European Review for Medical and Pharmacological Sciences 2021; 25: 3162-3172

In silico analysis of *Naegleria fowleri* cathepsin B paralogs: important drug targets

M. AURONGZEB[1,3], S.A. HAQ[2], Y. RASHID[2], S.H. AHMED NAQVI[3], S.I. HUSSAIN[4], A.U. REHMAN[4], I. KALEEM[6], S. BASHIR[7]

[1]Jamil ur Rehman Center for Genome Research, Dr. Punjwani Center for Molecular Medicine and Drug Research, International Center for Chemical and Biological Sciences, University of Karachi, Karachi, Pakistan
[2]Department of Biochemistry, University of Karachi, Karachi, Pakistan
[3]Institute of Biotechnology and Genetic Engineering, University of Sindh, Jamshoro, Pakistan
[4]Department of Zoology, Federal Urdu University of Arts, Science & Technology, Gulshan-e-Iqbal, Karachi, Pakistan
[5]Department of Biochemistry, Abdul Wali khan University Mardan, Mardan, Khyber Pakhtunkhwa, Pakistan
[6]Department of Bioinformatics and Bioscience, COMSATS University (CUI), Islamabad, Pakistan
[7]Neuroscience Center, King Fahad Specialist Hospital, Dammam, Saudi Arabia

Muhammad Aurongzeb and Sadia Mohammed contributed equally

Abstract. – *Naegleria fowleri* is a deadly human pathogen that causes primary amoebic meningoencephalitis (PAM). In this study, in silico investigations of two important *N. fowleri* cathepsin B paralogs, i.e., copies of genes resulting from a gene duplication event, were carried out using comparative modeling and molecular dynamics (MD) simulations. Comparative models of both paralogs showed significant architectural similarity with their template, i.e., rat cathepsin B. However, in *N. fowleri* cathepsin B (UniProt ID: X5D761) and putative cathepsin B (UniProt ID: M1HE19) enzymes, eleven and fifteen residues in the occluding loop regions were deleted, respectively, suggesting that these enzymes have a short occluding loop. Thus, it is concluded that *N. fowleri* cathepsin B and putative cathepsin B enzymes lack exopeptidase activity but possess enhanced endopeptidase activity and an affinity for macromolecular inhibitors. MD simulations further confirmed that prosegments (macromolecular inhibitors) bond more tightly with both enzymes than with wild-type cathepsin B. Additionally, a mutation was identified at an important N-glycosylation site; this mutation is believed to affect cathepsin B targeting inside the cell and make cathepsin B available in the extracellular environment. Due to this important N-glycosylation site mutation, these enzymes are secreted in the extracellular environment via an alternative, still unknown, posttranslational processing strategy. The present study is the first to predict the three-dimensional folds of *N. fowleri* cathepsin B paralogous enzymes, including a detailed description of the active site architecture and information about propeptide binding mode. This information can contribute to the discovery of novel and selective treatments that are effective against *N. fowleri*.

Key Words:
Primary amoebic meningoencephalitis, *Naegleria fowleri*, Cysteine proteases, Homology modeling, All-atom molecular dynamic simulations.

Abbreviations

PAM: Primary amoebic meningoencephalitis; VMD: Visual Molecular Dynamics; MD simulation: Molecular Dynamics simulation; CatB-M1: Cathepsin B Model-1; CatB-M2: Cathepsin B Model-2; PBL: Prosegment-Binding Loop; ECM: Extracellular Matrix.

Introduction

Primary amoebic meningoencephalitis (PAM) is a destructive brain disease caused by the thermotolerant protozoa *Naegleria fowleri*. This free-living amoeba is found in recreational waters, mainly ponds, lakes, rivers, hot springs, and swimming pools. It has also been found in water reservoirs and water storage tanks in houses[1,2]. Due to the high mortality rate, PAM is a serious medical concern worldwide[3]. In Pakistan, Karachi is the most affected city by PAM, and by October 2019, 146 cases have been reported.

Corresponding Author: Yasmeen Rashid, MD; e-mail: yasmeen.rashid@uok.edu.pk

N. fowleri multiplies rapidly in high temperatures, usually in summer. Infection occurs when amoebae adhere to the nasal mucosa, travel along the olfactory nerve, and cross the cribriform plate to enter into the brain, where they cause PAM[4-6]. Certain cytolytic molecules, including hydrolases, neuraminidases, phospholipases, and phospholipolytic enzymes, are responsible for the pathogenicity of *N. fowleri*. These molecules damage the cells and nerves of the host, often leading to death[7]. Current treatments for PAM rarely ensure survival. The primary drug of choice for treating PAM is amphotericin B; however, this treatment requires a high dosage, which can lead to renal toxicity[8]. The identification and structural characterization of factors that influence the pathogenicity of *N. fowleri* are crucial and can support the development of more effective therapeutic interventions.

The important role of cysteine proteinases in protozoan pathogenesis, including tissue invasion and the intracellular survival of several pathogenic parasites, has been reported by many studies[9-12]. *N. fowleri* genome contains three paralogs of cysteine cathepsin B proteases. Paralogous genes (or paralogs) are a type of homologous genes resulting from a gene duplication event within the same organism and may have the same or different function after the event[13]. *N. fowleri* cathepsin B proteases may be involved in the proteolytic degradation of collagen, fibronectin, immunoglobulins, albumin, and hemoglobin. This suggests that these proteases play an important role in the pathogenesis of *N. fowleri* by facilitating host tissue attachment, evading host immunity, and easing nutrient uptake[14]. Cathepsin B proteases have been found in *N. fowleri* lysate and excretory-secretory protein fractions, demonstrating their crucial role in the pathogenesis of *N. fowleri*[15-17].

The present study aims to elucidate three-dimensional structures of *N. fowleri* cathepsins B enzymes, delineate their active site architectures, and observe enzyme propeptide binding modes via in silico investigations.

Materials and Methods

Sequence Analyses and Template Search

Amino acid sequences of *N. fowleri* cysteine proteases (UniProt ID: X5D761; UniProt ID: M1HE19; UniProt ID: X5D911) were obtained from the UniProt database. For the identification of structural homologs, PSI-BLAST was used against Protein Data Bank (PDB). Depending on the high percentage similarities and the low number of gaps, full-length structural homologs were selected as templates for X5D761 and M1HE19 cysteine proteases, whereas cathepsin B-like protease (UniProt ID: X5D911) could not be aligned well with any of the structures present in PDB. Multiple sequence alignment was carried out using homologous sequences from different species using ClustalX software[18] to determine the consensus patterns in the cathepsin B cysteine protease family. The pattern was used to optimize the pairwise sequence alignment of target and template.

Secondary Structure Prediction

PSIPRED program was used for secondary structure prediction of X5D761 and M1HE19. The information obtained was used to build structure-based pairwise sequence alignment between the target and template protein sequences.

Comparative Modeling

Comparative models of both *N. fowleri* cathepsin B paralogs (UniProt ID: X5D761; UniProt ID: M1HE19) were built using the same template, i.e., the crystal structure of rat procathepsin B (PDB ID: 1MIR). All steps of homology modeling and refinement were carried out using the program MODELLER v9.20. Evaluation of homology models was done using the stand-alone programs PROCHECK[19] and ProSA[20]. Structural superposition was performed using options available in the SUPERPOSE script file of the MODELLER[21].

Protein Modeling and Structural Analysis

Analysis of three-dimensional protein structures is a more mature field of study than sequence analysis. DS Visualizer and Visual Molecular Dynamics (VMD)[22] software programs were used for the molecular analysis of these enzymes.

All-Atom Molecular Dynamics Simulation

Molecular Dynamics (MD) simulations can help shed light on protein-protein and protein-ligand interactions at the molecular level[23]. In the present study, all-atom MD simulations and reasonable analysis procedures were conducted using the Amber software package, version 18[24]. Initially, two systems were prepared, exemplified as Model-1 and Model-2. The LEaP module was used to add hydrogen atoms to

both systems. Counterions (Na^+ and Cl^-) were added to maintain system neutrality. All systems were solvated in a truncated octahedral box of TIP3P water model with 10 Å buffer. The Particle Mesh Ewald (PME) method[25] was used to treat long-range electrostatic interactions, and *ff99IDPs* force field was used for all simulations[26,27]. All the bonds involving hydrogen atoms were constrained using the SHAKE algorithm[28]. The PMEMD of CUDA version was used to accelerate all the MD simulations[29]. The steepest descent method was used to minimize the solvated systems for 20,000 steps, then 400 ps heating, and 200 ps equilibration in the NVT ensemble. The production ran under the NPT ensemble at 298 K with a time step of 2 ps in Berendsen thermostat and barostat. The CPPTRAJ package was used to analyze the trajectories in Amber 18. Finally, all analyses were conducted using the CPPTRAJ module implemented in Amber, i.e., root-mean-square deviation (RMSD); root-mean-square fluctuation (RMSF); the CA distance between the bound peptide and CatB protein; the radius of gyradius (Rg); hydrogen bonding population among peptides, respectively.

Results

Based on the importance of cysteine cathepsins to *N. fowleri* pathogenesis and PAM progression, we aimed to perform protein sequence analyses and identify three-dimensional architectures of *N. fowleri* cathepsin B using standard bioinformatics tools. UniProt database search provided three *N. fowleri* cathepsin B protein sequences, including cathepsin B (UniProt ID: X5D761), putative cathepsin B (UniProt ID: M1HE19), and cathepsin B-like protein (Uniprot ID: X5D911). The crystal structure of rat procathepsin B (PDB ID: 1MIR) was found to be a common optimal template for both *N. fowleri* cathepsin B and putative cathepsin B, with sequence similarities of 54% and 51%, respectively. However, no suitable template was found for the *N. fowleri* cathepsin B-like protein (UniProt ID: X5D911) in a PDB. Structure-based pairwise sequence alignments of cathepsin B and putative cathepsin B, along with their common template of rat procathepsin B (PDB ID: 1MIR), were optimized with consideration of protein secondary structures and conservation patterns across the cathepsin B protease family.

Homology Modeling of N. fowleri Cathepsin B and Putative Cathepsin B Proteases

Homology models of *N. fowleri* cathepsin B (CatB-M1) and putative cathepsin B (CatB-M2) were constructed using optimized pairwise sequence alignments (Figures 1A and 1B). Both models were found to be of good quality using different evaluation procedures. Overall, the three-dimensional folds of *N. fowleri* CatB-M1 (Figures 2A and 2B) and CatB-M2 (Figures 2C and 2D) were comparable to the crystal structure of rat procathepsin B. Similar to the rat procathepsin B fold, both the *N. fowleri* cathepsin B models were found to be comprised of two distinct domains: a five-stranded beta-sheet domain and an alpha-helical domain with three alpha-helices and a small beta-sheet structure.

A proregion lacking any globular structure tends to wrap around cathepsin B while interacting with three key areas, including a prosegment-binding loop (PBL), a substrate-binding cleft, and an occluding loop crevice. The prosegment and mature cathepsin B interactions were analyzed in both *N. fowleri* cathepsin B models and in the rat procathepsin B by selecting an area of 5 Å around their respective prosegments (Figure 3). The center of PBL that includes Y183 and Y188 forms a depression on the surface of cathepsin B. This prosegment interacts with PBL *via* a hairpin element that directs the β-strand and α-helix of the proregion on the other side of PBL, thereby positioning the W24 side chain in the center of the depression. The side chain of W24 forms an H-bond with the carbonyl group of K189 residue of PBL. All the interactions between *N. fowleri* prosegment and PBL were well conserved in both models suggesting that the interactions of PBL with prosegment are generally similar in both *N. fowleri* cathepsin B models and in the crystal structure the template, rat procathepsin B. Residues of the substrate-binding cleft, including G27, C29, G73, G74, W30, Y75, H110, F174, V176, L181, M196, G197, G198, H199, A200, and W221, bind with the prosegment (Figure 3). All residues were strictly conserved in both *N. fowleri* cathepsin B models except for H110, which is missing due to the deletion of eleven and fifteen residues in the occluding loop region of *N. fowleri* CatB-M1 and CatB-M2, respectively (Figure 4).

Active Site Analyses

The active site triad, i.e., Cys29-His199-Asp219, was conserved and found to be located

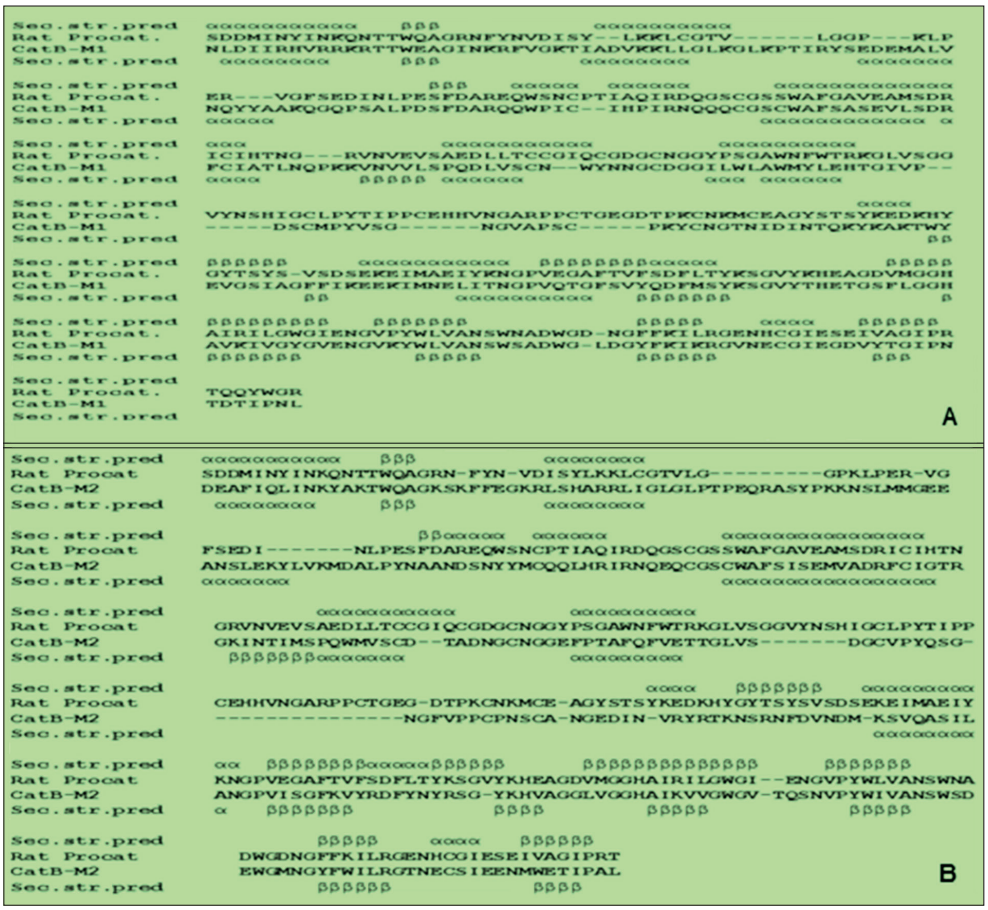

Figure 1. Optimized pairwise sequence alignments of *N. fowleri* cathepsin B (**A**) and putative cathepsin B (**B**) with rat procathepsin B (PDB ID: 1MIR).

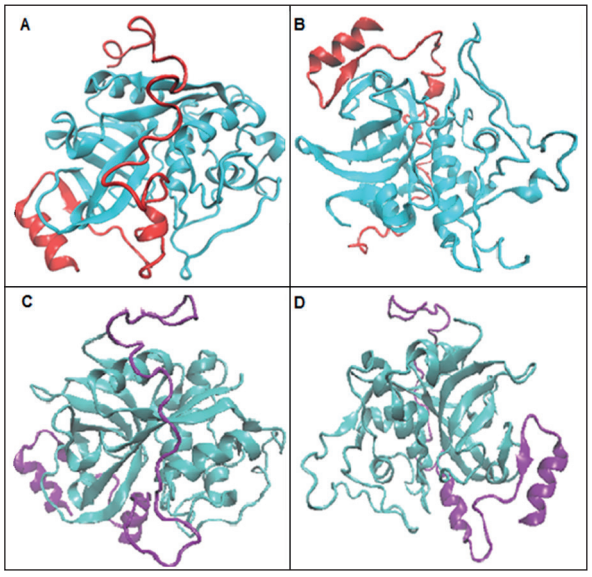

Figure 2. Ribbon representation of three-dimensional models of *N. fowleri* CatB-M1 (**A,** front view; **B,** back view) and CatB-M2 (**C,** front view; **D,** back view). The proregion is shown in *red* and *purple* in CatB-M1 and CatB-M2, respectively.

at the domain interface, like rat procathepsin B. An area of 5 Å around the catalytic triad was found to consist of 36 amino acid residues in both models (Figure 5). Twenty-seven residues, including the catalytic triad, were strictly conserved in *N. fowleri* CatB-M1 compared to the crystal structure of rat procathepsin B. Six of these residues showed conservative substitutions: I200→V, A173→G, T175→S, R200→K, N222→S, and F230→Y. Only three nonconservative substitutions were observed: G23→S, E171→Q, and G172→T. In CatB-M2, 28 residues, including the catalytic triad, were strictly conserved compared to rat procathepsin B. Five active site residues showed conservative substitutions: A34→I, A173→G, R20→K, N222→S, and F177→Y. Only three nonconservative substitutions were observed: T175→K, E171→I, and G172→S.

Active site analysis of both models revealed a similar abundance of active site residues in the *N. fowleri* CatB-M1 and CatB-M2 models

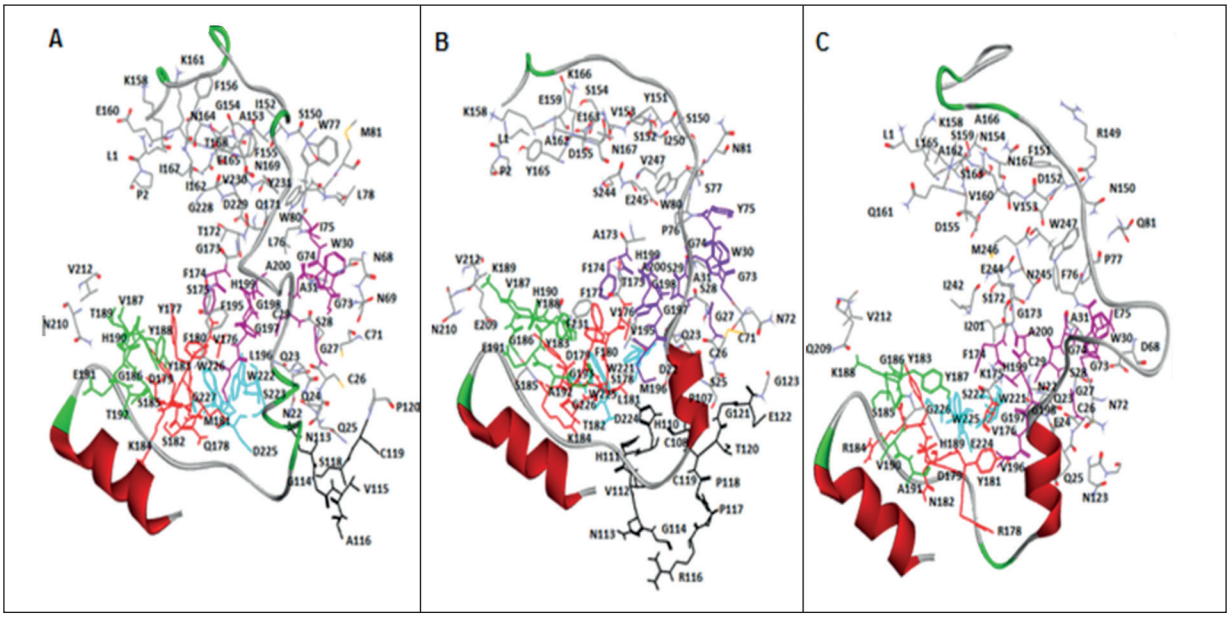

Figure 3. Interactions of proregion in CatB-M1 **(A),** 1MIR **(B)**, and CatB-M2 **(C)** with the three key areas of the mature cathepsin, including prosegment-binding loop (*green color*), substrate-binding cleft (*purple color*), and occluding loop crevice (*black color*). The residues common among the three regions are shown in *red*, whereas the residues of the base part are shown in *cyan*.

compared to rat procathepsin B. Here, we believe that the overall architectures of both *N. fowleri* cathepsin B models, including active sites related to endopeptidase activity, are quite similar to that of rat procathepsin B, suggesting a similar mechanism of action. The missing residues, which are important for exopeptidase activity, especially H110 and H111 that were missing in both models, suggest that *N. fowleri* cathepsin B lacks exopeptidase activity.

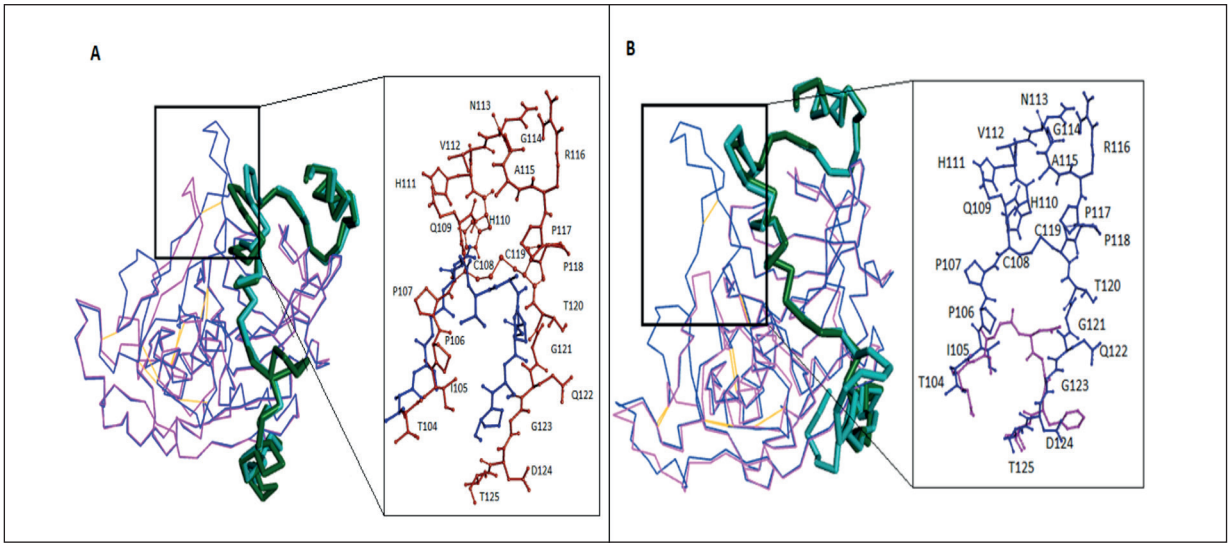

Figure 4. Superposition of rat procathepsin B (1MIR) with CatB-M1 **(A)** and CatB-M2 **(B)** of *N. fowleri*. The proregions of the template and the model are shown in c-alpha stick presentation in *dark green* and *cyan*, respectively. Cathepsin B parts of the template and the models are depicted in c-alpha wire presentation in *blue* and *pink*, respectively. The residues of the occluding loop of the template and model are shown in *brown* and *blue* in CatB-M1 **(A)**, respectively, and *blue* and *pink* in CatB-M2 **(B)**, respectively.

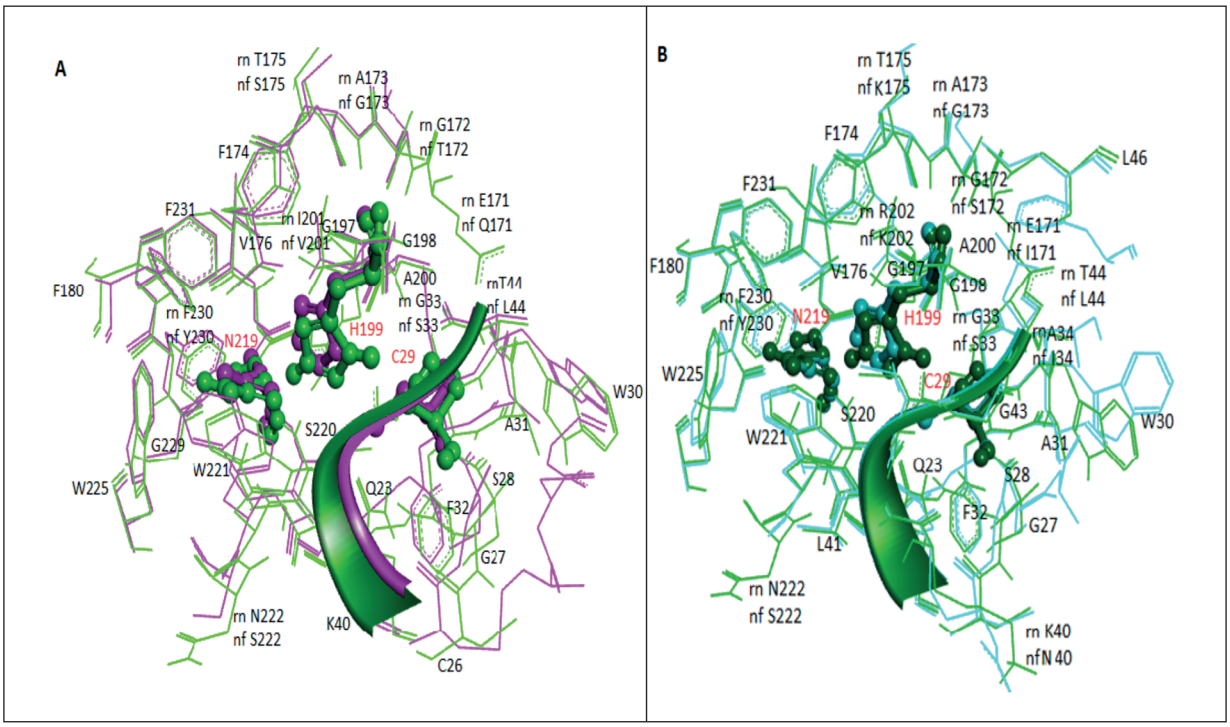

Figure 5. Analysis of the active site residues of *N. fowleri* cathepsin B **(A)** and putative cathepsin B **(B)** homology models after superposition. 1MIR is shown in *light* **(A)** and *dark green* (B), whereas CatB-M1 active site residues are shown in *pink* **(A)** and CatB-M2 active site residues are shown in *cyan*. The catalytic triad (C29, H199, and N219) is shown in a ball-and-stick presentation.

Localization of N. fowleri Cathepsin B Proteases

Multiple sequence alignment of human, rat, and *N. fowleri* cathepsin B proteases showed strict conservation of the catalytic triad (Figure 6). Normally, it is expected that cathepsin B-containing N-glycosylation site, i.e., N38 (human cathepsin B numbering[30]), is transported to the lysosome via the vesicular MPR pathway[31]. However, the nonconservative substitutions N38→R and N38→A were observed at the N-glycosylation site in *N. fowleri* cathepsin B and putative cathepsin B, respectively (Figure 6). These substitutions suggest the extralysosomal distribution of cathepsin B proteases via an unknown alternate mechanism in *N. fowleri*. Cathepsin B proteases are known to be differentially expressed and secreted in the trophozoite stage of *N. fowleri*[14], which suggests that they play a role in its pathogenesis.

All-Atom Molecular Dynamics Simulation

The three-dimensional folds of *N. fowleri* CatB-M1 and CatB-M2, comparable to the crystal structure of rat procathepsin B**,** were MD simulated in an explicit water environment for 100 ns. The deviation of backbone atoms was examined using RMSD to ensure the stability of the simulation process. CatB-M1 showed a gradual increase starting 1 Å into the MD simulation time, whereas CatB-M2 initially increased drastically in deviation reaching 3.9 Å at 18 ns, and then showed a consistent behavior and oscillated till 100 ns. A smaller RMSD curve indicates high stability and vice versa (Figure 7A). When comparing 1MIR-M1 with 1MIR-M2, the RMSD results revealed the inconsistent behavior of the M1 complex throughout MD simulation, whereas the 1M1R-M2 showed a consistent behavior that favors the stability and reliability of the complex. RMSF were analyzed to determine the flexibility of individual residues in the CatB-M1 and CatB-M2 systems. This analysis showed that the CatB-M1 complex exhibited local fluctuations, especially in residues 100-200, with residue fluctuations of up to 2.5 Å, 2.9 Å, and 4.3 Å (Figure 7B). These residue fluctuations were reduced in the CatB-M2 complex, clearly indicating the stability and tightness of the prosegment in this complex. The RMS analysis

Figure 6. Multiple sequence alignment of human, rat, *N. fowleri* cathepsin B (X5D761), and putative cathepsin B (M1HE19) proteases. The prosegment residues are shown in red, showing the important glycosylation site. The occluding loop region is shown in a red box, whereas the catalytic triad is shown in *blue*.

indicates that the prosegments in both models exhibit consistent, steady binding, particularly marked in the CatB-M2 complex, based on the RMSD/F results.

The alpha carbon distance between the peptide and mature enzyme in both models was also analyzed and compared. The alpha carbon distance for the CatB-M2 complex remains steady, i.e., around 1.5 Å. This steadiness indicates that the propeptide in CatB-M2 binds tightly and possesses a higher affinity for binding than the CatB-M1 complex (Figure 7C). A hydrogen population analysis further revealed that the propeptide showed a high affinity to bind CatB-M1

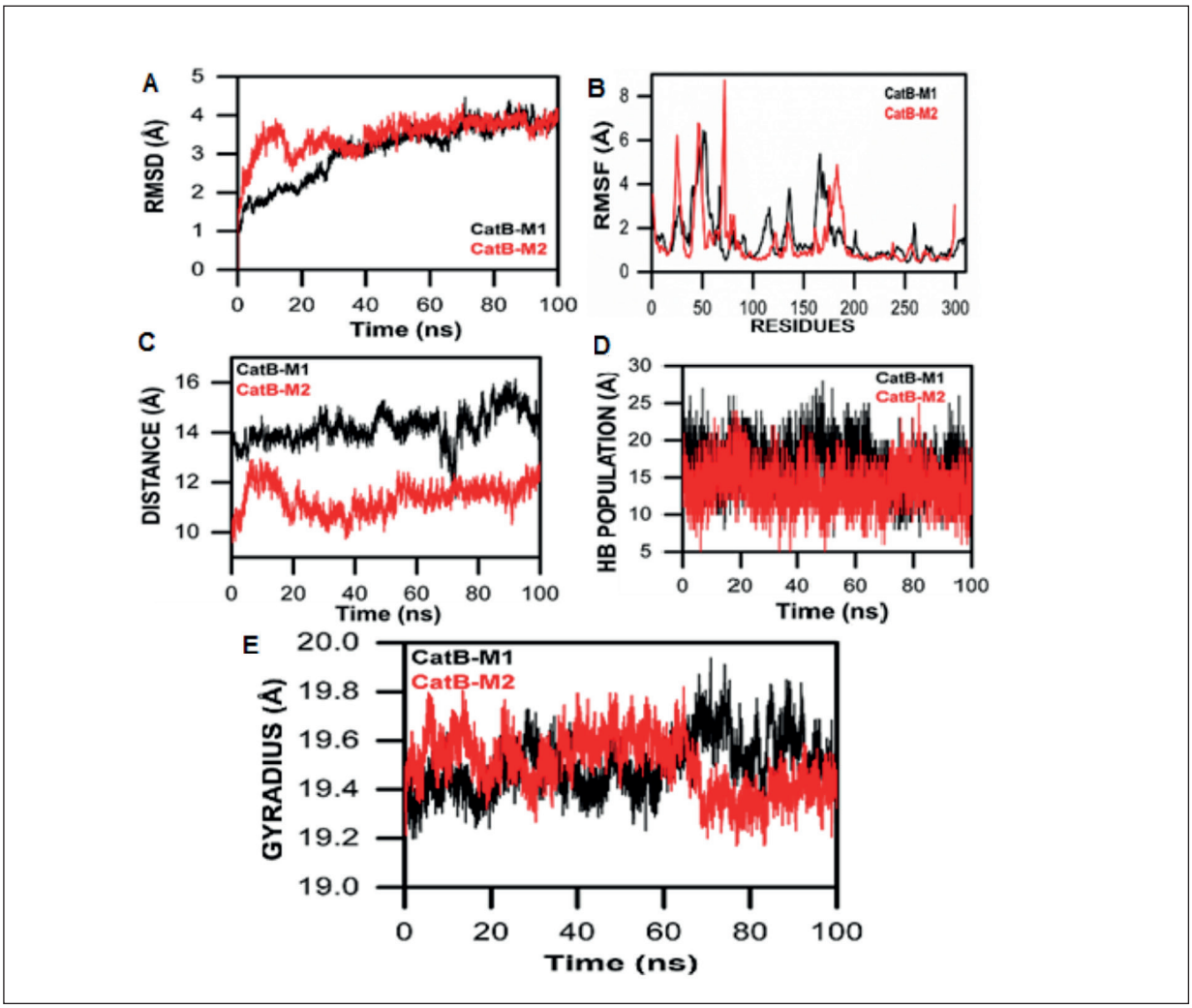

Figure 7. All-atom MD simulation studies of *N. fowleri* cathepsin B (X5D761) and putative cathepsin B (M1HE19) proteases.

enzyme and consistent behavior throughout the MD simulation. In contrast, the CatB-M2 exhibited dramatic behaviors, including high HB interactions between the enzyme and the propeptide during the 1-10 ns and 40-70 ns MD simulations (Figure 7D). The consequences of the backbone fluctuations and deviations in both CatB-M1 and CatB-M2 indicated the need for further exploration of the compactness of the overall conformation. We also analyzed the radius of gyration (Rg), which indicated that the CatB-M2 was initially more compact than the CatB-M1 for 60 ns to 100 ns, and the compactness of CatB-M2 decreased. The compactness of the CatB-M1 increased and decreased throughout the MD simulations (Figure 7E). The analysis revealed that the CatB-M2 complex showed consistent behaviors in overall conformation, probably due to the tight binding of the propeptide to the enzyme.

Discussion

N. fowleri is the only etiological agent of PAM in human beings. Although some information about the genome sequence and important pathogenic molecules of *N. fowleri* is available, there is still no selective potential treatment for PAM[32]. *N. fowleri* invades the CNS via nasal mucosa by crossing the cribriform plate, where it causes the characteristic meningoencephalitis. Cytolytic molecules are responsible for the pathogenicity of *N. fowleri*; these molecules are known to cause

nerve and cell destruction in the host, often resulting in death[7]. Cysteine proteases, especially cysteine cathepsins secreted by *N. fowleri*, can destroy host tissues and host immunity[1]. In addition to their primary function of cellular protein degradation and catabolism, cysteine cathepsins are involved in numerous important physiological processes, including ECM turnover, immune invasion, digestion, and parasite invasion and escape, and certain pathologies, such as rheumatoid arthritis, atherosclerosis, leishmaniasis, amebiasis, and malaria[31,33,34]. Since it is a lysosomal enzyme, cathepsin B is normally believed to be transported to the lysosomes via mannose-6-phosphate receptor pathway[35]. However, this distribution of cathepsin B shifts towards the cell periphery in tumor cells suggesting the extracellular secretion of cathepsin B proteases[36-40].

Cathepsin B proteases are known to be differentially expressed and secreted in the trophozoite stage of *N. fowleri*, which suggests that they play a significant role in its pathogenesis. *N. fowleri* cathepsin B enzymes play a pivotal role in the progression of PAM[41]. Therefore, the present study involved in silico protein sequence analyses and sought to determine the three-dimensional structures of *N. fowleri* cathepsin B and putative cathepsin B proteases. Detailed analyses revealed that the overall folds of *N. fowleri* cathepsin B and putative cathepsin B are quite similar to that of rat procathepsin B, suggesting a similar mechanism of action. Three segments of mature cathepsin B, a part of an occluding loop (108-122), a part of the PBL, and a loop (221-226) at the base, interact together to form a crevice. The latter two regions were conserved in both models. However, some residues of the occluding loop involved in forming the crevice were not found in both models. Notably, the H110 and H111 residues (human cathepsin B numbering[30]) are crucial parts of the occluding loop and are responsible for the exopeptidase activity in higher eukaryotes. The cathepsin B protease of *Giardia* lacks the occluding loop, suggesting that the loop was either deleted or inserted later during the course of evolution[42].

Accordingly, eleven and fifteen residues are deleted in the occluding loop regions of *N. fowleri* cathepsin B and putative cathepsin B, respectively. These deletions, along with the missing H110 and H111 residues, suggest that the occluding loop is short for cathepsin B and even shorter in putative cathepsin B (Figure 4). Deletion of the occluding loop eliminates exopeptidase activity but increases endopeptidase activity. It also increases the ability of macromolecular inhibitors to bind to the active site[43]. MD simulation studies confirmed tight binding of the prosegment, its natural inhibitor, with both mature cathepsin B models. This suggests that the prosegment plays a protective role against uncontrolled proteolytic degradation within the pathogen. Additionally, nonconservative substitutions at important N-glycosylation sites in *N. fowleri* cathepsin B and putative cathepsin B suggest their extracellular secretion and role in tissue damage to the host. This article is the first to present the tertiary architectures of *N. fowleri* cathepsin B proteases, along with information about their active sites and the mechanics of cathepsin B propeptide. We believe that this information will support the development of novel lead compounds against the deadly protozoa and could lead to more effective treatments for PAM.

Conclusions

Cysteine proteinases, such as cathepsin B, play a vital role in protozoan pathogenesis, including tissue invasion and intracellular pathogenic survival. *In silico* studies of two paralogous *N. fowleri* cathepsin B enzymes showed a large deletion of eleven and fifteen residues, respectively, compared with rat cathepsin B (the template), in their occluding loop regions, suggesting little or no exopeptidase activity but increased endopeptidase activity and increased affinity between the active site and the inhibitory propeptide. MD simulation studies of *N. fowleri* cathepsin B enzymes confirmed the tight binding of prosegments with their active sites. Additionally, mutation of typical N-glycosylation site indicates extracellular secretion of *N. fowleri* cathepsin B enzymes via an alternative yet unknown posttranslational processing strategy. Here, we suggest that both of these *N. fowleri* cathepsin B enzymes have high endopeptidase activity. When these highly active cathepsin B enzymes are inside the cell, they are kept inactive via tight binding of their propeptides. Once outside the cell, the enzymes achieve their fully active form by cleaving off their propeptide segments and become disastrous for the host tissue.

Conflict of Interest

The Authors declare that they have no conflict of interests.

Ethical Approval

Ethical approval is not applicable as the present study does not involve any animal or human participant.

Informed Consent

Informed Consent is not applicable as the present study does not involve any human participant.

References

1) Aldape K, Huizinga H, Bouvier J, McKerrow J. Naegleria fowleri: characterization of a secreted histolytic cysteine protease. Exp Parasitol 1994; 78: 230-241.
2) Ghanchi NK, Khan E, Khan A, Muhammad W, Malik FR, Zafar A. Naegleria fowleri meningoencephalitis associated with public water supply, Pakistan, 2014. Emerg Infect Dis 2016; 22: 1835-1837.
3) Tung MC, Hsu BM, Tao CW, Tung MC, Hsu BM, Tao CW, Lin WC, Tsai HF, Ji DD, Shen SM, Chen JS, Shih FC, Huang YL. Identification and significance of Naegleria fowleri isolated from the hot spring which related to the first primary amebic meningoencephalitis (PAM) patient in Taiwan. Int J Parasitol 2013; 43: 691-696.
4) Visvesvara GS, Moura H, Schuster FL. Pathogenic and opportunistic free-living amoebae: acanthamoeba spp., Balamuthia mandrillaris, Naegleria fowleri, and Sappinia diploidea. FEMS Immunol Med Microbiol 2007; 50: 1-26.
5) Jamerson M, da Rocha-Azevedo B, Cabral GA, Marciano-Cabral F. Pathogenic Naegleria fowleri and non-pathogenic Naegleria lovaniensis exhibit differential adhesion to, and invasion of, extracellular matrix proteins. Microbiology (Reading, England) 2012; 158: 791-803.
6) Baig AM. Primary amoebic meningoencephalitis: neurochemotaxis and neurotropic preferences of naegleria fowleri. ACS Chem Neurosci 2016; 7: 1026-1029.
7) Grace E, Asbill S, Virga K. Naegleria fowleri: pathogenesis, diagnosis, and treatment options. Antimicrob Agents Chemother 2015; 59: 6677-6681.
8) Stevens AR, Shulman ST, Lansen TA, Cichon MJ, Willaert E. Primary amoebic meningoencephalitis: a report of two cases and antibiotic and immunologic studies. J Infect Dis 1981; 143: 193-199.
9) Karrer KM, Peiffer SL, DiTomas ME. Two distinct gene subfamilies within the family of cysteine protease genes. Proc Natl Acad Sci USA 1993; 90: 3063-3067.
10) McKerrow JH, Caffrey C, Kelly B, Loke P, Sajid M. Proteases in parasitic diseases. Annu Rev Pathol 2006; 1: 497-536.
11) McKerrow JH. The diverse roles of cysteine proteases in parasites and their suitability as drug targets. PLoS Negl Trop Dis 2018; 12: e0005639.
12) Keene WE, Petitt MG, Allen S, McKerrow JH. The major neutral proteinase of Entamoeba histolytica. J Exp Med 1986; 163: 536-549.
13) Moreira D, López-García P. Paralogous Gene. In: Gargaud M, Amils R, Quintanilla JC, Cleaves HJ, Irvine WM, Pinti D, Viso M, eds. Encyclopedia of Astrobiology. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011; 1215.
14) Lee J, Kim JH, Sohn HJ, Yang HJ, Na BK, Chwae YJ, Park S, Kim K, Shin HJ. Novel cathepsin B and cathepsin B-like cysteine protease of Naegleria fowleri excretory-secretory proteins and their biochemical properties. Parasitol Res 2014; 113: 2765-2776.
15) Shin HJ, Cho MS, Jung SY, Kim HI, Park S, Kim HJ, Im KI. Molecular cloning and characterization of a gene encoding a 13.1 kDa antigenic protein of Naegleria fowleri. J Eukaryot Microbiol 2001; 48: 713-717.
16) Kim JH, Yang AH, Sohn HJ, Kim D, Song KJ, Shin HJ. Immunodominant antigens in Naegleria fowleri excretory--secretory proteins were potential pathogenic factors. Parasitol Res 2009; 105: 1675-1681.
17) Kim JH, Kim D, Shin HJ. Contact-independent cell death of human microglial cells due to pathogenic Naegleria fowleri trophozoites. Korean J Parasitol 2008; 46: 217-221.
18) Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 1997; 25: 4876-4882.
19) Laskowski RA, Moss DS, Thornton JM. Main-chain bond lengths and bond angles in protein structures. J Mol Biol 1993; 231: 1049-1067.
20) Sippl MJ. Recognition of errors in three-dimensional structures of proteins. Proteins 1993; 17: 355-362.
21) Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 1993; 234: 779-815.
22) Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph 1996; 14: 33-38.
23) Rehman AU, Khan MT, Liu H, Wadood A, Malik SI, Chen HF. Exploring the pyrazinamide drug resistance mechanism of clinical mutants T370P and W403G in ribosomal protein S1 of mycobacterium tuberculosis. J Chem Inf Model 2019; 59: 1584-97.
24) Avagliano D, Sánchez-Murcia PA, González L. Directional and regioselective hole injection of spiropyran photoswitches intercalated into A/T-duplex DNA. Phys Chem Chem Phys 2019; 21: 17971-17977.
25) Darden T, York D, Pedersen L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. J Chem Phys 1993; 98: 10089.

26) Wang W, Ye W, Jiang C, Luo R, Chen HF. New force field on modeling intrinsically disordered proteins. Chem Biol Drug Des 2014; 84: 253-269.

27) Ye W, Ji D, Wang W, Luo R, Chen HF. Test and evaluation of ff99IDPs force field for intrinsically disordered proteins. J Chem Inf Model 2015; 55: 1021-1029.

28) Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 1977; 23: 327-341.

29) Salomon-Ferrer R, Götz AW, Poole D, Le Grand S, Walker RC. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. J Chem Theory Comput 2013; 9: 3878-3888.

30) Musil D, Zucic D, Turk D, Engh RA, Mayr I, Huber R, Popovic T, Turk V, Towatari T, Katunuma N. The refined 2.15 A X-ray crystal structure of human liver cathepsin B: the structural basis for its specificity. EMBO J 1991; 10: 2321-2330.

31) Moin K, Demchik L, Mai J, Duessing J, Peters C, Sloane BF. Observing proteases in living cells. Adv Exp Med Biol 2000; 477: 391-401.

32) Liechti N, Schürch N, Bruggmann R, Wittwer M. Nanopore sequencing improves the draft genome of the human pathogenic amoeba Naegleria fowleri. Sci Rep 2019; 9: 16040.

33) Verma S, Dixit R, Pandey KC. Cysteine proteases: modes of activation and future prospects as pharmacological targets. Front Pharmacol 2016; 7: 107.

34) Vasiljeva O, Reinheckel T, Peters C, Turk D, Turk V, Turk B. Emerging roles of cysteine cathepsins in disease and their potential as drug targets. Curr Pharm Des 2007; 13: 387-403.

35) Kornfeld S. Trafficking of lysosomal enzymes. FASEB J 1987; 1: 462-468.

36) Calkins CC, Sameni M, Koblinski J, Sloane BF, Moin K. Differential localization of cysteine protease inhibitors and a target cysteine protease, cathepsin B, by immuno-confocal microscopy. J Histochem Cytochem 1998; 46: 745-751.

37) Sameni M, Elliott E, Ziegler G, Fortgens PH, Dennison C, Sloane BF. Cathepsin B and D are localized at the surface of human breast cancer cells. Pathol Oncol Res 1995; 1: 43-53.

38) Sloane B, Moin K, Lah T. Regulation of lysosomal endopeptidases in malignant neoplasia. In: Aspects of the Biochemistry and Molecular Biology of Tumors. New York: Academic Press, 1994.

39) Erdel M, Trefz G, Spiess E, Habermaas S, Spring H, Lah T, Ebert W. Localization of cathepsin B in two human lung cancer cell lines. J Histochem Cytochem 1990; 38: 1313-1321.

40) Krepela E, Bartek J, Skalkova D, Vicar J, Rasnick D, Taylor-Papadimitriou J, Hallowes RC. Cytochemical and biochemical evidence of cathepsin B in malignant, transformed and normal breast epithelial cells. J Cell Sci 1987; 87: 145-154.

41) Zyserman I, Mondal D, Sarabia F, McKerrow JH, Roush WR, Debnath A. Identification of cysteine protease inhibitors as new drug leads against Naegleria fowleri. Exp Parasitol 2018; 188: 36-41.

42) Ward W, Alvarado L, Rawlings ND, Engel JC, Franklin C, McKerrow JH. A primitive enzyme for a primitive cell: the protease required for excystation of Giardia. Cell 1997; 89: 437-444.

43) Illy C, Quraishi O, Wang J, Purisima E, Vernet T, Mort JS. Role of the occluding loop in cathepsin B activity. J Biol Chem 1997; 272: 1197-1202.