ± 0.03	2.33 ± 0.03	1.78 ± 0.03
Cl-CM	1.77 ± 0.03	2.24 ± 0.03	3.10 ± 0.03
Cl-HD1 (His159)	3.50 ± 0.59	4.06 ± 0.36	2.40 ± 0.20
Cl-HE1 (Trp181)	5.10 ± 0.78	5.76 ± 0.62	3.78 ± 0.76
Cl-H (H_2O)	3.91 ± 0.73	3.36 ± 0.74	2.50 ± 0.26
SG-H (Trp26)	2.38 ± 0.54	2.95 ± 0.29	3.08 ± 0.20
SG-H (His159)	1.92 ± 0.10	5.14 ± 0.29	3.26 ± 0.46
SG-H (Gly160)	2.92 ± 0.66	2.51 ± 0.21	3.03 ± 0.43
OT-H (Cys25)	4.38 ± 0.98	2.55 ± 0.32	3.07 ± 0.64
OT-H (Gln19)	5.74 ± 1.69	2.54 ± 0.33	3.90 ± 0.82
O1-H (Gly66)	3.34 ± 0.73	3.58 ± 0.44	3.92 ± 0.36
O1-H (Trp26)	2.59 ± 1.15	1.65 ± 0.31	1.52 ± 0.20
O2-H (Gly66)	4.42 ± 0.45	2.99 ± 0.45	4.13 ± 0.29
O3-H (Gly66)	3.46 ± 0.68	2.99 ± 0.45	3.01 ± 0.42
N1-H (Gly66)	3.03 ± 0.38	2.74 ± 0.22	2.82 ± 0.24
H1-O (Gly66)	2.01 ± 0.23	2.06 ± 0.20	1.99 ± 0.17
H2-O (Asp158)	3.67 ± 0.47	3.21 ± 0.49	3.43 ± 0.38

As commented in the Introduction section, Powers and co-workers[47] proposed a second mechanism of inhibition of cruzain by PHKs (mechanism II in Scheme 1) involving the formation of a thiohemiketal intermediate (TH) and a three-membered sulfonium intermediate (TMSI). Nevertheless, all the attempts carried out to explore the mechanism by means of AM1d/MM PESs were unsuccessful. TH has not been found to be a stable intermediate in the PES. This has been confirmed when exploring the evolution of the PMF from reactants to the TH using the antisimmetric combination d(SG-CT)-d(CT-OT) as a reaction coordinate. The resulting PMFs obtained for both inhibitors at M06-2X/6-31+G(d,p)/MM (see Figure 4) confirm the instability of the proposed TH. The corresponding profiles at AM1d/MM level, reported in Figure S4 of Supporting Information, render the same conclusions.

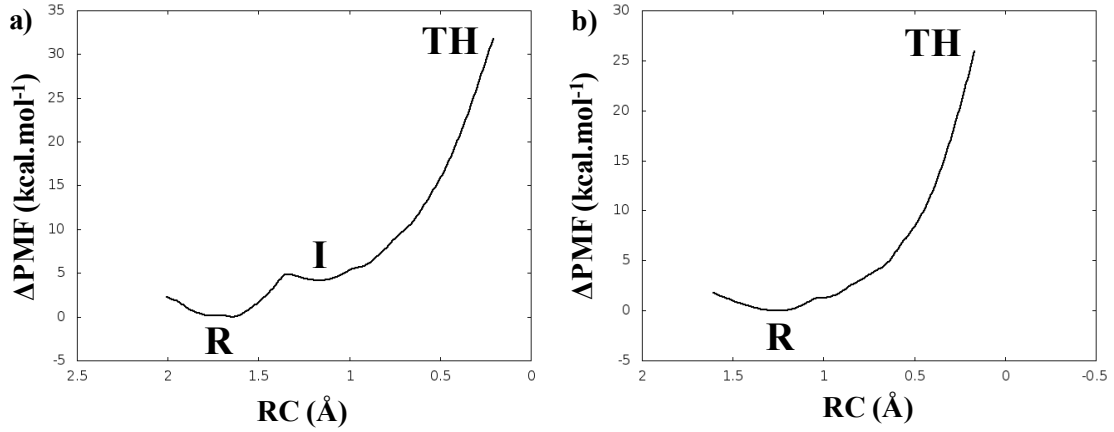


Figure 4. PMFs for the formation of TH intermediate from reactants computed at M06-2X/6-31+G(d,p)/MM level: a) Bz-Tyr-Ala-CH_2F inhibitor, and b) Bz-Tyr-Ala-CH_2Cl inhibitor. RC corresponds to d(SG-CT)-d(OT-CT).

The minimum that appears in Figure 4a at a RC value of 1.17 Å corresponds to an intermediate with distances SG-CT and SG-CM equal to 2.41 ± 0.03 and 3.06 ± 0.10 Å, respectively. Obviously, this structure does not correspond to the TH. Moreover, further attempts to explore this mechanism II were based on the exploration of the step from products to the TMSI. The PMFs obtained as a function of the SG-CT distance render profiles that do not allow localizing stable structures for this intermediate, as shown in Figure 5. The corresponding profiles at AM1d/MM level (Figure S5 of Supporting Information) render the same conclusions. Finally, any attempt to equilibrate the PHK-cruzain systems at TH or TMSI intermediates by means of constrained QM/MM MD

simulations reveal the instability of these structures as soon as all atoms of the system allowed to move (see Figures S6, S7 and S8 of Supporting Information)

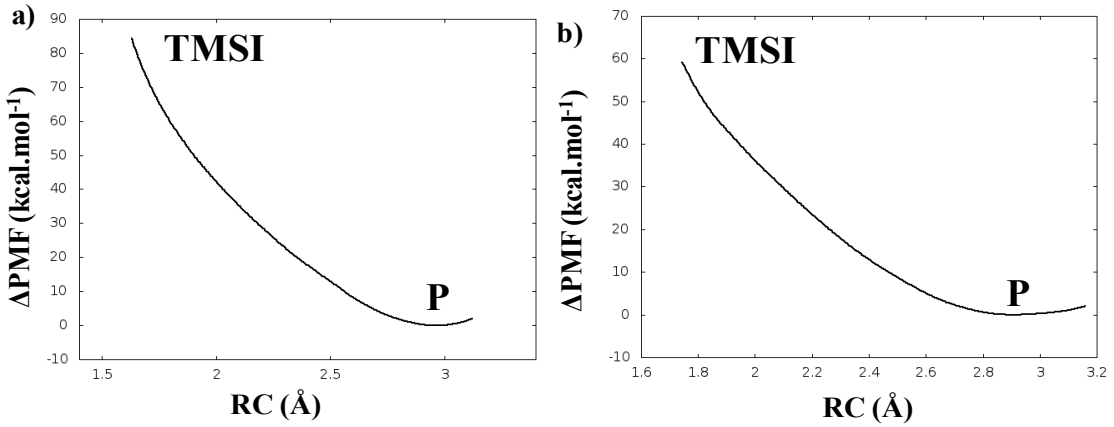


Figure 5. PMFs for the formation of TMSI intermediate from products, computed at M06-2X/6-31+G(d,p)/MM level: a) Bz-Tyr-Ala-CH_2F inhibitor, and b) Bz-Tyr-Ala-CH_2Cl inhibitor. RC corresponds to the distance d(SG-CT).

These results suggest that inhibition of cruzain by PHKs can not take place through the proposed mechanism II, a conclusion that is in agreement with previous theoretical studies of different systems showing that the attack of S^- on a C=O bond do not involve a stable anionic tetrahedral structure.[48-53] Moreover, notwithstanding experimental studies suggest the possibility of the mechanism II taking place in serine proteases through the formation of TH between the serine-OH and the C=O from the inhibitor;[78] no experimental evidence exist of this mechanism in cysteine proteases.

Finally, a third mechanism through a protonated thiohemiketal, THH, mechanism III depicted in Scheme 1, has been explored. From the computational point of view this mechanism appears to be more complex due to the participation of an acid, His159, that protonates the OT carbonyl oxygen of the inhibitor. Consequently, and as described in Computational Methods section, 2D-PMFs were computed to unequivocally define the mechanism. Thus, the first step corresponds to the attack of Cys25 on CM carbon atom and the proton transfer from His159 to OT oxygen leading to THH. The resulting FESs obtained at M06-2X/6-31+G(d,p)/MM level are shown in Figure 6. The corresponding surfaces at AM1d/MM level are reported in Figure S9 of Supporting Information. The first conclusion derived from Figure 6 is that this step takes place without any barrier when inhibition takes place with PClK. In the case of PFK, the proton from His159 is transferred to the OT atom of the inhibitor also without an energy barrier, but rendering an stable intermediate (RH in Figure 6a). Then, the attack of Cys25 to the CT atom

proceeds with a free energy barrier defined by TS1. Interestingly, both 2D-PMFs show the protonated thiohemiketal, THH, as an stable species and confirms the previous results (see Figure 4) that indicated the instability of the TH. Representative snapshots of this THH species obtained with both inhibitors are presented in Figure 7.

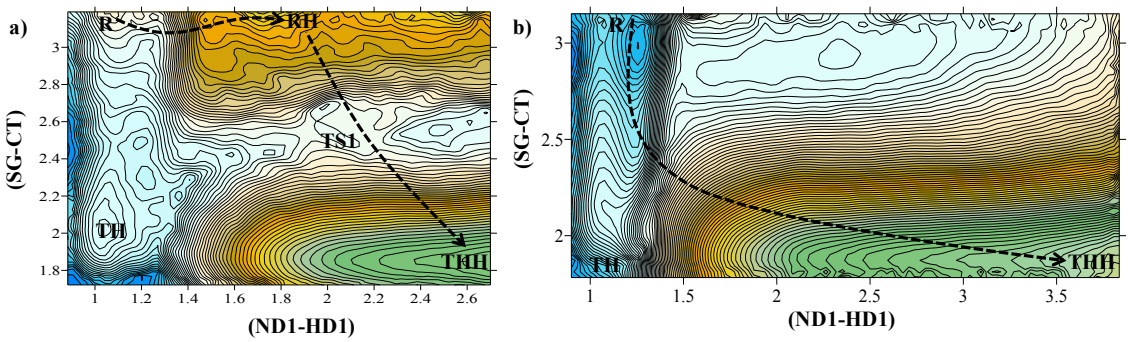


Figure 6. 2D-PMFs for the formation of protonated thiohemiketal intermediate, THH, from reactants, computed at M06-2X/6-31+G(d,p)/MM level: a) Bz-Tyr-Ala-CH_2F inhibitor, and b) Bz-Tyr-Ala-CH_2Cl inhibitor. Distances are in Å and iso-energetic lines are displayed every 0.5 kcal·mol^{-1}.

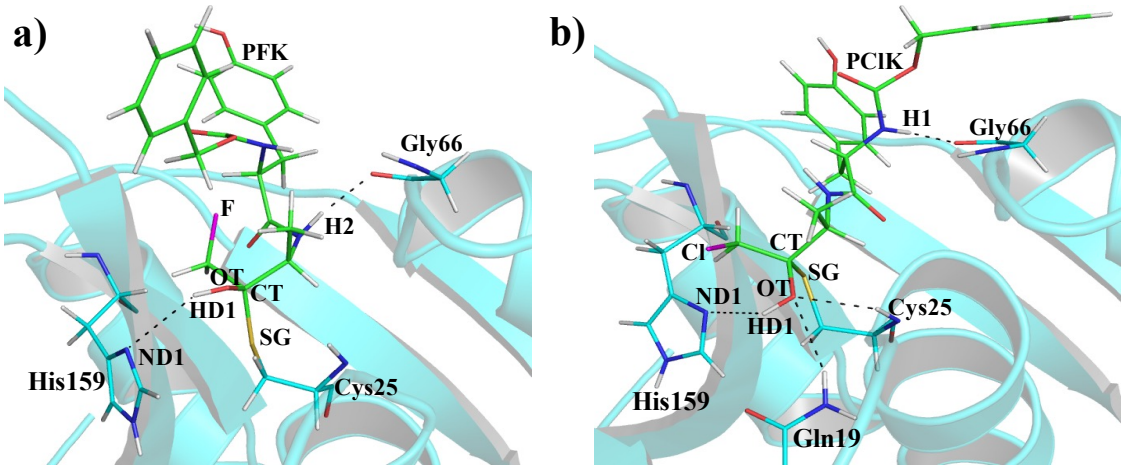


Figure 7. Representative snapshots of the protonated thiohemiketal intermediate, THH, located on the reaction mechanism III of cruzain inhibition by (a) PFK and (b) PClK.

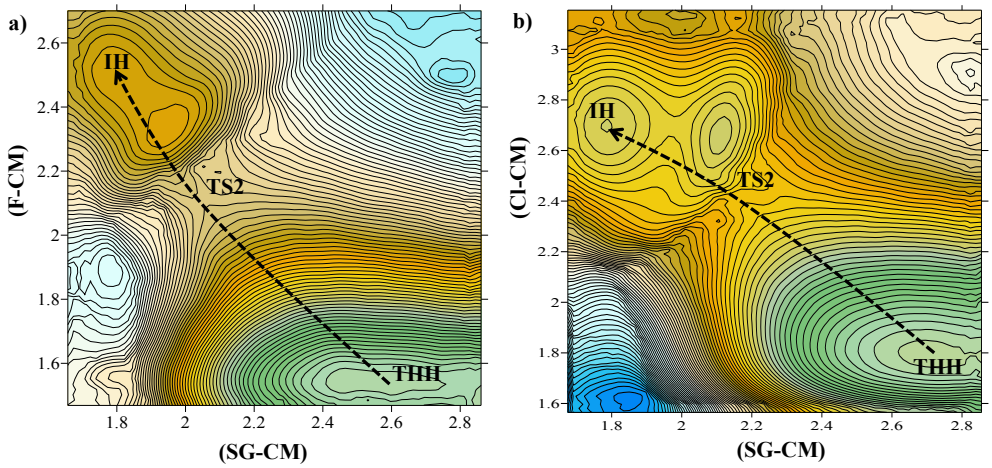


Figure 8. 2D-PMFs for the transformation from protonated thiohemiketal, THH, to intermediate IH, computed at M06-2X/6-31+G(d,p)/MM level: a) Bz-Tyr-Ala-CH_2F inhibitor, and b) Bz-Tyr-Ala-CH_2Cl inhibitor. Distances are in Å and iso-energetic lines are displayed every 1.0 kcal·mol^{-1}.

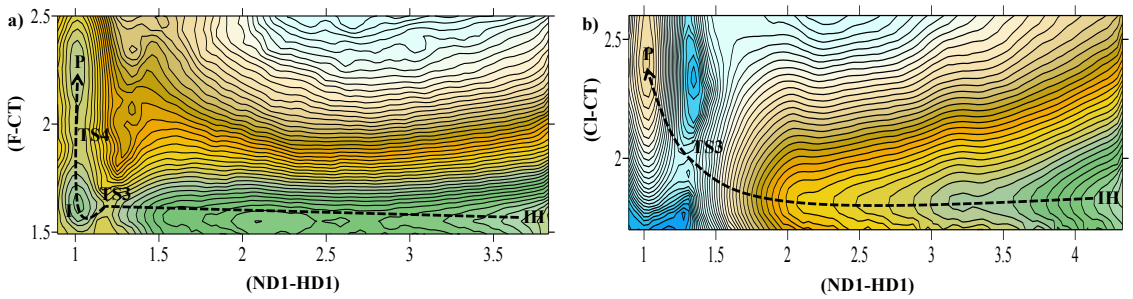


Figure 9. 2D-PMFs for the formation of products from IH intermediate computed at M06-2X/6-31+G(d,p)/MM level: a) Bz-Tyr-Ala-CH_2F inhibitor, and b) Bz-Tyr-Ala-CH_2Cl inhibitor. Distances are in Å and iso-energetic lines are displayed every 1.0 kcal·mol^{-1}.

The second step has been also explored by means of 2D-PMFs, controlling the X-CM and SG-CM distances. The resulting M06-2X/6-31+G(d,p)/MM FESs are presented in Figure 8 (the corresponding surfaces at AM1d/MM are reported in Figure S10 of Supporting Information). In this step, the mechanisms of both inhibitors proceed in an almost equivalent way. Finally, when exploring the last step of the inhibition of cysteine protease by PFK and PClK inhibitors, differences are found. By comparing Figure 9a and 9b, it is observed how the transformation from IH to products takes place in a concerted way for PClK but through a stable intermediate, I, for PFK. Same conclusions are derived form the AM1d/MM free energy surfaces reported in Supporting Information as Figure S11.

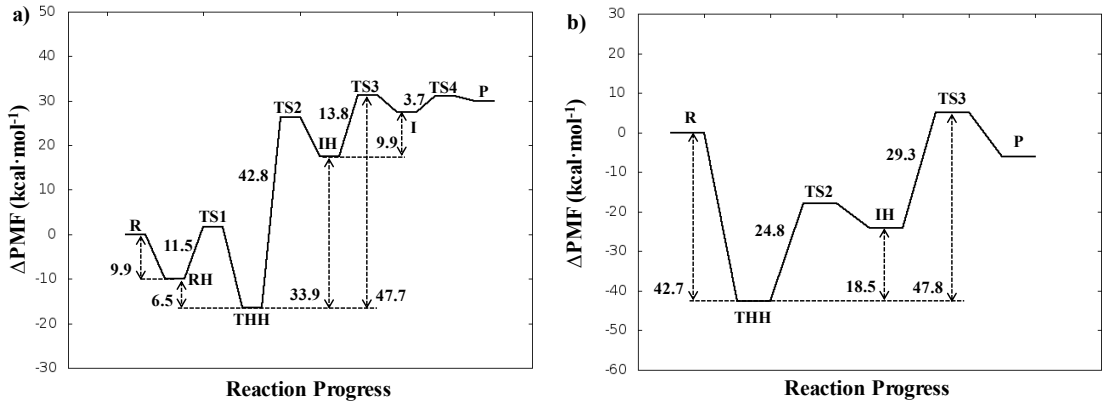


Figure 10. Free energy profiles for the inhibition of cruzain through mechanism III computed at M06-2X/6-31+G(d,p)/MM level by: a) Bz-Tyr-Ala-CH_2F inhibitor, and b) Bz-Tyr-Ala-CH_2Cl inhibitor.

The complete free energy profiles for the inhibition of cruzain through mechanism III, derived from the M06-2X/MM 2D-PMFs reported in Figure 6, 8 and 9, are summarized in Figure 10. It can be concluded that, as was observed for mechanism I, the inhibition of cruzain with PClK is much more favorable than the inhibition with PFK, once again in agreement with experimental studies of papain-like enzyme studies,[77] concluding the higher reactivity of a PClK in comparison with a PFK.[27,30,31] The rate limiting step of mechanism III would correspond to TS3 for both inhibitors with effective free energy barriers, computed from the stabilized THH, of 47.7 and 47.8 kcal·mol^{-1} for PFK and PClK, respectively. Nevertheless, as observed on the profiles, if reactants is considered as the reference state, the inhibition of cruzain with PClK is clearly more feasible than the inhibition with PFK.

From the thermodynamics point of view, the energetic results derived from the FESs in this mechanism would describe an endorgenic process when inhibition takes place with PFK but an exergonic one for the reaction with the PClK. Anyway, as mentioned above, the reaction must be irreversible because the leaving anion will diffuse into the surrounding water shell due to the concentration gradients. Quantitative differences obtained from the PMFs can be associated to the limits of the distinguished reaction coordinate in the last step that, obviously, do not cover the full diffusion process of the leaving anion.

Analysis of geometries of stationary point structures (reported in Table S1 of Supporting Information) reveal more hydrogen bond interactions between the inhibitor

and the residues of the active site in the TS3 of PClK than in the TS3 of PFK, which is in agreement with the lower energies obtained for the former reaction. Again, and as observed in the results of mechanism I, the larger amount of favorable inhibitor-protein interactions obtained in products of the PClK than in the products of PFK inhibitor is in agreement with the higher stabilization of the PClK-protein complex (see Table S1).

Conclusions

This paper reports the theoretical study of the inhibition of cruzain by two irreversible peptidyl halomethyl ketones inhibitors, Bz-Tyr-Ala-CH_2F, or PFK, and Bz-Tyr-Ala-CH_2Cl, or PClK, carried out by means of MD simulations using hybrid AM1d/MM and M06-2X/MM potentials. The reaction mechanisms previously proposed in the literature[47] have been explored. According to our results, the nucleophilic attack of the unprotonated Cys25 on CM atom of the inhibitor (PFK or PClK) and the halogen-CM breaking bond take place in a concerted way. Analysis of free energy barriers and reaction free energies, computed at both levels of theory, shows that the inhibition by PClK would be kinetically and thermodynamically more favourable. In fact, the difference between the barriers of both reactions is almost three folds (10.0 and 29.5 $kcal \cdot mol^{-1}$ for PClK and PFK, respectively) and even larger differences are obtained when comparing the reaction energies stabilized (-27.7 and -2.3 $kcal \cdot mol^{-1}$ for PClK and PFK, respectively). It is important to point out that the use of a semiempirical Hamiltonian to treat those atoms described quantum mechanically in our QM/MM scheme provoke an important overestimation of the free energy barriers, by comparison with the use of DFT methods. Our results confirm the great chemical reactivity of PClK as an irreversible inhibitor of cruzain. Despite no experimental kinetic data are available for this particular reaction, our results would be in agreement with experimental studies in other papain-like enzymes (Cathepsin B) that concluded the higher reactivity of a PClK in comparison with a PFK. A deeper analysis of the results suggest that the origin of these differences can be on the different stabilizing interactions established between the inhibitors and the residues of the active site of the protein. Any attempt to explore the viability of the inhibition process through a step-wise mechanism involving the formation of a thiohemiketal intermediate and a three-membered sulfonium intermediate have been unsuccessful. All QM/MM calculations show that these intermediates are unstable thus not recommending the use of the intermediate as a

template to design an efficient inhibitor. Nevertheless, an alternative step-wise mechanism has been located through a protonated thiohemiketal intermediate. This species, which is generated though a proton transfer from a histidine residue located in the active site of cruzain protein, His159, and the attack of Cys25 to CT atom of the inhibitor, appears to be dramatically stabilized with respect to the initial reactants species. Our results suggest that, in fact, this covalently bounded inhibitor-protein complex could be a final state of the inhibition process since significant barriers are obtain in going from this complex to the final product complex.

Altogether, our results suggest that benzoil-tyrosine-alanine-chloromethyl ketone is a good candidate to develop a proper inhibitor of cruzain, which in turn could be used in therapies against Chagas desease.

Acknowledgements

K.A. thanks the Spanish *Ministerio de Economía y Competitividad* for a predoctoral contract. The authors also acknowledge the Servei d´Informatica, Universitat Jaume I for generous allotment of computer time.

Supporting Information: Figure S1; Time evolution of the root-mean-square-deviation along the QM/MM MD simulation for the backbone atoms of the protein (red line) and atoms of the inhibitor (blue line) for both systems. Figure S2; 1D-PMFs for cruzain inhibition by PFK through mechanism I computer at M062X/MM level with 6-31+G(d,p) and 6-311+G(2df,2p) basis set. Figure S3; AM1d/MM 1D-PMFs of the cruzain inhibition by PHKs through mechanism I. Figure S4; AM1d/MM 1D-PMFs of the formation of TH intermediate from reactants for both inhibitors. Figure S5; AM1d/MM 1D-PMFs of the formation of TMSI intermediate from products for both inhibitors. Figure S6; Time dependent evolution of selected distances on the QM/MM MD simulation corresponding with the TH of both inhibitors. Figure S7; Time dependent evolution of selected distances on the QM/MM MD simulation corresponding with the TMSI of the Bz-Tyr-Ala-CH_2F inhibitor. Figure S8; Time dependent evolution of selected distances on the QM/MM MD simulation corresponding with the TMSI of the Bz-Tyr-Ala-CH_2Cl inhibitor. Figure S9; AM1d/MM 2D-PMFs for the formation of THH intermediate from reactants for both inhibitors. Figure S10; AM1d/MM 2D-PMFs for the formation of IH intermediate from

THH for both inhibitors. Figure S11; AM1d/MM 2D-PMFs for the formation of P from IH intermediate for both inhibitors. Table S1. Averaged distances for key states located along the inhibition of cruzain by PHK, through a mechanism III. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Clayton, J. (2010) Chagas disease 101, *Nature 465*, S4-S5.

(2) http://www.who.int/mediacentre/factsheets/fs340/en. (2014).

(3) Jackson, Y., Angheben, A., Carrilero Fernandez, B., Jansa i Lopez del Vallado, J. M., Jannin, J. G., and Albajar-Vinas, P. (2009) Management of Chagas disease in Europe. Experiences and challenges in Spain, Switzerland and Italy, *Bulletin de la Societe de pathologie exotique (1990) 102*, 326-329.

(4) Castro, J. A., de Mecca, M. M., and Bartel, L. C. (2006) Toxic side effects of drugs used to treat Chagas' disease (American trypanosomiasis), *Human & Experimental Toxicology 25*, 471-479.

(5) Filardi, L. S., and Brener, Z. (1987) Susceptibility and natural-resistance of trypanosoma-cruzi strains to drugs used clinically in chagas-disease, *Transactions of the Royal Society of Tropical Medicine and Hygiene 81*, 755-759.

(6) McKerrow, J. H., Engel, J. C., and Caffrey, C. R. (1999) Cysteine protease inhibitors as chemotherapy for parasitic infections, *Bioorganic & Medicinal Chemistry 7*, 639-644.

(7) McKerrow, J. H., McGrath, M. E., and Engel, J. C. (1995) The cysteine protease of trypanosoma-cruzi as a model for antiparasite drug design, *Parasitology Today 11*, 279-282.

(8) Renslo, A. R., and McKerrow, J. H. (2006) Drug discovery and development for neglected parasitic diseases, *Nature Chemical Biology 2*, 701-710.

(9) Eakin, A. E., Mills, A. A., Harth, G., McKerrow, J. H., and Craik, C. S. (1992) The sequence, organization, and expression of the major cysteine protease (cruzain) from trypanosoma-cruzi, *Journal of Biological Chemistry 267*, 7411-7420.

(10) Itow, S., and Camargo, E. P. (1977) Proteolytic activities in cell-extracts of trypanosoma-cruzi, *Journal of Protozoology 24*, 591-595.

(11) Campetella, O., Henriksson, J., Aslund, L., Frasch, A. C. C., Pettersson, U., and Cazzulo, J. J. (1992) The major cysteine proteinase (cruzipain) from trypanosoma-cruzi is encoded by multiple polymorphic tandemly organized genes located on different chromosomes, *Molecular and Biochemical Parasitology 50*, 225-234.

(12) Fujii, N., Mallari, J. P., Hansell, E. J., Mackey, Z., Doyle, P., Zhou, Y. M., Gut, J., Rosenthal, P. J., McKerrow, J. H., and Guy, R. K. (2005) Discovery of potent thiosemicarbazone inhibitors of rhodesain and cruzain, *Bioorganic & Medicinal Chemistry Letters 15*, 121-123.

(13) Scharfstein, J., Schechter, M., Senna, M., Peralta, J. M., Mendoncapreviato, L., and Miles, M. A. (1986) Trypanosoma-cruzi characterization and isolation of a 57/51,000 MW surface glycoprotein (GP57/51) expressed by epimastigotes and bllod-stream trypomastigotes, *Journal of Immunology 137*, 1336-1341.

(14) Schnapp, A. R., Eickhoff, C. S., Sizemore, D., Curtiss, R., and Hoft, D. F. (2002) Cruzipain induces both mucosal and systemic protection against Trypanosoma cruzi in mice, *Infection and Immunity 70*, 5065-5074.

(15) Cazzulo, J. J. (2002) Proteinases of Trypanosoma cruzi: patential targets for the chemotherapy of Changas desease, *Current topics in medicinal chemistry 2*, 1261-1271.

(16) Cazzulo, J. J., Stoka, V., and Turk, V. (1997) Cruzipain, the major cysteine proteinase from the protozoan parasite Trypanosoma cruzi, *Biological Chemistry 378*, 1-10.

(17) Engel, J. C., Doyle, P. S., Hsieh, I., and McKerrow, J. H. (1998) Cysteine protease inhibitors cure an experimental Trypanosoma cruzi infection, *Journal of Experimental Medicine 188*, 725-734.
(18) Engel, J. C., Doyle, P. S., Palmer, J., Hsieh, I., Bainton, D. F., and McKerrow, J. H. (1998) Cysteine protease inhibitors alter Golgi complex ultrastructure and function in Trypanosoma cruzi, *Journal of Cell Science 111*, 597-606.
(19) Harth, G., Andrews, N., Mills, A. A., Engel, J. C., Smith, R., and McKerrow, J. H. (1993) Peptide-fluoromethyl ketones arrest intracellular replication and intercellular transmission of trypanosoma-cruzi, *Molecular and Biochemical Parasitology 58*, 17-24.
(20) McKerrow, J. H. (1999) Development of cysteine protease inhibitors as chemotherapy for parasitic diseases: insights on safety, target validation, and mechanism of action, *International Journal for Parasitology 29*, 833-837.
(21) Polticelli, F., Zaini, G., Bolli, A., Antonini, G., Gradoni, L., and Ascenzi, P. (2005) Probing the cruzain S-2 recognition subsite: A kinetic and binding energy calculation study, *Biochemistry 44*, 2781-2789.
(22) Barr, S. C., Warner, K. L., Kornreic, B. G., Piscitelli, J., Wolfe, A., Benet, L., and McKerrow, J. H. (2005) A cysteine protease inhibitor protects dogs from cardiac damage during infection by Trypanosoma cruzi, *Antimicrobial Agents and Chemotherapy 49*, 5160-5161.
(23) Brak, K., Doyle, P. S., McKerrow, J. H., and Ellman, J. A. (2008) Identification of a new class of nonpeptidic inhibitors of cruzain, *Journal of the American Chemical Society 130*, 6404-6410.
(24) Doyle, P. S., Zhou, Y. M., Engel, J. C., and McKerrow, J. H. (2007) A cysteine protease inhibitor cures Chagas' disease in an immunodeficient-mouse model of infection, *Antimicrobial Agents and Chemotherapy 51*, 3932-3939.
(25) Jacobsen, W., Christians, U., and Benet, L. Z. (2000) In vitro evaluation of the disposition of a novel cysteine protease inhibitor, *Drug Metabolism and Disposition 28*, 1343-1351.
(26) Roush, W. R., Cheng, J. M., Knapp-Reed, B., Alvarez-Hernandez, A., McKerrow, J. H., Hansell, E., and Engel, J. C. (2001) Potent second generation vinyl sulfonamide inhibitors of the trypanosomal cysteine protease cruzain, *Bioorganic & Medicinal Chemistry Letters 11*, 2759-2762.
(27) Otto, H. H., and Schirmeister, T. (1997) Cysteine proteases and their inhibitors, *Chemical Reviews 97*, 133-171.
(28) Lecaille, F., Kaleta, J., and Bromme, D. (2002) Human and parasitic papain-like cysteine proteases: Their role in physiology and pathology and recent developments in inhibitor design, *Chemical Reviews 102*, 4459-4488.
(29) Schoellmann, G., and Shaw, E. (1963) Direct evidence for presence of histidine in active center of chymotrypsin *Biochemistry 2*, 252-255.
(30) Rasnick, D. (1985) Systhesis of peptide fluoromethyl ketones and the inhibition of human cathepsin-B *Analytical Biochemistry 149*, 461-465.
(31) Rauber, P., Angliker, H., Walker, B., and Shaw, E. (1986) The synthesis of peptidylfluoromethanes and their properties as inhibitors of serine proteinases and cysteine proteinases, *Biochemical Journal 239*, 633-640.
(32) Gillmor, S. A., Craik, C. S., and Fletterick, R. J. (1997) Structural determinants of specificity in the cysteine protease cruzain, *Protein Science 6*, 1603-1611.
(33) McGrath, M. E., Eakin, A. E., Engel, J. C., McKerrow, J. H., Craik, C. S., and Fletterick, R. J. (1995) The crystal-structure od cruzain- A therapeutic target for chagas-disease, *Journal of Molecular Biology 247*, 251-259.
(34) Gilles, A. M., and Keil, B. (1984) Evidence for an active-center cysteine in the SH-proteinase alpha-clostripan through use of N-tosyl-L-lysine chloromethyl ketone, *Febs Letters 173*, 58-62.
(35) Polgar, L. (1973) Mode of activation of catalytically essential sulfhydryl group of papain, *European Journal of Biochemistry 33*, 104-109.

(36) Polgar, L., and Halasz, P. (1982) Current problems in mechanistic studies of serine and cysteine proteinases, *Biochemical Journal 207*, 1-10.
(37) Keillor, J. W., and Brown, R. S. (1992) Attack of zwitterionic ammonium thiolates on a distorted anilide as a model for the acylation of papain by amides- A simple demonstration of a bell.shaped pH rate profile, *Journal of the American Chemical Society 114*, 7983-7989.
(38) Brocklehurst, K. (1979) Specific covalent modification of thiols-applications in the study of enzymes and other biomolecules, *International Journal of Biochemistry 10*, 259-274.
(39) Arafet, K., Ferrer, S., Martí, S., and Moliner, V. (2014) Quantum Mechanics/Molecular Mechanics Studies of the Mechanism of Falcipain-2 Inhibition by the Epoxysuccinate E64, *Biochemistry 53*, 3336-3346.
(40) Barreiro, G., deAlencastro, R. B., and Neto, J. D. D. (1997) A semiempirical study on leupeptin: An inhibitor of cysteine proteases, *International Journal of Quantum Chemistry 65*, 1125-1134.
(41) Mendez-Lucio, O., Romo-Mancillas, A., Medina-Franco, J. L., and Castillo, R. (2012) Computational study on the inhibition mechanism of cruzain by nitrile-containing molecules, *Journal of Molecular Graphics & Modelling 35*, 28-35.
(42) Mladenovic, M., Ansorg, K., Fink, R. F., Thiel, W., Schirmeister, T., and Engels, B. (2008) Atomistic insights into the inhibition of cysteine proteases: First QM/MM calculations clarifying the stereoselectivity of epoxide-based inhibitors, *Journal of Physical Chemistry B 112*, 11798-11808.
(43) Shankar, R., Kolandaivel, P., and Senthilkumar, K. (2010) Reaction Mechanism of Cysteine Proteases Model Compound HSH With Diketone Inhibitor PhCOCOCH3-nXn, (X = F, Cl, n=0, 1, 2), *International Journal of Quantum Chemistry 110*, 1660-1674.
(44) Tarnowska, M., Oldziej, S., Liwo, A., Kania, P., Kasprzykowski, F., and Grzonka, Z. (1992) MNDO study of the mechanism of the inhibition of cysteine proteinases by diazomethyl ketones *European Biophysics Journal with Biophysics Letters 21*, 217-222.
(45) Vicik, R., Helten, H., Schirmeister, T., and Engels, B. (2006) Rational design of aziridine-containing cysteine protease inhibitors with improved potency: Studies on inhibition mechanism, *Chemmedchem 1*, 1021-1028.
(46) Vijayakumar, S., and Kolandaivel, P. (2008) Reaction mechanism of HSH and CH3SH with NH2CH2COCH2X (X = F and Cl) molecules, *International Journal of Quantum Chemistry 108*, 927-936.
(47) Powers, J. C., Asgian, J. L., Ekici, O. D., and James, K. E. (2002) Irreversible inhibitors of serine, cysteine, and threonine proteases, *Chemical Reviews 102*, 4639-4750.
(48) Arad, D., Langridge, R., and Kollman, P. A. (1990) A simulation of the sulfur attack in the cataliytic pathway of papain using molecular mechanics and semiempirical quantum-mechanics, *Journal of the American Chemical Society 112*, 491-502.
(49) Howard, A. E., and Kollman, P. A. (1988) OH-versus-SH nucleophilic attack on amides-dramatically different gas-phase and solvation energetics, *Journal of the American Chemical Society 110*, 7195-7200.
(50) Han, W. G., Tajkhorshid, E., and Suhai, S. (1999) QM/MM study of the active site of free papain and of the NMA-papain complex, *J. Biomol. Struct. Dyn. 16*, 1019-1032.
(51) Byun, K., and Gao, J. L. (2000) A combined QM/MM study of the nucleophilic addition reaction of methanethiolate and N-methylacetamide, *Journal of Molecular Graphics & Modelling 18*, 50-55.
(52) Strajbl, M., Florian, J., and Warshel, A. (2001) 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, *Journal of Physical Chemistry B 105*, 4471-4484.

(53) Wei, D. H., Huang, X. Q., Liu, J. J., Tang, M. S., and Zhan, C. G. (2013) Reaction Pathway and Free Energy Profile for Papain-Catalyzed Hydrolysis of N-Acetyl-Phe-Gly 4-Nitroanilide, *Biochemistry 52*, 5145-5154.
(54) Warshel, A., and Levitt, M. (1976) Theoretical studies of enzymic reactions - dielectric, electrostatic and steric stabilization of carbonium-ion reaction lysozyme *Journal of Molecular Biology 103*, 227-249.
(55) Field, M. J., Bash, P. A., and Karplus, M. (1990) A combined quantum-mechanical and molecular mechanical potential for molecular-dynamics simulations, *Journal of Computational Chemistry 11*, 700-733.
(56) Gao, J. L., Amara, P., Alhambra, C., and Field, M. J. (1998) A generalized hybrid orbital (GHO) method for the treatment of boundary atoms in combined QM/MM calculations, *Journal of Physical Chemistry A 102*, 4714-4721.
(57) Field, M. J., Albe, M., Bret, C., Proust-De Martin, F., and Thomas, A. (2000) The Dynamo library for molecular simulations using hybrid quantum mechanical and molecular mechanical potentials, *J. Comp. Chem. 21*, 1088-1100.
(58) Bas, D. C., Rogers, D. M., and Jensen, J. H. (2008) Very fast prediction and rationalization of pK(a) values for protein-ligand complexes, *Proteins-Structure Function and Bioinformatics 73*, 765-783.
(59) Li, H., Robertson, A. D., and Jensen, J. H. (2005) Very fast empirical prediction and rationalization of protein pK(a) values, *Proteins-Structure Function and Bioinformatics 61*, 704-721.
(60) Olsson, M. H. M., Sondergaard, C. R., Rostkowski, M., and Jensen, J. H. (2011) PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pK(a) Predictions, *Journal of Chemical Theory and Computation 7*, 525-537.
(61) www.pymol.org. (2009).
(62) Nam, K., Cui, Q., Gao, J., and York, D. M. (2007) Specific reaction parametrization of the AM1/d Hamiltonian for phosphoryl transfer reactions: H, O, and P atoms, *Journal of Chemical Theory and Computation 3*, 486-504.
(63) Jorgensen, W. L., Maxwell, D. S., and TiradoRives, J. (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids, *Journal of the American Chemical Society 118*, 11225-11236.
(64) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys. 79*, 926-935.
(65) Singh, U. C., and Kollman, P. A. (1986) A combined abinitio quantum-mechanical and molecular mechanical method for carryng out simulations on complex molecular-systems - Applications to the CH3Cl + Cl- Exchange-reaction and gas-phase protonation of polyethers, *Journal of Computational Chemistry 7*, 718-730.
(66) Martí, S., Moliner, V., and Tuñón, I. (2005) Improving the QM/MM description of chemical processes: A dual level strategy to explore the potential energy surface in very large systems, *Journal of Chemical Theory and Computation 1*, 1008-1016.
(67) Kumar, S., Bouzida, D., Swendsen, R. H., Kollman, P. A., and Rosenberg, J. M. (1992) The weighted histogram analysis method for free-energy calculations on biomolecules. 1. The method, *J. Comp. Chem. 13*, 1011-1021.
(68) Torrie, G. M., and Valleau, J. P. (1977) Non-physical sampling distributions in monte-carlo free-energy estimation - Umbrella Sampling, *J. Comp. Phys. 23*, 187-199.
(69) Ruiz-Pernia, J. J., Silla, E., Tuñón, I., Martí, S., and Moliner, V. (2004) Hybrid QM/MM potentials of mean force with interpolated corrections, *Journal of Physical Chemistry B 108*, 8427-8433.
(70) Chuang, Y. Y., Corchado, J. C., and Truhlar, D. G. (1999) Mapped interpolation scheme for single-point energy corrections in reaction rate calculations and a critical evaluation of dual-level reaction path dynamics methods, *Journal of Physical Chemistry A 103*, 1140-1149.

(71) Zhao, Y., and Truhlar, D. G. (2008) The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals, *Theoretical Chemistry Accounts 120*, 215-241.

(72) Hehre, W. J. R., L.; Schleyer, P. v.; Pople, J. A. (1986) *Ab Initio Molecular Orbital Theory*, New York.

(73) Lynch, B. J., Zhao, Y., Truhlar, D. G. (2003) Effectiveness of Diffuse Basis Functions for Calculating Relative Energies by Density Functional Theory, *Journal of Physical Chemistry A 107*, 1384-1388.

(74) Frisch, M. J., Pople, J. A., and Binkley, J. S. (1984) Self-Consistent Molecular-Orbital Methods.25. Supplementary functions for Gaussian-Basis Sets, *Journal of Chemical Physics 80*, 3265-3269.

(75) Krishnan, R., Binkley, J. S., Seeger, R., and Pople, J. A. (1980) Self-Consistent Molecular-Orbital Methods .20. Basis set for correlated wave-functions, *Journal of Chemical Physics 72*, 650-654.

(76) Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., and Nakatsuji, H. C., M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. . (2009) Gaussian 09. Revision A.1.

(77) Rawlings, N. D., Barrett, A. J., and Bateman, A. (2012) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors, *Nucleic Acids Research 40*, D343-D350.

(78) Malthouse, J. P. G., Mackenzie, N. E., Boyd, A. S. F., and Scott, A. I. (1983) Detection of a tetrahedral adduct in a trypsin-chloromethyl ketone specific inhibitor complex by C-13 NMR, *Journal of the American Chemical Society 105*, 1685-1686.

Graphic for the Table of Contents