
Article

pH-dependent structural dynamics of cathepsin D-family aspartic peptidase of *Clonorchis sinensis*

Jung-Mi Kang^{1,2}, Hương Giang Lê^{1,2}, Byoung-Kuk Na^{1,2} and Won Gi Yoo^{1,2,*}

¹Department of Parasitology and Tropical Medicine, Institute of Health Sciences, Gyeongsang National University College of Medicine, Jinju 52727, Republic of Korea; jmkang@gnu.ac.kr (J.-M.K.); gianglee291994@gmail.com (H.G.L.); bkna@gnu.ac.kr (B.-K.Na)

²Department of Convergence Medical Science, Gyeongsang National University, Jinju 52727, Republic of Korea

* Correspondence: wgyoo@gnu.ac.kr (W.G.Y.); Tel.: +82-55-772-8101

Abstract: Cathepsin D (CatD; EC 3.4.23.5) family peptidases of parasitic organisms are regarded as potential drug targets as they play critical roles in the physiology and pathobiology of parasites. Previously, we characterized the biochemical features of cathepsin D isozyme 2 (CatD2) in the carcinogenic liver fluke *Clonorchis sinensis* (CsCatD2). In this study, we performed all-atomic molecular dynamics simulations by applying different systems for the ligand-free/bound forms under neutral and acidic conditions to investigate the pH-dependent structural alterations and associated functional changes in CsCatD2. CsCatD2 showed several distinctive characteristics as follows: 1) CsCatD2-inhibitor complex formed more hydrogen bonds; 2) acidic pH caused major conformational transitions from open to closed state in this enzyme; 3) neutral pH induced displacement of the N-terminal part to hinder the accessibility of the active site and open allosteric site of this enzyme; and 4) the flap dynamics metrics, including distance (d_1), $\text{TriC}\alpha$ angles (θ_1 and θ_2), and dihedral angle (ϕ), account for the asymmetrical twisting motion of the active site of this enzyme. These findings provide an in-depth understanding of the pH-dependent structural dynamics of CsCatD2 and basic information for the rational design of an inhibitor as a drug targeting CsCatD2.

Keywords: *Clonorchis sinensis*; cathepsin D; aspartic peptidase; molecular dynamics simulation; pH effect; flap dynamics

1. Introduction

Clonorchiasis is a parasitic disease caused by the liver fluke *Clonorchis sinensis*. This parasite is prevalent in far east Asian countries, including China, Korea, and northern Vietnam, and it is reported to infect approximately 35 million people worldwide [1]. Infections in humans usually occur because of consumption of raw or inadequately cooked freshwater fish carrying the *C. sinensis* metacercariae. After the infection, the metacercariae exist in the duodenum of the human body and the juvenile worms migrate through the ampulla of Vater to reach the common bile duct. *C. sinensis* worms induce a series of pathological changes in the bile duct, resulting in epithelial hyperplasia, periductal

fibrosis, obstructive jaundice, dyspepsia, and cirrhosis of the liver [2]. Chronic clonorchiasis can induce diverse complications, such as periductal inflammation, fibrosis, cholangitis, cholelithiasis, and cholangiectasis [1,3,4]. *C. sinensis* has been classified as Group I biocarcinogen that promotes cholangiocarcinoma in humans by the World Health Organization [5].

Peptidases of parasitic helminths are one of the most extensively studied molecules because of their pivotal pathobiological roles in the physiology and nutrition of parasites as well as host-parasite interactions [6–15]. Among various peptidases, cathepsin D (CatD; EC 3.4.23.5) is considered an attractive target for the design of vaccines or anthelmintic drugs [16–19], as it is an essential component for the degradation of the host hemoglobin or proteins within the gut of blood-feeding flukes [18,20–22]. Moreover, helminth CatDs are pathogenic factors that can trigger molecular pathogenesis [23] and perform diverse pathophysiological roles across different developmental stages in the parasitic helminths [24]. Transcriptome analysis of *C. sinensis* revealed that most CatDs show relatively higher abundance in metacercariae than in adults and eggs, and their expression patterns are quite distinct from those of other species of peptidases, such as cysteine peptidases, metallopeptidases, serine peptidases, and threonine peptidases [25]. Recently, we characterized the biochemical and immunological properties of two CatDs of *C. sinensis*, cathepsin D isozyme 1 (CsCatD1) and cathepsin D isozyme 2 (CsCatD2) [26]. CsCatD2 is likely to play an important role as a digestive enzyme as it is mainly expressed in the intestinal epithelial cells of *C. sinensis* adult worms. This enzyme is also expressed in the eggs, suggesting its possible role in the production or maturation of eggs of the parasite.

Enzyme-substrate binding is a dynamic process in which enzymes undergo structural changes, resulting in the disruption of chemical bonds in the substrates and/or formation of new ones to yield the end products [27]. The active site pocket of typical CatDs is covered by the Y flap region (commonly known as the antiparallel β -sheet) that plays an important role in binding of the substrate [28]. Understanding the flap dynamics of CatDs is essential to gain an in-depth knowledge of the molecular aspects of enzyme-substrate interactions, particularly in pH-dependent structural alterations of the enzyme. It may also provide crucial information for the design of optimized aspartic peptidase inhibitors as drugs targeting the CatDs of the parasite.

In this study, we performed all-atomic molecular dynamics (MD) simulations in the ligand-free and bound forms of the CsCatD2 homology models under acidic and neutral conditions to explore the molecular features of flap dynamics and study the pH-induced structural alterations of CsCatD2.

2. Results and Discussion

2.1. Homology models and features of mature CsCatD2

The models for free-CsCatD2 (CsCatD2_{free}) and bound-CsCatD2 (CsCatD2_{bound}) were built based on the experimental X-ray characterized structures of *Ixodes ricinus* cathepsin D 1 (IrCD1), without the inhibitor

[Protein Data Bank (PDB) code: 5n7n] and in complex with the pepstatin (PDB code: 5n7q) from the tick *I. ricinus* (Figures S1 and S2). The IrCD1 structures were chosen as the best templates via the protein-Basic Local Alignment Search Tool (BLASTP) against the PDB database [29]. To generate the CsCatD2_{bound} structure model, the pepstatin from the template (PDB code: 5n7q) was superimposed on the homology model of CsCatD2 during the modeling process. Mature CsCatD2 showed a conserved bilobed structure consisting of two antiparallel β -sheet domains packed against each other. The interdomain connects the N- and C-terminal domains and each contributes one catalytic aspartate residue, aspartate 33 (Asp₃₃) and aspartate 219 (Asp₂₁₉), respectively, thereby forming the active site pocket. The Y flap region and polyproline loop covered this active site pocket to facilitate substrate binding (Figure 1a). Glycine 79 (Gly₇₉) was situated as a flap tip residue, whereas methionine 297 (Met₂₉₇) was placed as a hinge residue on the opposite side.

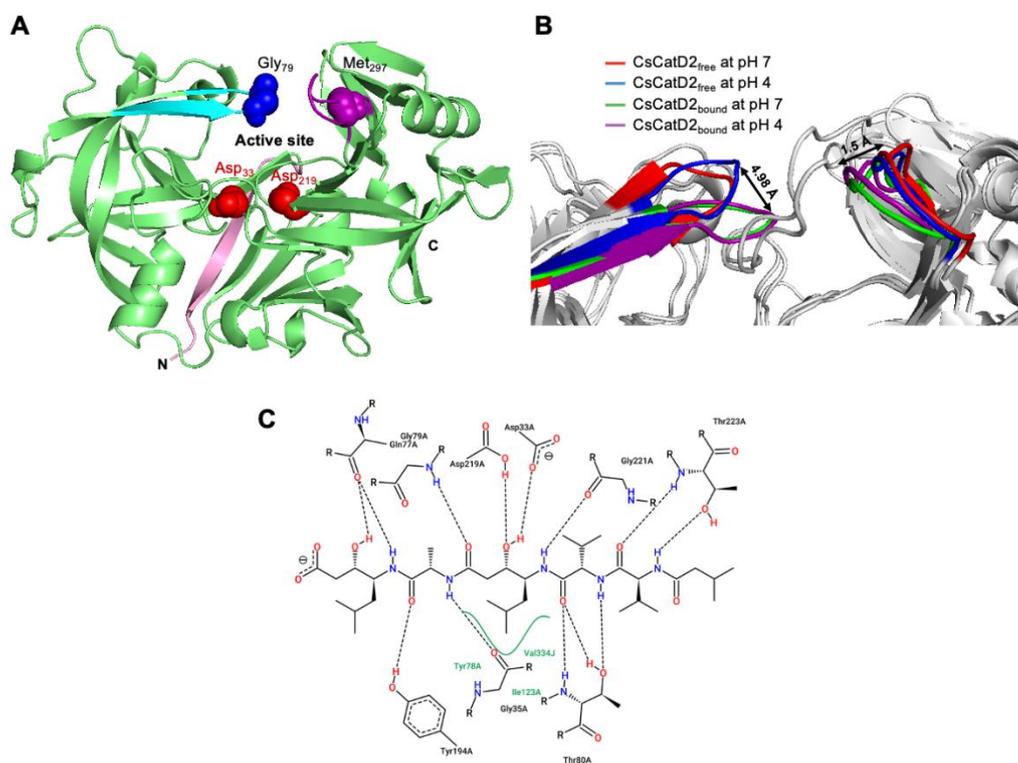


Figure 1. Conserved flap parameters and four entrance loops of mature CsCatD2. (a) In mature CsCatD2, the Y flap region (β -sheet in cyan) and polyproline loop (purple) and two catalytic aspartic residues (red sphere) are marked, respectively. The flap tip residue and hinge residue of the polyproline loop are indicated as blue and purple spheres, respectively. (b) Distances of tip-to-tip and hinge-to-hinge between all CsCatD2 models. Flap tip and hinge residues indicate Gly₇₉ and Met₂₉₇, respectively. (c) Two-dimensional map showing interaction between CsCatD2 and pepstatin.

2.2. Hydrogen bond formation between CsCatD2 and pepstatin

Pepstatin is an inhibitor of both subfamilies of aspartic proteases, A1 and A2 [30]. To explore potential anti-*C. sinensis* compounds targeting CsCatD2, it is necessary to understand the structure of mature CsCatD2 complexed with pepstatin. Binding of the inhibitor to the active site of

CsCatD2 induced conformational changes in the flap region containing conserved Gly⁷⁹ and the polyproline loop, which recoiled toward the inhibitor. Between CsCatD2_{free} and CsCatD2_{bound}, the tip-to-tip distance and the hinge-to-hinge distance at pH 4.0 were 4.98 Å and 1.5 Å, respectively (Figure 1b). The hydrogen bonds and hydrophobic interactions were predicted to be predominantly responsible for the intermolecular interactions that maintained the stability of the CsCatD2–pepstatin complex (Figure 1c). The bound form adopted a closed conformation as direct hydrogen bonds between the backbone carboxyl and amine groups of flap residues (glycine 77 (Gln⁷⁷) and Gly⁷⁹) and the bound pepstatin fixed the flap in a closed state. The intermolecular interactions between CsCatD2 and pepstatin were similar to those found in *Homo sapiens* CatDs (HsCatDs) complexed with respective inhibitors [31]. However, binding of CsCatD2 to pepstatin increased the number of hydrogen bonds because of additional interactions between the inhibitor and the side chain of the flap residue Gln⁷⁷.

2.3. Four systems for MD simulations according to the enzyme conformations and pH values

The proteolytic activity of CatDs is generally optimum at low pH and can be detected at neutral pH for a limited period of time [32]. There have been several reports regarding the optimal pH for a CatD enzymatic activity which are mentioned as follows: *Trichomonas vaginalis* CatD proteolytic activity is high at an acidic pH (3.5–4.5) and low at a pH close to 7 [33]; optimal pH of HsCatD is in the pH range 3.0–4.5 [34]. Considering the optimal pH ranges studied previously, we selected acidic (pH 4) and neutral (pH 7) conditions to assess the influence of different pH values on the stability and flexibility of CsCatD2. In addition, it was reported that a 50 nanosecond (ns)-MD simulation is a sufficiently long simulation time to observe all the conformational changes of aspartic peptidases [31,35]. The MD trajectories were analyzed in terms of stability, flexibility, and compactness, including the root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration (Rg), and solvent-accessible surface area (SASA).

2.4. CsCatD2_{free} adopts a semi-closed conformation for ligand-binding at acidic pH

The conformational stability of CsCatD2_{free} was evaluated by calculating the backbone RMSD by superposing the MD trajectories onto each initial structure at 0 ns. Simulations at the bound states were stable regardless of pH, as evidenced by low and relatively constant RMSD values (1.48 ± 0.01 , pH 4, and 1.53 ± 0.01 , pH 7) of the trajectories (Figure 2a). CsCatD2_{free} was relatively unstable at pH 7 (average RMSD: 2.36 ± 0.02) throughout the simulation, which was revealed by the deflection at the beginning of the simulation. However, it remained stable at pH 4 (average RMSD: 1.90 ± 0.02), comparable to the RMSD level of the bound form until approximately 32.0–35.5 ns. At these specific points-of-time, the RMSD was abruptly increased by 0.8 Å during which a transition in the tertiary structure may have occurred. The RMSD was then maintained at a steady level upon adopting a new conformation, as supported by the increasing number of hydrogen bonds in CsCatD2_{free} at pH 4 (Figure S3). These results suggest that acidification can contribute

to major conformational changes in CsCatD2_{free} from open to closed/semi-closed conformations.

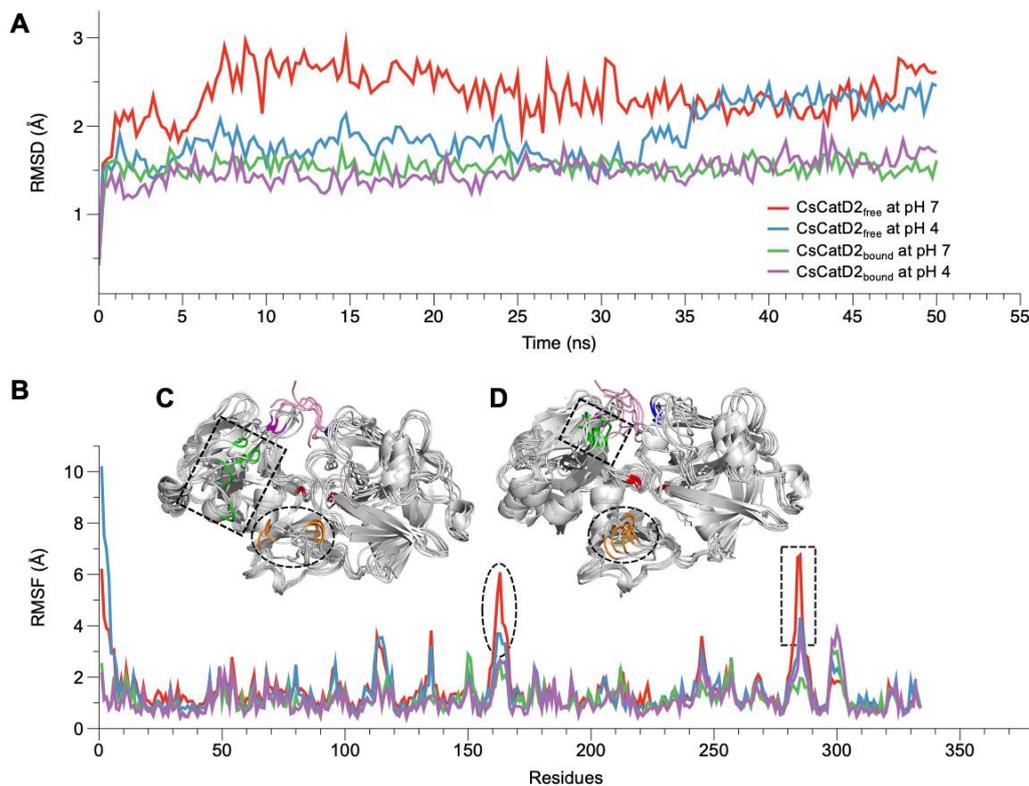


Figure 2. RMSD and RMSF variations of CsCatD2_{free} and CsCatD2_{bound} models at both pH conditions. During the 50-ns MD simulation at neutral (pH 7) and acidic (pH 4) condition, fluctuations of RMSD (a) and RMSF (b) are shown based on the time and residues, respectively. (c–d) Structural variations of CsCatD2_{free} at pH 7 (c) and pH 4 (d) values during the MD simulation. Backbone snapshots of both proteins are in shades of gray, each of which were obtained at 0, 10, 20, 30, 40 and 50 ns. The back view is rotated by 180° about a vertical axis, relative to the view shown in Figure 3a. The dotted circle and rectangle correspond with each region having high RMSF value, such as aa162–163 (orange part) and aa284–285 (green part). The remaining color legend is provided in the legend to Figure 3a.

2.5. Both ligand-binding and acidic pH limit the conformational flexibility

RMSF provides insights regarding the flexible or disordered regions of molecules in a biological system during MD simulations [36]. In this study, residual atomic fluctuations in all systems showed similar trends below 4 Å throughout the course of the simulations, except for CsCatD2_{free} at pH 7 (Figure 2b). Overall, the free forms at pH 7 (average RMSF: 1.53) and pH 4 (average RMSF: 1.44) showed higher flexibility than bound forms at both pH values (average RMSF: 1.11). Surprisingly, two regions (aa 162–163 and aa 284–285) exhibited larger deviations of up to 7 Å from its starting structure in CsCatD2_{free} at pH 7 than at pH 4 (Figure 2c and d). ₁₆₂DV₁₆₃ lies in the flexible loop region between the N- and C-termini, whereas ₂₈₄YR₂₈₅ belongs to the β-loop region and is located near the polyproline loop. No such RMSF patterns were observed in the bound forms under both pH conditions. This was in accordance with a previous report that flexibility occurs in some regions of free HsCatD owing to the absence of ligands [31]. Thus, it may be suggested

that pH itself can affect major conformational fluctuations in CsCatD2_{free} as shown by the increase in pH toward neutral conditions, which caused great fluctuations in the two regions in CsCatD2_{free}. However, the active site residue Asp₃₃ (RMSF: 0.86 Å) showed a slightly higher atomic fluctuation than Asp₂₁₉ (RMSF: 0.81 Å) in CsCatD2_{free} and vice versa in CsCatD2_{bound} (RMSF: 0.44 Å vs 0.48 Å, respectively). The salt bridge effects of the Asp dyad upon binding of the inhibitor may result in a decrease in flexibility (Figure 1c).

2.6. Neutral pH may induce the opening of potential allosteric inhibitory sites

Based on the average B-factor models for all systems, the N-terminal part covered the entrance of the active site in CsCatD2_{free} (Figure 3a and b), whereas there was no steric hindrance involving the entrance in the CsCatD2_{bound} (Figure 3c and d). In the free forms, three flexible regions (N-terminus, aa 162–163, and aa 284–285) were detected and the latter two regions showed higher mobility at pH 7 than at pH 4 in the bound forms as shown by the RMSF values (Figure 3a and b). The bound forms have flexible loops, including a polyproline loop (aa 296–302), as compared to free forms (Figure 3c and d). Remarkably, the region flanking aa 149–152 exhibited high flexibility only in CsCatD2_{bound} at pH 7. This highly mobile region was located at the boundary of the potential allosteric inhibitory pocket, all of which were matched between CsCatD2 [glycine 14 (Gln₁₄), tyrosine 15 (Tyr₁₅), tyrosine 16 (Tyr₁₆), phenylalanine 32 (Phe₃₂), alanine 97 (Ala₉₇), valine 149 (Val₁₄₉), glutamine (Glu₁₇₀), isoleucine 171 (Ile₁₇₁), and phenylalanine 173 (Phe₁₇₃)] and IrCatD1 [valine 39 (Val₃₉), tyrosine 40 (Tyr₄₀), tyrosine 41 (Tyr₄₁), phenylalanine 57 (Phe₅₇), alanine 120 (Ala₁₂₀), leucine 172 (Leu₁₇₂), glutamine 193 (Glu₁₉₃), valine 194 (Val₁₉₄), and phenylalanine 196 (Phe₁₉₆)] with a template modeling (TM) score of 93% (Figure 3c and Figure S4).

For CsCatD2_{free}, as described above, no hydrogen bond formation of flap tip Gly₇₉ (average RMSF: 1.55) with the inhibitor can have a critical impact on the overall flexibility and polyproline loop as compared to the CsCatD2_{bound} (average RMSF: 0.66, Gly₇₉) under both pH conditions. These findings are consistent with the fact that the flexible loop folds inward over the active site and interacts with the flap as the HsCatD closes [31,37]. Surprisingly, several highly flexible regions, such as aa 162–163 and aa 284–285 in CsCatD2_{free} and CsCatD2_{bound}, respectively, at pH 7 and aa 149–152 in CsCatD2_{bound} at pH 7, were found. These findings together with recent reports indicating that the allosteric inhibition mechanism regulates pepsin-family peptidases [32,38] suggest that neutral pH can induce the opening of the allosteric inhibitory site of CsCatD2 similar to HsCatD and IrCatD [32,39]. Ordinarily, a pH shift toward neutral condition displaces the N-terminus of mature CatDs into the active site and opens the allosteric site as the potency of the allosteric inhibitor is pH-dependent [32]. Thus, an inhibitor (pepstatin) may preferentially bind to the active site of CatDs at an acidic pH, whereas allosteric inhibitors (peptidyl inhibitors) favorably bind to allosteric inhibitory sites.

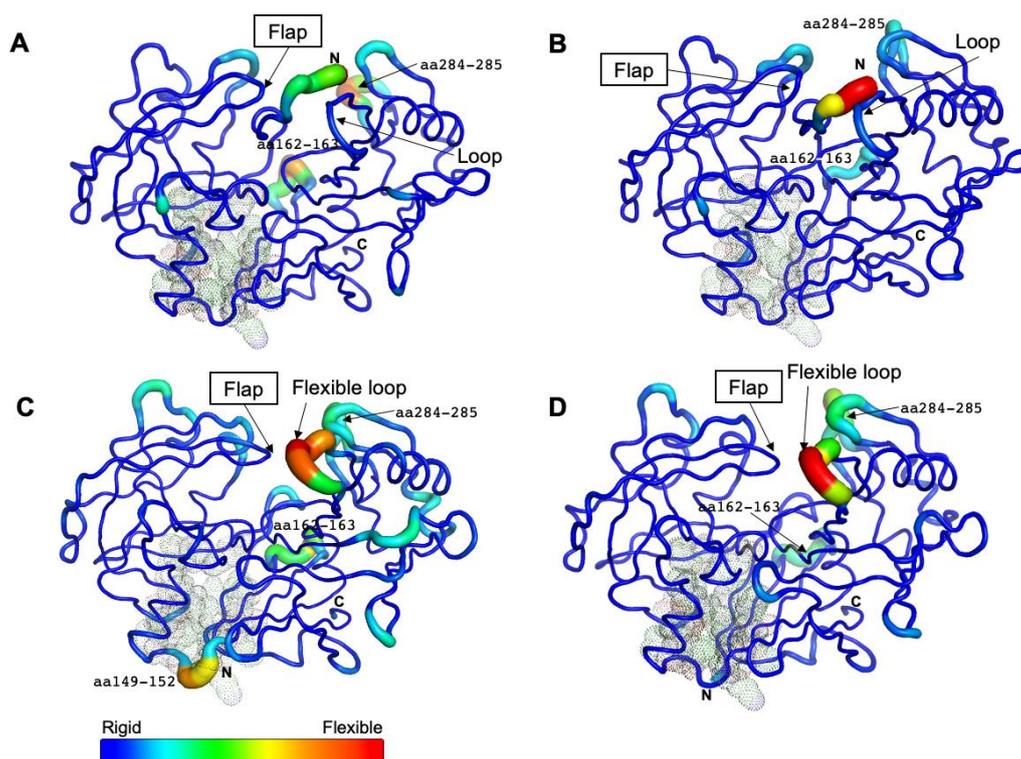


Figure 3. Residual flexibility of the free and bound forms at different pH conditions. The average B-factor-derived mobilities of CsCatD2_{free} at pH 7 (a) and at pH 4 (b), CsCatD2_{bound} at pH 7 (c) and pH 4 (d) were visualized. All residues (at positions 14–16, 32, 97, 149, 170, 171 and 173) of potential allosteric inhibitory site are indicated with black dots (See Section 2.6.).

2.7. Acidic pH enhances the compactness of CsCatD2_{free} to converge to the level of bound forms

To monitor the compactness of CsCatD2_{free} and CsCatD2_{bound} structures, Rg and SASA were analyzed as the indicators of protein structure compactness in all systems. Rg indicates the mass-weighted RMSD between the common center of mass and a collection of atoms [40]. Although the low Rg values of CsCatD2_{bound} were consistent regardless of pH, the Rg values of CsCatD2_{free} tended to decline (Figure 4a). Remarkably, the Rg value of CsCatD2_{free} at pH 4 approached close to the level of CsCatD2_{bound} and eventually converged at approximately 44 ns. These results were further supported by the SASA value, which measures the exposed surface of the entire enzyme that is accessible to solvent molecules [41]. The overall trend of the SASA corresponded to fluctuations in the Rg values for all systems. At approximately 44 ns, the convergence of the SASA in CsCatD2_{free} at pH 4 occurred at the level of CsCatD2_{bound} (Figure 4b). These processes were in line with the abrupt changes in the RMSD values. The low Rg and SASA values throughout the remainder of the simulations were also consistent with the adoption of a new conformation.

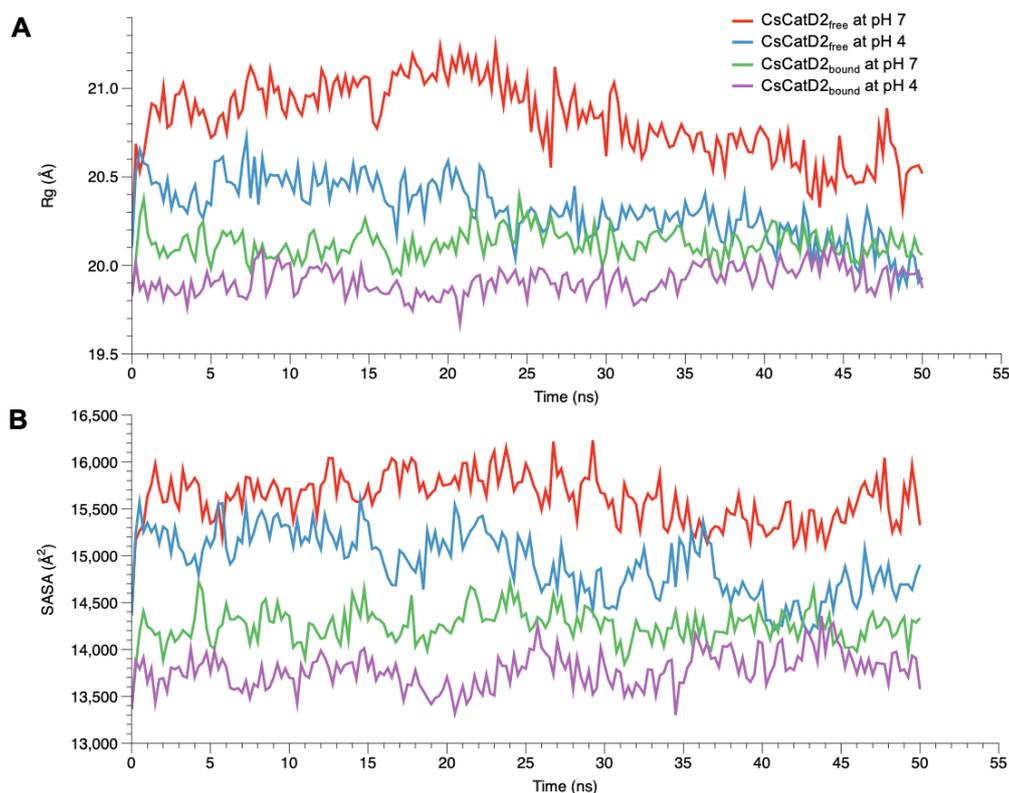


Figure 4. Conformational changes of free and bound CsCatD2 at both pH conditions. Rg (a) and SASA (b) of CsCatD2_{free} and CsCatD2_{bound} were calculated at neutral (pH 7) and acidic (pH 4) condition throughout the trajectory of 50 ns.

2.8. Twisting of the active site pocket by the bound inhibitor and acidic pH

Flap dynamics imply the time-dependent changes in the physical interactions and conformational flexibilities of the flap region and loops. Based on the structural dynamics, it indicates time-dependent conformational changes in proteins [27,28]. Specific structural metrics have been previously suggested for the flap dynamics of CatD [28]. The distance d_1 (Gly₇₉-Met₂₉₇) between the flap tip and hinge residue, TriCα angles, θ_1 (Gly₇₉-Asp₃₃-Met₂₉₇) and θ_2 (Gly₇₉-Asp₂₁₉-Met₂₉₇), which account for the opening and closing of the active site, and dihedral angle ϕ (Gly₇₉-Asp₃₃-Asp₂₁₉-Met₂₉₇) are responsible for the twisting motion (Figure 5a).

The dynamics of the flap structure of CsCatD2 were analyzed. Based on the average structures obtained through a 50-ns MD simulation, each parameter was measured (Figure 5b). The conformational states of CsCatD2 were determined according to the distance d_1 as suggested by Kumalo and Soliman [42]. It was found that the open state was higher than 13 Å, the closed state was lower than 9 Å, and either the semi-open or semi-closed state existed between the two values. In addition to ligand-binding, acidic pH decreased the distance d_1 as follows: from open (13.8 Å) to semi-open (12.6 Å) state in CsCatD2_{free}; and from semi-closed (9.7 Å) to closed (8.7 Å) state in CsCatD2_{bound}. Compared to the initial structures, the acidic pH and bound inhibitor steered the dihedral angle ϕ and reduced the TriCα angles (θ_1 and θ_2) simultaneously. Both the distance d_1 and TriCα angles are likely to decrease. For instance, the

dihedral angle ϕ of CsCatD2_{free} at pH 7 shifted from +2.1° to -4.7° at acidic pH, to -8.0° for the bound ligand, and to -3.8° for both factors. These results collectively provide evidence of the “twisting” of the binding pocket, which can be seen by the asymmetrical movement of the pocket by the bound inhibitor and acidic pH.

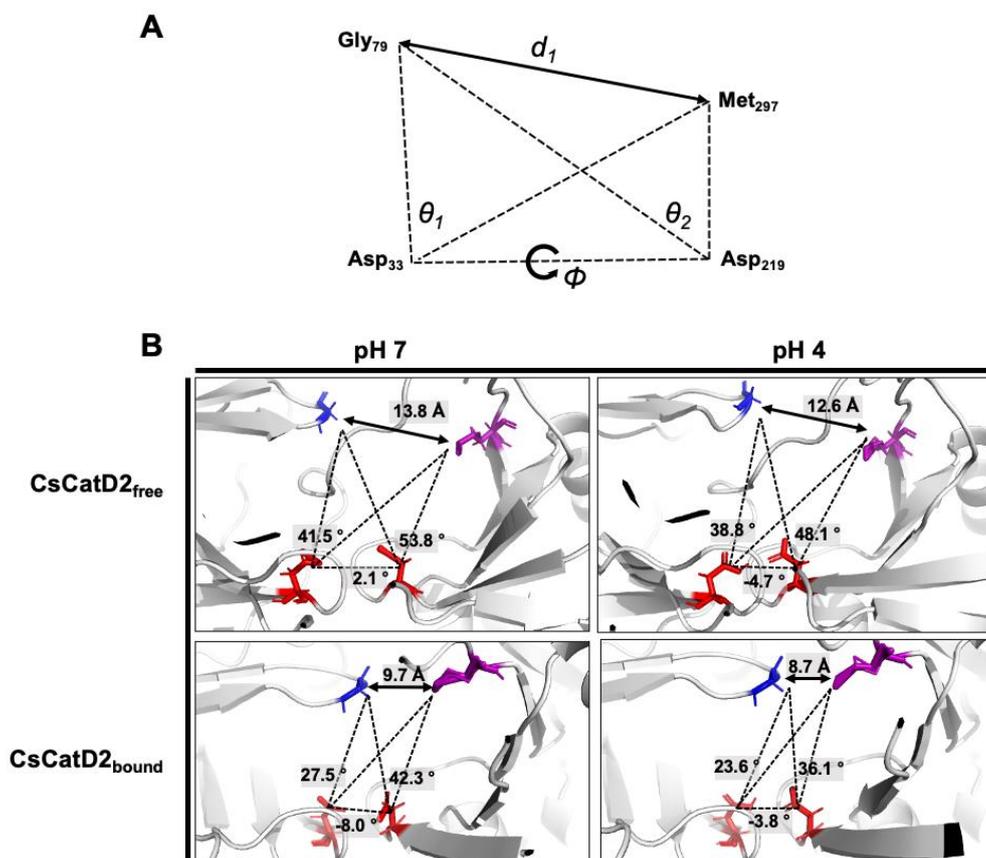


Figure 5. Schematic representation to define the flap-structure dynamics in the free and bound CsCatD2. (a) Parameters used to describe the flap dynamics of the active site and flap region: the distance (d_1) between the flap tip and flexible region hinge residue, the dihedral angle (ϕ) and the TriC α angles (θ_1 and θ_2). (b) Parameter values obtained from the average free and bound CsCatD2 during MD simulations at neutral (pH 7) and acidic (pH 4) conditions. Four residues in active site are shown in sticks.

2.9. Effect of acidic pH on flap dynamics and its correlation with the twisting motion of the active site pocket

Time-series distance d_1 was calculated based on the snapshots obtained at 0, 10, 20, 30, 40, and 50 ns during the 50-ns MD trajectories of all systems (Figure 6). The mean and standard error of distance d_1 were calculated for CsCatD2_{free} at pH 7 (14.85 ± 0.54), CsCatD2_{free} at pH 4 (12.66 ± 0.10), CsCatD2_{bound} at pH 7 (10.52 ± 0.46), and CsCatD2_{bound} at pH 4 (8.69 ± 0.04) (Tables S1 and S2). At the beginning of the simulation, the distance d_1 showed semi-open conformation (11.30 Å at pH 7 and 11.29 Å at pH 4) in CsCatD2_{free}, whereas d_1 indicated a semi-closed conformation (9.20 Å at pH 7 and 9.66 Å at pH 4) in CsCatD2_{bound}. Toward the end of the 50-ns MD simulation, CsCatD2_{free} showed only two open states at pH 4 and then transitioned to a semi-closed state (10.60 Å) at 50 ns. Although the

bound forms showed a similar trend at both pH conditions, CsCatD2_{bound} had a lower minimum d_1 of 7.74 Å at pH 4 than the d_1 (8.74 Å) at pH 7 at 50 ns, revealing three semi-closed conformations or three closed states. Notably, the acidic pH induced distance d_1 of CsCatD2_{free} close to the level of CsCatD2_{bound} at pH 7 around the simulation times (5, 21, 34, and 46 ns) (Figure 7a, Tables S1 and S2).

A similar trend was also observed at the dihedral angle ϕ in CsCatD2_{free} at pH 4 (Figure 7b). CsCatD2_{free} at pH 7 (average ϕ : 2.29 ± 0.43) showed a few small negative values of ϕ on an intermittent basis, whereas CsCatD2_{free} at pH 4 (average ϕ : -4.68 ± 0.43) made the preferable twisted conformation as represented by a majority of negative values of ϕ . Acidic pH induced opposite patterns of ϕ in CsCatD2_{free} during 40–50 ns via conformational shifts to the semi-open state (ranging from 16.24 to 5.59) at pH 7 and semi-closed state (from -18.36 to -6.54) at pH 4. The bound forms had negative values of dihedral angles ϕ , -7.56 ± 0.38 at pH 7 and -3.78 ± 0.19 at pH 4, throughout the entire simulation time. Among the TriC α angles θ_1 and θ_2 (Figure 7c and d), the angle θ_1 was more consistent with distance d_1 , suggesting that these factors may work together to ensure the timely conformational changes of CsCatD2. Therefore, the three parameters (d_1 , ϕ , and θ_1) of flap dynamics were correlated with each other to account for the discrepancies in the open or closed states of the binding cleft.

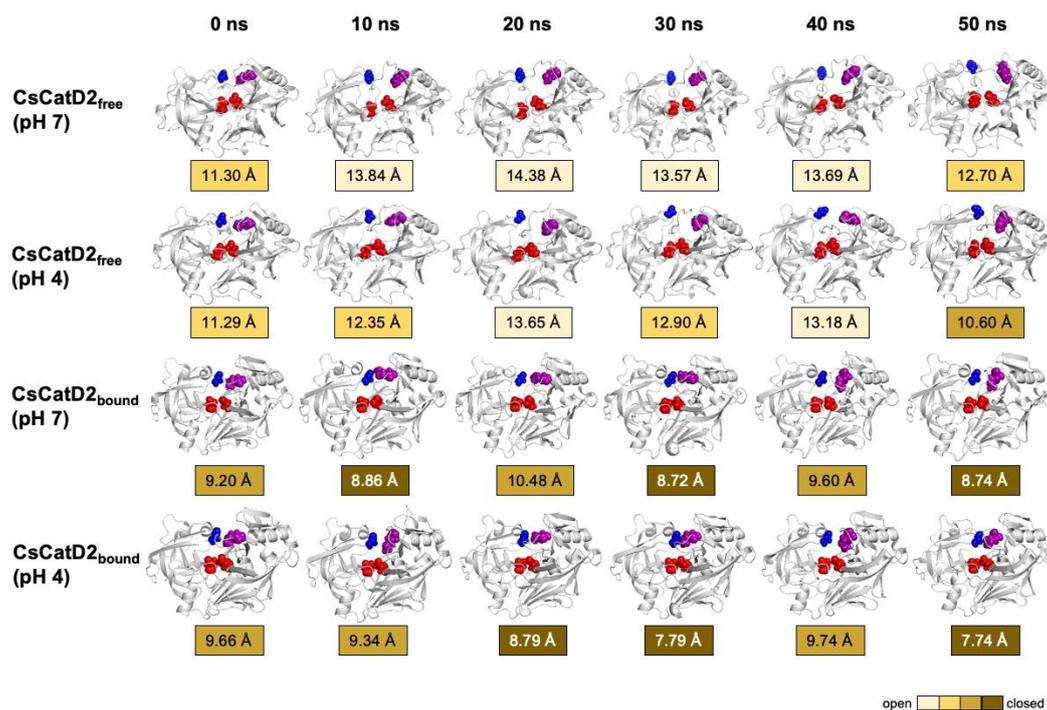


Figure 6. Time-series snapshots of d_1 , the distance between the flap tip and hinge residue, throughout the 50-ns MD simulation. Structural changes of distance d_1 of CsCatD2_{free} and CsCatD2_{bound} were visualized based on the six time points, such as 0, 10, 20, 30, 40 and 50 ns. The conformational states were determined according to the distance d_1 as suggested by Kumalo and Soliman [42], such as open state of $d_1 > 13$ Å, closed state of $d_1 < 9$ Å and semi-conformational state of $9 \leq d_1 \leq 13$. Four residues in active site are shown in spheres, as same color in Figure 3a.

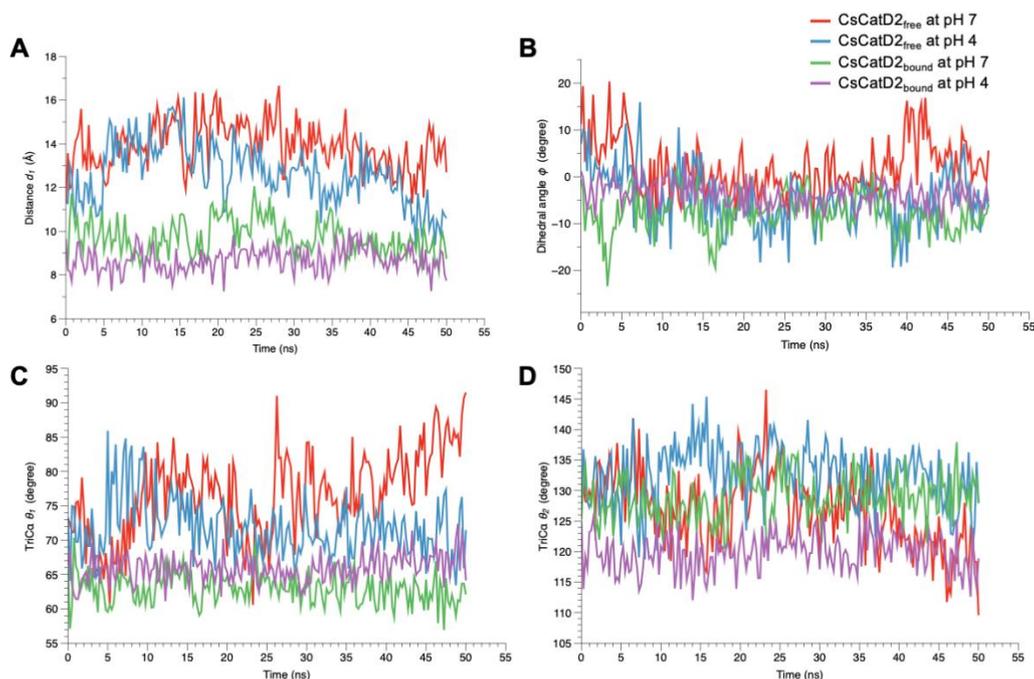


Figure 7. Flap dynamics of free and bound CsCatD2 throughout the trajectory of 50 ns. Flap parameters of CsCatD2_{free} and CsCatD2_{bound} were determined according to both pH conditions, pH 7 and pH 4. (a) distance d_1 , (b) dihedral angle ϕ , (c) TriCa angle θ_1 and (d) TriCa angle θ_2 .

3. Materials and Methods

3.1. Homology modeling and verification

The complete mRNA sequence of CsCatD2 was retrieved from GenBank (Acc. No.: GU433605) [26], and the amino acid sequence was deduced using the Lasergene software (DNASTAR, Madison, Wisconsin, USA). The 3D homology models of CsCatD2_{free} and CsCatD2_{bound} were constructed using Yet Another Scientific Artificial Reality Application (YASARA) Structure v20.12.24 [43] as described by Kang et al. [26]. A hybrid homology model was chosen by combining the best scoring parts of the reliable models for CsCatD2_{free} and CsCatD2_{bound} based on the experimental X-ray characterized structures of free cathepsin D (PDB code: 5n7n) and inhibitor-bound cathepsin D (PDB code: 5n7q) from the tick *I. ricinus* [32]. The qualities of all models were further evaluated using the Ramachandran plot [44], ERRAT [45], and protein structure analysis (ProSA) [46].

3.2. Structural comparison

Structural superpositions were carried out using TM-align [47] with default parameter and “cealign” method script (<https://pymolwiki.org/index.php/Cealign>). The TM-score, calculated by TM-align, indicates the fold similarity between the two structures. All structure graphics were prepared using PyMOL (PyMOL Molecular Graphics System, v.2.4.0, Schrödinger, LLC). The molecular interaction between CsCatD2 and pepstatin was visualized using PoseView [48], as implemented in ProteinPlus [49].

3.3. MD simulation and post-MD analysis

All MD simulations were performed using a pre-installed “md_runfast.mcr” macro file within the YASARA Structure v18.4.24 [50] for both the free and bound forms of CsCatD2. The AMBER14 force field was applied under periodic boundary conditions. The simulation cell was allowed to include 20 Å surrounding the protein and filled with water, as a solvent, at a density of 0.997 g/ml. The initial energy minimization was carried out under relaxed constraints using steepest descent minimization *in vacuo*. To mimic the physiological conditions as previously described [33,34,51], all simulated systems were maintained at pH 4 or pH 7 by adding counter ions to replace the water containing 0.9% sodium chloride (NaCl). Simulations were performed in water at a temperature of 298 K at constant pressure. The cut-off radius for long-range electrostatics was set to 8 Å. The snapshots were saved at 250 picoseconds (ps) intervals for the duration of the 50-ns simulation, as previously suggested [31,35,52]. Subsequently, snapshots were analyzed using the built-in “md_analyze.mcr” and “md_analyzeres.mcr” macro files within the YASARA package for the RMSD, RMSF, Rg, SASA, and flap parameters, such as the distance d_1 (Gly⁷⁹-Met²⁹⁷) between the flap tip and hinge residue, TriC α angles θ_1 (Gly⁷⁹-Asp³³-Met²⁹⁷) and θ_2 (Gly⁷⁹-Asp²¹⁹-Met²⁹⁷) for the conformational changes of the active site, and the dihedral angle ϕ (Gly⁷⁹-Asp³³-Asp²¹⁹-Met²⁹⁷) for the twisting motion of the binding pocket. Plots were generated using DataGraph v4.7 β (Visual Data Tools Inc., Chapel Hill, NC, USA).

4. Conclusions

Parasitic CatD homologues have gained significant attention for being attractive targets for therapeutic drug design as CatD plays a critical role in the degradation of blood to facilitate the survival of the parasite. Here, we report the structure and flap dynamics of CsCatD2 from the blood-feeding parasite *C. sinensis*, which can cause food-borne parasitism and cholangiocarcinoma. The structure-inhibition relationships and validated computational simulations determined in this study will provide important insights into the binding mechanisms of parasitic CatD to potential inhibitors during the pH shift, thereby enabling the generation of a set of new derivatives that possess increased stability and better binding properties and can effectively adjust to the changes in pH.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, Table S1: Fluctuations of distance d_1 of free and bound CsCatD2 at acidic and neutral conditions during 50-ns MD simulations, Table S2: Conformational states of free and bound CsCatD2 at acidic and neutral conditions, Figure S1: Quality verification of tertiary model of CsCatD2free, Figure S2: Quality verification of tertiary model of CsCatD2bound, Figure S3: Variations of the number of hydrogen bonds through the course of 50-ns MD simulation, Figure S4: Superposed results between CsCatD2free and IrCatD1 (PDB code: 5n7n).

Funding: This research was supported by the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (grant no. NRF2016R1C1B1009348 and NRF2018R1C1B6005581) (<http://www.nrf.re.kr>).

Acknowledgments: We gratefully thank Dr. Elmar Krieger at the YASARA Biosciences GmbH (<http://www.yasara.org>; Vienna, Austria) for technical help about post-molecular dynamics analysis and also for his fruitful advice about data interpretation.

Author Contributions: Conception and design, B.-K.N., W.G.Y.; methodology and data analysis, J.-M.K., W.G.Y.; data curation and interpretation, W.G.Y., J.-M.K. and H.G.L.; writing, J.-M.K. and W.G.Y.; review and editing, B.-K.N. and W.G.Y.; supervision, B.-K.N. and W.G.Y. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data generated or analyzed during this study are included in this published article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lun, Z.R.; Gasser, R.B.; Lai, D.H.; Li, A.X.; Zhu, X.Q.; Yu, X.B.; Fang, Y.Y. Clonorchiasis: A key foodborne zoonosis in china. *Lancet Infect Dis* **2005**, *5*, 31-41.
2. Na, B.K.; Pak, J.H.; Hong, S.J. Clonorchis sinensis and clonorchiasis. *Acta Trop* **2020**, *203*, 105309.
3. Choi, B.I.; Han, J.K.; Hong, S.T.; Lee, K.H. Clonorchiasis and cholangiocarcinoma: Etiologic relationship and imaging diagnosis. *Clin Microbiol Rev* **2004**, *17*, 540-552, table of contents.
4. Vennervald, B.J.; Polman, K. Helminths and malignancy. *Parasite Immunol* **2009**, *31*, 686-696.
5. Bouvard, V.; Baan, R.; Straif, K.; Grosse, Y.; Secretan, B.; El Ghissassi, F.; Benbrahim-Tallaa, L.; Guha, N.; Freeman, C.; Galichet, L., *et al.* A review of human carcinogens-part b: Biological agents. *Lancet Oncol* **2009**, *10*, 321-322.
6. Delcroix, M.; Sajid, M.; Caffrey, C.R.; Lim, K.C.; Dvorak, J.; Hsieh, I.; Bahgat, M.; Dissous, C.; McKerrow, J.H. A multienzyme network functions in intestinal protein digestion by a platyhelminth parasite. *J Biol Chem* **2006**, *281*, 39316-39329.
7. Sajid, M.; McKerrow, J.H. Cysteine proteases of parasitic organisms. *Mol Biochem Parasitol* **2002**, *120*, 1-21.
8. Na, B.K.; Kim, S.H.; Lee, E.G.; Kim, T.S.; Bae, Y.A.; Kang, I.; Yu, J.R.; Sohn, W.M.; Cho, S.Y.; Kong, Y. Critical roles for excretory-secretory cysteine proteases during tissue invasion of *paragonimus westermani* newly excysted metacercariae. *Cell Microbiol* **2006**, *8*, 1034-1046.
9. Na, B.K.; Kang, J.M.; Sohn, W.M. Cscf-6, a novel cathepsin f-like cysteine protease for nutrient uptake of *clonorchis sinensis*. *Int J Parasitol* **2008**, *38*, 493-502.
10. Robinson, M.W.; Dalton, J.P.; Donnelly, S. Helminth pathogen cathepsin proteases: It's a family affair. *Trends Biochem Sci* **2008**, *33*, 601-608.
11. Pearson, M.S.; Ranjit, N.; Loukas, A. Blunting the knife: Development of vaccines targeting digestive proteases of blood-feeding helminth parasites. *Biol Chem* **2010**, *391*, 901-911.
12. Cwiklinski, K.; Donnelly, S.; Drysdale, O.; Jewhurst, H.; Smith, D.; Verissimo, C.D.; Pritsch, I.C.; O'Neill, S.; Dalton, J.P.; Robinson, M.W. The cathepsin-like cysteine peptidases of trematodes of the genus *fasciola*. *Adv Parasit* **2019**, *104*, 113-164.
13. Horn, M.; Fajtova, P.; Rojo Arreola, L.; Ulrychova, L.; Bartosova-Sojkova, P.; Franta, Z.; Protasio, A.V.; Opavsky, D.; Vondrasek, J.; McKerrow, J.H., *et al.* Trypsin- and chymotrypsin-like serine proteases in *schistosoma mansoni*-- 'the undiscovered country'. *PLoS Negl Trop Dis* **2014**, *8*, e2766.
14. Kang, J.M.; Ju, H.L.; Ju, J.W.; Sohn, W.M.; Kim, T.S.; Bahk, Y.Y.; Hong, S.J.; Na, B.K. Comparative biochemical and functional properties of two leucine aminopeptidases of *clonorchis sinensis*. *Mol Biochem Parasitol* **2012**, *182*, 17-26.

15. Williamson, A.L.; Brindley, P.J.; Abbenante, G.; Datu, B.J.; Prociv, P.; Berry, C.; Girdwood, K.; Pritchard, D.I.; Fairlie, D.P.; Hotez, P.J., *et al.* Hookworm aspartic protease, na-apr-2, cleaves human hemoglobin and serum proteins in a host-specific fashion. *J Infect Dis* **2003**, *187*, 484-494.
16. Caffrey, C.R.; Placha, L.; Barinka, C.; Hradilek, M.; Dostal, J.; Sajid, M.; McKerrow, J.H.; Majer, P.; Konvalinka, J.; Vondrasek, J. Homology modeling and sar analysis of *schistosoma japonicum* cathepsin d (sjcd) with statin inhibitors identify a unique active site steric barrier with potential for the design of specific inhibitors. *Biol Chem* **2005**, *386*, 339-349.
17. Sojka, D.; Hartmann, D.; Bartosova-Sojkova, P.; Dvorak, J. Parasite cathepsin d-like peptidases and their relevance as therapeutic targets. *Trends Parasitol* **2016**, *32*, 708-723.
18. Dougall, A.M.; Skwarczynski, M.; Khoshnejad, M.; Chandrudu, S.; Daly, N.L.; Toth, I.; Loukas, A. Lipid core peptide targeting the cathepsin d hemoglobinase of *schistosoma mansoni* as a component of a schistosomiasis vaccine. *Hum Vaccin Immunother* **2014**, *10*, 399-409.
19. Hartman, A.M.; Mondal, M.; Radeva, N.; Klebe, G.; Hirsch, A.K. Structure-based optimization of inhibitors of the aspartic protease endothiasepsin. *Int J Mol Sci* **2015**, *16*, 19184-19194.
20. Brinkworth, R.I.; Prociv, P.; Loukas, A.; Brindley, P.J. Hemoglobin-degrading, aspartic proteases of blood-feeding parasites: Substrate specificity revealed by homology models. *J Biol Chem* **2001**, *276*, 38844-38851.
21. Li, J.; Chi, Z.; Liu, Z.; Yue, L.; Peng, Y.; Wang, L. Cloning and characterization of a novel aspartic protease gene from marine-derived *metschnikowia reukauffii* and its expression in *e. Coli*. *Appl Biochem Biotechnol* **2009**, *159*, 119-132.
22. Williamson, A.L.; Brindley, P.J.; Loukas, A. Hookworm cathepsin d aspartic proteases: Contributing roles in the host-specific degradation of serum proteins and skin macromolecules. *Parasitology* **2003**, *126*, 179-185.
23. Suttiprapa, S.; Mulvenna, J.; Huong, N.T.; Pearson, M.S.; Brindley, P.J.; Laha, T.; Wongkham, S.; Kaewkes, S.; Sripa, B.; Loukas, A. Ov-apr-1, an aspartic protease from the carcinogenic liver fluke, *opisthorchis viverrini*: Functional expression, immunolocalization and subsite specificity. *Int J Biochem Cell Biol* **2009**, *41*, 1148-1156.
24. Morales, M.E.; Rinaldi, G.; Gobert, G.N.; Kines, K.J.; Tort, J.F.; Brindley, P.J. Rna interference of *schistosoma mansoni* cathepsin d, the apical enzyme of the hemoglobin proteolysis cascade. *Mol Biochem Parasitol* **2008**, *157*, 160-168.
25. Yoo, W.G.; Kim, D.W.; Ju, J.W.; Cho, P.Y.; Kim, T.I.; Cho, S.H.; Choi, S.H.; Park, H.S.; Kim, T.S.; Hong, S.J. Developmental transcriptomic features of the carcinogenic liver fluke, *clonorchis sinensis*. *PLoS Negl Trop Dis* **2011**, *5*, e1208.
26. Kang, J.M.; Yoo, W.G.; Le, H.G.; Thai, T.L.; Hong, S.J.; Sohn, W.M.; Na, B.K. Partial characterization of two cathepsin d family aspartic peptidases of *clonorchis sinensis*. *Korean J Parasitol* **2019**, *57*, 671-680.
27. Boehr, D.D.; D'Amico, R.N.; O'Rourke, K.F. Engineered control of enzyme structural dynamics and function. *Protein Sci* **2018**, *27*, 825-838.
28. McGillewie, L.; Ramesh, M.; Soliman, M.E. Sequence, structural analysis and metrics to define the unique dynamic features of the flap regions among aspartic proteases. *Protein J* **2017**, *36*, 385-396.
29. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The protein data bank. *Nucleic Acids Res* **2000**, *28*, 235-242.
30. Marcinišzyn, J., Jr.; Hartsuck, J.A.; Tang, J. Mode of inhibition of acid proteases by pepstatin. *J Biol Chem* **1976**, *251*, 7088-7094.
31. Arodola, O.A.; Soliman, M.E. Molecular dynamics simulations of ligand-induced flap conformational changes in cathepsin-d-a comparative study. *J Cell Biochem* **2016**, *117*, 2643-2657.
32. Hanova, I.; Brynda, J.; Houstecka, R.; Alam, N.; Sojka, D.; Kopacek, P.; Maresova, L.; Vondrasek, J.; Horn, M.; Schueler-Furman, O., *et al.* Novel structural mechanism of allosteric regulation of

- aspartic peptidases via an evolutionarily conserved exosite. *Cell Chem Biol* **2018**, *25*, 318-329 e314.
33. Mancilla-Olea, M.I.; Ortega-Lopez, J.; Figueroa-Angulo, E.E.; Avila-Gonzalez, L.; Cardenas-Guerra, R.E.; Miranda-Ozuna, J.F.T.; Gonzalez-Robles, A.; Hernandez-Garcia, M.S.; Sanchez-Ayala, L.; Arroyo, R. *Trichomonas vaginalis* cathepsin d-like aspartic proteinase (tv-catd) is positively regulated by glucose and degrades human hemoglobin. *Int J Biochem Cell Biol* **2018**, *97*, 1-15.
 34. Minarowska, A.; Gacko, M.; Karwowska, A.; Minarowski, L. Human cathepsin d. *Folia Histochem Cytobiol* **2008**, *46*, 23-38.
 35. Karubiu, W.; Bhakat, S.; McGillevie, L.; Soliman, M.E. Flap dynamics of plasmepsin proteases: Insight into proposed parameters and molecular dynamics. *Mol Biosyst* **2015**, *11*, 1061-1066.
 36. Berjanskii, M.V.; Wishart, D.S. A simple method to predict protein flexibility using secondary chemical shifts. *J Am Chem Soc* **2005**, *127*, 14970-14971.
 37. Dostal, J.; Pecina, A.; Hruskova-Heidingsfeldova, O.; Mareckova, L.; Pichova, I.; Rezacova, P.; Lepsik, M.; Brynda, J. Atomic resolution crystal structure of sapp2p, a secreted aspartic protease from *candida parapsilosis*. *Acta Crystallogr D Biol Crystallogr* **2015**, *71*, 2494-2504.
 38. Masa, M.; Maresova, L.; Vondrasek, J.; Horn, M.; Jezek, J.; Mares, M. Cathepsin d propeptide: Mechanism and regulation of its interaction with the catalytic core. *Biochemistry* **2006**, *45*, 15474-15482.
 39. Lee, A.Y.; Gulnik, S.V.; Erickson, J.W. Conformational switching in an aspartic proteinase. *Nat Struct Biol* **1998**, *5*, 866-871.
 40. Lobanov, M.; Bogatyreva, N.S.; Galzitskaia, O.V. [radius of gyration is indicator of compactness of protein structure]. *Mol Biol (Mosk)* **2008**, *42*, 701-706.
 41. Richmond, T.J. Solvent accessible surface area and excluded volume in proteins. Analytical equations for overlapping spheres and implications for the hydrophobic effect. *J Mol Biol* **1984**, *178*, 63-89.
 42. Kumalo, H.M.; Soliman, M.E. A comparative molecular dynamics study on bace1 and bace2 flap flexibility. *J Recept Signal Transduct Res* **2016**, *36*, 505-514.
 43. Krieger, E.; Nabuurs, S.B.; Vriend, G. Homology modeling. *Methods Biochem Anal* **2003**, *44*, 509-523.
 44. Lovell, S.C.; Davis, I.W.; Arendall, W.B., 3rd; de Bakker, P.I.; Word, J.M.; Prisant, M.G.; Richardson, J.S.; Richardson, D.C. Structure validation by calpha geometry: Phi,psi and cbeta deviation. *Proteins* **2003**, *50*, 437-450.
 45. Colovos, C.; Yeates, T.O. Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Sci* **1993**, *2*, 1511-1519.
 46. Wiederstein, M.; Sippl, M.J. Prosa-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* **2007**, *35*, W407-410.
 47. Zhang, Y.; Skolnick, J. Tm-align: A protein structure alignment algorithm based on the tm-score. *Nucleic Acids Res* **2005**, *33*, 2302-2309.
 48. Stierand, K.; Rarey, M. Drawing the pdb: Protein-ligand complexes in two dimensions. *ACS Med Chem Lett* **2010**, *1*, 540-545.
 49. Fahrrolfes, R.; Bietz, S.; Flachsenberg, F.; Meyder, A.; Nittinger, E.; Otto, T.; Volkamer, A.; Rarey, M. Proteinsplus: A web portal for structure analysis of macromolecules. *Nucleic Acids Res* **2017**, *45*, W337-W343.
 50. Krieger, E.; Vriend, G. New ways to boost molecular dynamics simulations. *J Comput Chem* **2015**, *36*, 996-1007.
 51. Mukherjee, B.; Tessaro, F.; Vahokoski, J.; Kursula, I.; Marq, J.B.; Scapozza, L.; Soldati-Favre, D. Modeling and resistant alleles explain the selectivity of antimalarial compound 49c towards apicomplexan aspartyl proteases. *EMBO J* **2018**, *37*.

52. McGillewie, L.; Soliman, M.E. Flap flexibility amongst plasmepsins i, ii, iii, iv, and v: Sequence, structural, and molecular dynamics analyses. *Proteins* **2015**, *83*, 1693-1705.