Schistosoma Mansonii aspartic protease expression and refolding trials

By

Bruno de Lima Damasceno

Submitted to the graduate degree program in Molecular Biosciences and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Masters

_____

Chairperson: Dr. Mark Richter

_____

Dr. Fusao Takusagawa

_____

Dr. Krzysztof Kuczera

Date Defended: 12/09/2011

The Dissertation Committee for Bruno de Lima Damasceno

certifies that this is the approved version of the following dissertation:

Schistosoma Mansonii aspartic protease expression and refolding trials

_____

Chairperson: Dr. Mark Richter

Date approved: 12/09/2011

Abstract

Schistosomiasis is a parasitic disease that causes considerable socio-economic losses in affected areas due to loss of productive capacity of affected individuals and high rates of morbidity and mortality.

Therapeutic controls for this parasitic disease have shown some drawbacks with resistance emergence to praziquantel, the drug of choice for treatment, being reported in recent years. Thus, new chemotherapeutic targets are been investigated, aiming to develop drugs that are more effective, with lower cost and fewer adverse reactions. Among these new targets, aspartic proteases are among the most promising, since their involvement in other diseases such as Alzheimer's and diabetes have been proven and a chemotherapeutic arsenal for AIDS treatment has been developed based on the structure and function of the HIV1 aspartic protease. The rational design of new drugs requires knowledge of the structures of target proteins.

In this study we have examined the activation of pro-enzyme, recombinant forms of two putative cathepsin D-like aspartic proteases from the helminth *Schistosoma mansonii* (SmCDs). Extensive folding trials were undertaken in attempts to determine the potential for activation of the proteolytic activities of the enzymes. A recombinant protein disulfide isomerase from the same organism was also prepared for use in protease folding trials. Preliminary evidence suggests that an activated form of one of the two proteases (SmCD2) may be obtained by introducing a solubilizing maltose-binding tag on the N-terminal end resulting in soluble expression of the enzyme. This result has raised the prospect of developing an *in vitro* screening tool to identify potential lead compounds for new drug development. Although PDI samples were purified and active in rearrange disulfide bridges of insulin, no evidence of assistance in refolding SmCDs were observed. Since PDI itself may be considered as a drug target, crystallization attempts in order to obtain its structures were done. Both crystallization of PDI and its influence in SmCDs refolding must be further be investigated in different conditions.

Acknowledgements

I would like to dedicate this work to my family and friends, both in USA and Brazil, which provided me generosity and understanding in my moments of anxiety during this work.

Next I would like to express my deepest gratitude to my formal and informal advisors, Dr. Mark Richter Dr. Floriano Paes Silva Junior for their trust, guidance, support and friendship.

I would like to acknowledge Phillip Gao and Na Zhang, Kim Colvert, Diane Wyatt and Alex Moise for sharing their knowledge, time and collaborations.

Special thanks to my Lab colleagues, Gale Hasslan, Lindsey, Daxin, Shyan, Caroline, Denise; graduate colleagues Vi and Amanda; Marcos, Pedro, Rafael, Lucia, Luciana e several friends from the brazilian community of Lawrence; Manmeet, Mauricio, Suzana, Daniel, Luis, Nadia, Martha, Cristian, Veronica, Akshaia and several good friends from the international student community of Lawrence. Thank you all for all the excellent moments together.

Table of contents

List of figures and tables

Chapter 1- Schistosomiasis

1A- Schistosomiasis and drug discovery.

According to the World Health Organization (WHO-UN), parasitic diseases such as malaria and schistosomiasis are a major focus of public health concern in developing countries. Most are classified as neglected diseases because of the lower interest of pharmaceutical industries in development of treatments. (Fenwick *et al.*, 2003; Renslo, 2006).

Adaptation of transmission cycles of neglected diseases to new geographical areas **(Figure 1**), climate changes and human migration, limitations of treatment availability (economic and / or high toxicity and / or reports of resistance) are factors that drive the search for new forms of therapies, especially considering newly acquired structure/function relations of vital proteins for parasite metabolism, such as proteases, chaperones and outer membrane channels. (Reithinger *et al.*, 2006; Renslo, 2006).

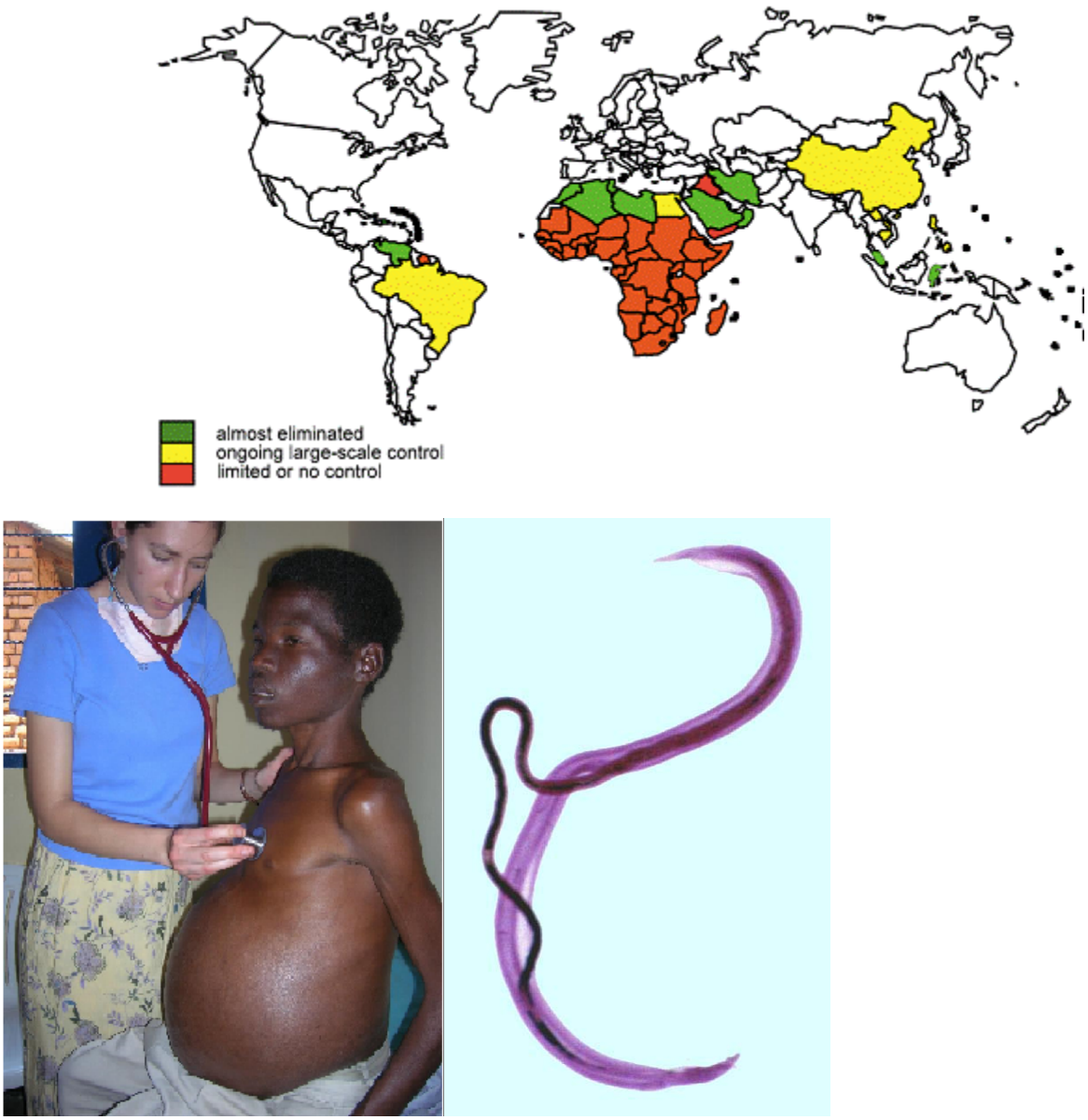


Figure 1: Upper panel: Distribution of schistosomiasis (Engels *et al*,2002); Lower panel, male and female parasite and patient with schistosomiasis (www.shedfoundation.org)

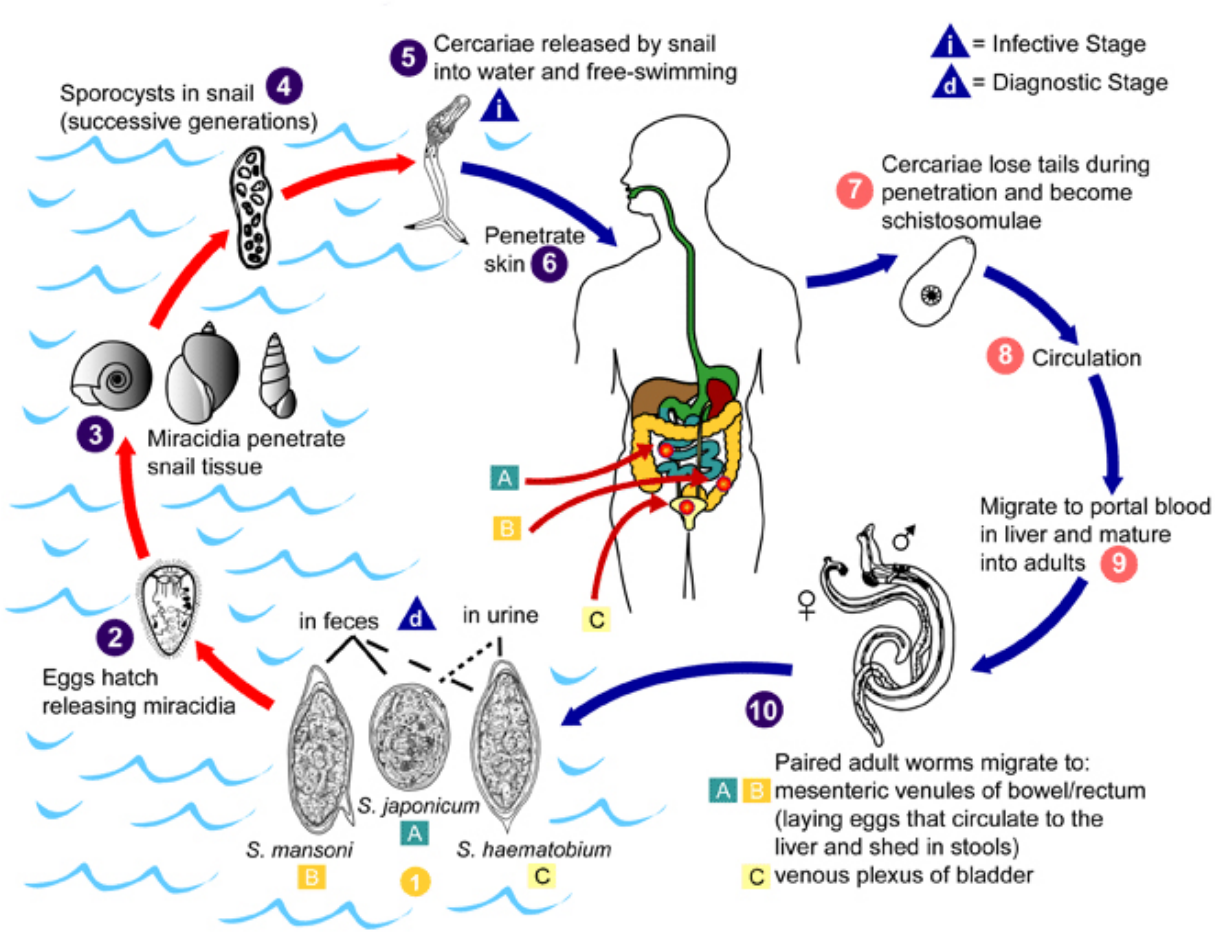


Figure 2: The Schistosoma life cycle: eggs are eliminated with feces or urine (1). Under optimal conditions the eggs hatch and release miracidia (2), which swim and penetrate snail intermediate hosts (3). The stages in the snail include 2 generations of sporocysts (4) and the production of cercariae (5). Upon release from the snail, the infective cercariae swim, penetrate the skin of the human host (6), and become schistosomulae (7). The schistosomulae migrate through several tissues to their residence site (8, 9). Adult worms in humans reside in the mesenteric venules in various locations (10). *S. japonicum* is more frequently found in the small intestine (A), and *S.*

mansoni occurs more often in the large intestine (B). However, both species can occupy either location, and they are capable of moving between sites. Various animals, such as dogs, cats, rodents, pigs, horses and goats, serve as hosts (http://www.niaid.nih.gov).

Schistosomiasis is one of the major neglected tropical diseases, affecting 200 million people in Latin America, North Africa, the Middle East, South China and Indochina. It is also known as barriga-dágua in Brazil or bilharzioses (tribute to Thomas Bilharz, who identified the parasites for the first time in 1851 in Egypt) and is caused by parasites of the genus Schistossoma (trematode). The main etiologic agents in humans are: S. *mansoni* (Africa and Latin America), S. *haematobium* (Africa and Middle East), S. *japonicum* (China and Philippines) and S. *mekongi* (Laos and Cambodia).

Control of Schistosomiasis is achieved through therapeutic administration of praziquantel and oxamniquin, drugs that are effective and present low toxicity and cost. However, schistosomiasis has been expanding in many regions and this seems to be associated with high rates of re-infection and repeated interventions that have limited the effectiveness of pharmacotherapy and allowed development of resistance to therapeutics. Because of the low number of alternatives to the drugs available and lack of a vaccine (Capron, 2005; Abdulla, 2007), the construction of a therapeutic arsenal against schistosomiasis is a major goal of public health authorities and scientists (Fenwick *et al*., 2003; Sibley & Hunt, 2003)

The search for new drugs has focused on obtaining new molecular targets through a better knowledge of the biochemistry and metabolism of the parasite.

Substantial effort has been made to identify the complete genome, proteome and glycome of the Schistosoma genus by two independent groups, one Chinese and one Brazilian (Curwen et al., 2004; Verjovski-Almeida *et al*., 2004; McManus *et al*., 2004; Wilson *et al.*, 2006).

One of the most intriguing discoveries about the infective process of parasites is their use of proteases for degradation of host proteins in order to obtain amino acids for their metabolism. Not only do those proteases have cellular localizations and structural properties different from those existing in vertebrates, but also the degradation cascade that occurs is quite different: In vertebrates, it involves several proteases in a coordinated manner (especially serine proteases), but in invertebrates, degradation events appear to be non coordinated and functionally redundant, with cysteine and aspartic proteases cleaving the same substrate (Delcroix *et al*., 2006). This creates the prospect of developing libraries of inhibitors targeting protease functions that could be used against some vital aspect of the host-parasite relationship, such as degradation of host proteins as a source of amino acids for parasite development (Renslo & McKerrow, 2006).

Historically, studies about therapies based on the structure of aspartic proteases began with HIV-1 protease and renin and led to a series of new drugs based on statin, such as AG-1343 designed as an inhibitor for HIV1 protease (Kaldor *et a*l., 1997) (**Figure 3**). Other studies with different aspartic proteases are currently being made in order to repeat this success and achieve new chemotherapeutics: (i) BACE-1 (B amyloid precursor protein-cleaving enzyme 1), involved in accumulation of amyloid

forms in the brain, causing Alzheimer's disease; (ii) Cathepsin E, protein present in the rough endoplasmic reticulum involved in processing major histocompatibility complex class II (MHC-II) and the endothelial vasoconstrictor peptide, in addition to having a possibly important role in cerebral ischemia; (iii) Cathepsin D, whose over-expression is found in cancerous tissues.

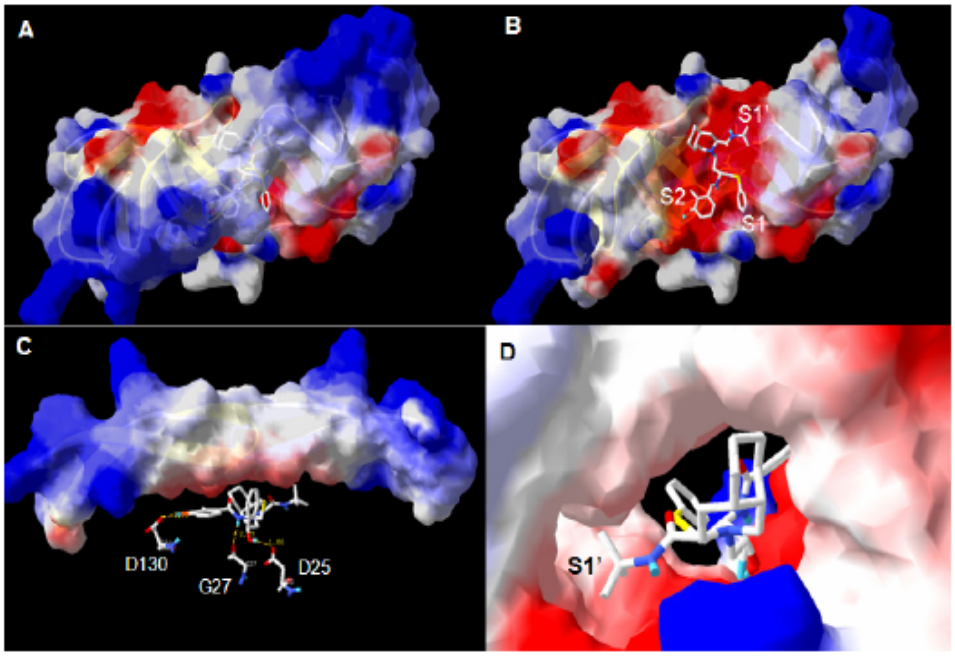


Figure 3: Interaction of AG-1343 with HIV-1 protease: A: View perpendicular, showing the concealment of inhibitor, B: Perpendicular view of the region covering the inhibitor-protein interaction, C: View of the top of the molecule, D: Surface subsites of the S1 (adapted from Silva-Jr & De Simone, 2005).

Exploration of Schistosoma aspartic protease function as a drug target originally started with the report by Becker *et al*. (1995) of a cDNA sequence of *S. japonicum* aspartic protease zymogen of 380 amino acid residues with 50-55% identity with the human cathepsin D (here called SjCD1). Through alignment with other enzymes, Becker *et al*. proposed an enzyme that has a mature form with 329 amino acid residues and 36kDa. Analysis of Southern blots allowed the same authors to infer that this gene would be in a single copy in *S. japonicum* and should be present in *S. mansoni* as well. In fact, Wong *et al*. (1997) reported the discovery of a putative aspartic protease from screening of cDNA libraries of the adult worms of *S. mansoni.*

This sequence (here called SmCD1) codes for an enzyme with 50-55% identity with other orthologous enzymes. The importance of these proteases for the parasite lies in the possibility that they have a broad spectrum of substrates, from proteins involved in the immune system as IgG and proteins of the complement system (Verity *et al*., 2002), serum albumin and hemoglobin. This could infer a role for aspartic proteases in defense of the parasite against the host immune system in addition to their use in the metabolism of amino acids.

Interestingly, *S. japonicum*'s AP has 3 different cleavage sites for human hemoglobin than the human cathepsin D orthologue while S*. mansoni*'s AP has 15 different sites (Brindley *et al*., 2001) and structures obtained *in silico* through molecular modeling by homology of SmCD1 showed similarities to HIV-1 protease and some differences to its human counterpart enzyme, cathepsin D (Silva-Jr *et al*., 2002). SmCD1 does not present a loop between residues L94 and G95 and a C-terminal

protuberance affecting the binding the substrate as human cathepsin D. Besides that, SmCD1 and human cathepsin D seem to have different volumes and surface electrostatic potentials in the binding subsites, according to diferences in the amino acid residue composition. These findings can open the path for development of antiretroviral drugs for anti-schistosomiasis therapy (Silva-Jr *et al*., 2002).

The aspartic proteases of *Schistossoma spp* have more functional and structural similarities with bovine cathepsin D, than with human cathepsin D (e.g. common epitopes found in the SmCD1 and in the bovine enzyme, but not in human cathepsin D) which could be interesting for future immunodiagnostic and drug discovery applications (Valdivieso *et al*., 2003), although it's potential as a vaccine was considered modest (Verity *et al*., 2002).

Families of genes of APs have been described in different groups of eukaryotic life forms, including the Apicomplexa parasite *Plasmodium sp*. (Jean *et al*., 2001) and *Toxoplasma gondii* (Shea *et al*., 2007), the nematode *Caenorhabditis elegans* (Tcherepanova *et al*., 2000) and *Onchocerca volvulus* (Jolodar & Miller, 1998). Therefore, it was not entirely surprising when the first results of transcriptome project of *S. mansoni* showed fragments of coding sequences of additional aspartyl proteases to SmCD1 (Merrick *et al*., 2003). Allied to this, a two-dimensional electrophoretic analysis (2DE) of an extract of adult worms linked to pepstatin-agarose, prepared at the Laboratory of Biochemistry of Proteins and Peptides (LBPP) of IOC / FIOCRUZ (Silva, Jr, 2005), revealed the presence of different protein species in the range of 36-50 kDa. Further research undertaken by the Brazilian group analyzing EST"s databanks of *S.*

mansoni indicated that a family of aspartyl proteases is present in this parasite (here collectively called SmAPs).

Over the past five years, efforts to characterize the new members of the aspartic protease family S.*mansoni* resulted in the complete sequencing of the cDNA of two new genes similar to Cathepsin D, which were named SmCD2 and SmCD3 (deposited in Genbank under numbers DQ256465 of access and EF373976, respectively). SmCD2 has an open reading frame (ORF) of 1203 bp and genomic structure with 7 exons, similar to SmCD1, while SmCD3 has an ORF of 1161 bp, and has only 6 exons. Both genes encode pre-proenzymes of 401 and 387 amino acid residues respectively, and contain the signature D (T / S) G residues typical of the active site of aspartic proteases from the pepsin family. The computational analysis and sequence conservation suggests a signal peptide of 19 residues for SmCD2 and a propeptide of 43 amino acid residues covering residues 20-62. Similar analysis suggested SmCD3 has a signal peptide of 15 residues and pro-region covering residues 16-52. SmCD2 has 5 potential sites for N-glycosylation at aminoacid residues 32, 43, 198, 246 and 348, while SmCD3 has only two sites at positions 54 and 136.

Currently, therapeutic approaches based on RNA interference technology, a technique that can knock out AP genes in order to observe gene function in the parasite's biology, led to opposite conclusions. Skelly *et al*. (2003) observed that the expression of the SmCD1 did not result in adverse effects on parasite growth cultivated *in vitro*, indicating that chemotherapy interventions need to consider the redundancy of protease functions involved in hemoglobin degradation. On the other hand, RNA

interference studies by Morales *et al*., 2008, suggest blockage of about 70 percent of the activity of the enzyme and a decrease in the rate of growth of S. *mansoni*, implying that degradation of hemoglobin would not be redundant.

1A- aspartic proteases of the pepsin family (A1)

Aspartic proteases, (APs), as all peptidases, catalyze the hydrolysis of peptide bonds, a fundamental chemical reaction in biology. Since APs are involved in a wide range of cellular, physiological and pathological processes (Eder *et al.*, 2007; Turk, 2006; Rawlings & Barret, 1993, 2000) these enzymes have been used in diverse industrial processes such as cheese and flavor enhancement compound production as well as new therapeutic research/production at pharmaceutical companies (Horimoto,2009).

The electronic database MEROPS, (http://merops.sanger.ac.uk/), shows in its catalog seven clans of aspartic proteases, divided into 14 families, 3855 described sequences and 37 three-dimensional structures deposited in the PDB from fungi, plants, animals, bacteria, viruses, protozoa and archea. Although family A22 (preselinin-like) has gained increasing attention due to new discoveries associating these aspartic proteases to neurodegenerative diseases (Parks & Curtis, 2007), the most important families are A1 (pepsin-like family) and A2 (HIV1 aspartic protease-like family). Family A1 apparently originated from a gene duplication event related to adaptations to new function (Carginale *et al*., 2004), (Dunn, 2002). **Table 1** summarizes the various biological functions and industrial applications of the family A1 enzymes.

Table 1-Biological and Biotechnological functions of AP family A1 enzymes

Enzyme	Function
Pepsin	Nutrition of vertebrates
MemAPsin I	Involved in Alzheimer's disease
MemAPsin II	Involved in Alzheimer's disease
Cathepsin D	Involved in cancer, inflammation, immune response, nutrition of parasites, apoptosys, autophagy
Cathepsin E	Involved in lysosomal disorders, cancer and immune response
Renin	Involved in blood pressure control
Chymosin	Involved in digestion of milk in neonates and production of cheese
Plasmepsin I	Nutrition of parasites (Plasmodium)
Plasmepsin II	Nutrition of parasites (Plasmodium)
Gastricsin	Nutrition of vertebrates

1B-Structure, folding and activity of aspartic proteases of cathepsin D-like proteases

1.B.1- General structure and Catalytic mechanism

Family A1 APs are very similar in terms of general structure, showing a high degree of structural homology with minor variations (Tyndall *et al*,2005). The secondary structure is entirely formed by β sheets with few α-helices and the tertiary structure shows the presence of two chains linked by a loop region known as a flap. The flap has greater structural flexibility and lays above the active site cleft (Koelsh, 1994; Dunn, 2002). The molecular weight is generally around 35-45 kDa. Despite these similarities, these APs have different functions and cellular localizations as shown in **Table 1.**

The catalytic site of aspartic proteases from the A1 family follows fundamental principles that guide the hydrolysis of peptide bonds such as the complementarity and electrostatic interaction between the surfaces of the proteins. The active sites of the protease are capable of accommodating a single sidechain of the substrate in surface clefts called subsites (Turk, 2006). Subsites are numbered S1-Sn towards the N-terminal and S1'-Sn' towards the C-terminal, beginning from the subsites closer to the scissile bond. The substrate side chains accommodated in this subsite are named as P1-Pn and P1'-Pn in the same way previously described, as it can be seen in **Figure 4** (Turk, 2006).

The AP catalytic site is larger compared with other enzymes and forces the substrate to adopt an extended β strand conformation, stretching the substrate polypeptide in order to facilitate reaction (Dunn,2000; Tindall *et al*, 2005). The main

reason for such a conformation is that linear extension of the substrate polypeptide exposes maximally all the main chain amide atoms of the substrate to the solvent and protease residues inside the subsites (Tyndall *et al*, 2005). The correct solvation of the product as consequence of this arrangement prevents peptide bonding between substrate and product, driving the equilibrium for reaction toward product formation (Tyndall *et al*,2005). On the other hand, helix/turn structures are too large to bind correctly in the subsites pockets confirmed by the fact that proteases do not recognize these structures unless they are unfolded. Such conformational selection explains why the resistance of folded regions of some proteins leads to active site selectivity to proteolytic activity (Tyndall *et al*,2005).

In fact, this structural aspect can be one of several ways cells have to precisely control proteolytic activity in space and time, considering the high diversity of biological processes in which these enzymes are involved. Other ways can be: (i) via a cascade of enzyme activities; (ii) specific AP degradation (iii) blockade of proteolytic activity by natural inhibitors; (iv) influence of cellular compartmentalization and chemical microenvironment; (v) regulation of gene expression; and (vi) regulation by substrate and/or product inhibition or other chemical compounds (Lazure, 2002)

The proposed reaction mechanism occurs in an acidic environment (pH 2-5), as shown in **Figure 4** where the aspartic residue carboxyl group is in its acidic form and the other is in its basic form in the enzyme substrate complex. This arrangement activates a water molecule which then attacks the carbonyl carbon of the scissile bond, resulting in the formation of a tetrahedral geminal diol intermediate. Subsequently

deprotonation of the hydroxyl group occurs by one of the catalytic aspartates, with a simultaneous activation of the leaving amine by the other protonated aspartic residue, which leads to the peptide bond cleavage (Eder *et al*, 2007).

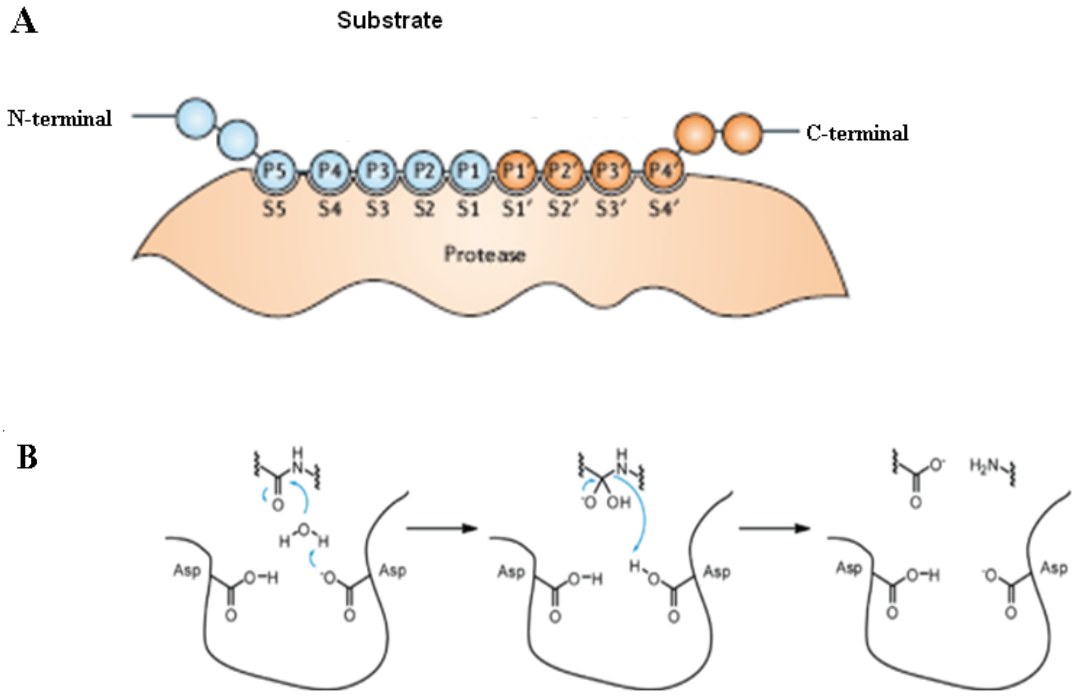


Figure 4: Schematic representation of molecular recognition of proteases and the reaction mechanism of aspartic proteases. A) The surface of the protease able to accommodate the substrate is called subsite. Subsites are named as S1–Sn towards the N terminus of the substrate, and S1'–Sn' towards the C terminus, beginning from the scissile bond. The substrate residues they accommodate are named as P1–Pn, and P1'–Pn', respectively(adapted from Turk, 2006). ; B) Mechanism for the general acid-base aspartic protease reaction: one aspartate residue activates a water molecule by

abstracting a proton. This enables water to attack the substrate scissile bond, generating an intermediate. Rearrangement of this intermediate leads to protonation of the scissile amide (adapted from Eder *et al*, 2007).

1.B.2 Importance of the prosegment

All vertebrates and most of the fungal APs are synthesized as zymogens, inactive precursors of the active and mature enzyme. This prosegment occludes the active site from the rest of the enzyme through a very intimate non-covalent bonding, in the same way as the substrates do for conversion of zymogen to the mature form (Lazure, 2002). Different mechanistic models have been proposed and those are summarized by Koelsh *et al* (1994) as: i) Complete self processing (pro-pepsin); ii) Partially self processing assisted by different proteins (lysosomal and vacuolar peptidases); and iii) Fully assisted processing (pro-renin). All of these mechanisms occur at acidic pH values so the disruption of electrostatic interactions can occur between the ~ 54 residue prosegment polypeptide and the rest of the enzyme (Horimoto *et al*, 2009).

AP prosegments can be divided into two types: Type 1, has the ability to act as an intramolecular chaperone, lowering the energy barrier among folded and unfolded states or stabilizing the folded state in a kinetic way; Type 2: besides working as an intramolecular chaperone, the prosegment also performs other possible functions. This feature can be related to the fact that prosegment regions have fewer similarities than other AP enzyme parts. (Fortenberry & Chigwin-1995). However, the prosegments of

Cathepsin D and Pepsin are able to interact with non-parental AP, suggesting common structural motifs (Fusek *et al*, 1991, Ma *et al*,2006).

In plasmepsin II, the prosegment seems to be very important for the correct folding and catalytic activity of this AP, since expression of the enzyme with a prosegment shorter than normal led to changes in the structure of the mature enzyme (Khan *et al*., 1999). Also a chimeric plasmepsin II with a pepsinogen prosegment fusion instead of the natural plasmepsin II prosegment showed instability and was considered misfolded (Horimoto *et al*, 2009). The zymogen conversion into the mature form seems to be different from other A1 aspartic proteases, since the prosegment of plasmepsin II associates with both C and N terminals to keep two domains apart from each other. This is considered a folding intermediate of the protein, once the lobes are assembled but not connected correctly due to charge repulsion of catalytic aspartates presented in this configuration. (Horimoto *et al*, 2009).

Pepsin, a human digestive aspartic protease, has its precursos form (pepsinogen) converted to the mature form when the pH shift towards acid causes destabilization of salt bridges that maintain the prosegment in place. In addition, the prosegment acts in the correct folding of the mature form pepsin, lowering the free energy barrier and accelerating the folding process to a practical timescale (Dee *et al*, 2009). In the cases of cathepsin D and cathepsin E, the prosegments may be involved in cellular localization and also can have different functions inside the cell (Horimoto *et al*, 2009)

1C- Expression and Refolding procedures for recombinant APs.

The study of structure-function relationships has enabled advances in understanding the logic of life at the molecular level. Some of the techniques used in this field, such as X ray crystallography, small Angle X ray scattering (SAXS) and nuclear magnetic resonance (NMR) spectroscopy require large protein quantities in order to solve the structure (Minor, 2007). Therefore, many proteins have to be produced through recombinant DNA technology, often utilizing hosts with a cellular apparatus that may not resemble the original organism.

SmCDs and related parasite proteins such as *Necator americanus* and *Ancylostoma duodenale* aspartic proteases (named respectively as Ac-APR-1 and Na-APR-1) were previously expressed in insect cells. SmCDs, Ac-APR-1 and Na-APR-1 genes were cloned into a pBacPak6 and co-transfected with baculovirus and SF9 insect cells. After generation of baculovirus stocks, Trichoplusia ni High Five cells were infected to express these aspartic proteases as secreted pro-enzymes. Those pro-enzymes were purified successfully using pepstatin-affinity columns (Brindley et al.2001) and the enzymatic activity confirmed via degradation of hemoglobin, IgG, synthetic fluorogenic peptides, and human collagens (I, III, IV and V), canine and human serum albumin, and canine and human fibrinogen (Beyer & Dunn, 1996 ; Brindley et al. 2001, Verity *et al,* 2001 Willianson *et al*, 2003). However, Differences in proteins expressed by in baculovirus infected insect cells and original life forms have been described in some cases. For example, inefficient secretion from insect cells and improperly folded and/or glycosylated proteins, or low levels of expression.

On the other side, SmCDs were not yet expressed in bacteria, although related aspartic proteases, such as human cathepsin D, Na-APR-1 and plasmepsin were already expressed in this heterolougous system (Ranjit *et al*, 2009 Pearson *et al*, 2010). *E. coli* is the main host for recombinant protein production due primarily to: (i) Modest nutritional requirements of the culture medium; (ii) Rapid accumulation of biomass; (iii) Tolerance of the fermentation process at high cell density; (iv) Process of scaling-up production is simple; and (v) Knowledge of E. *coli* cell physiology, molecular biology and metabolism. Some of the heterologous proteins are sometimes expressed in inactive and incorrectly folded form or even in a completely insoluble form known as inclusion bodies (Boney & Mujacic, 2004). Although the formation of inclusion bodies is interesting in terms of ease of purification steps from cell homogenates and protection from protease degradation, the process of solubilization and refolding is laborious and more often leads to a lower fraction of active protein.

Another major problem in refolding is the precise determination of what is the relevant protein activity to be assayed in order to evaluate the success of the refolding process. It is well known in modern protein folding theories that the polypeptide chains have plasticity and possible multifunctionality. This plasticity is due to the dynamic interaction of the polypeptide chain with microchemical and physical (temperature, pressure, volume, and viscosity) environments inside cellular and body compartments. This could lead to different functions for the same protein. Thus, a refolding procedure could make the protein achieve a different folding and function than those expected by

the scientist and, on the other hand, new possible functions could be not explored (Zhou, 2008, Shin *et al*, 2006 Frauenfelder *et al*, 2006, Dix, 2008).

Plasticity in function can be seen in human Cathepsin D, since its proteolitic activity is not the major reason of increased proliferation of cells in tumor tissues (Fusek and Vetvicka, 1994). The mitogenic function *in vitro seems to be* strictly connected to the use of procathepsin D and not the mature form, thus suggesting the importance of ligand binding activity of the polypeptide alongside the proteolysis function. Presence of pepstatin A into the sample tested did not impaired procathepsin D mitogenic activity . The argument was also supported by experiments showing maintenance of mitogenic activity in a proteolytic inactive mutated form of procathepsin D. Besides this, enzymatically inactive procathepsin D can be secreted by various cell types including macrophages, keratinocytes, mammary epithelial cells and cancer cells of different origins, acting as a paracrine and autocrine mitogen on cancer cells, fibroblasts and keratinocytes.

Therefore, it is possible that many proteins would fold correctly only during or after interaction with a binding partner, such as chaperones, protein disulfide isomerases or even different proteases along a proteolitic cascade.

In this study the major goal is to provide basic information for SmCD production in *E.coli* and refolding in large amounts. This is aimed at future drug screening trials to inhibit the proteolitic activity that may lead to novel chemotherapies against schistosomiasis. One strategy to refold SmCDs investigated was the use of *Schistosoma mansonii* protein disulfide isomerase (PDI) in order to rearrange disulfide

bonds of the aspartic protease as well as provide a chaperone environment able to assist refolding.

1D- Protein Disulfide Isomerases

Protein disulfide isomerase (PDI) is a member of the thioredoxin family, which is composed of several redox proteins (Dellagi *et al*, 2002). PDIs also display chaperone activity and are highly abundant ER luminal proteins in mammalian and yeast cells, essential tools for assisting unfolded or incorrectly folded proteins to their native states (Dellagi *et al*, 2002). Evidence exists that PDIs can serve as potential links between thiol–disulfide systems and oxidative protein folding, participating in other ER redox systems. The first evidence involves vitamin K epoxide reductase (VKOR), responsible for regeneration of vitamin K1, an essential redox cofactor for carboxylation of glutamic acid residues. The second evidence is a possible regulatory role in NAD(P)H oxidase production of reactive oxygen species in vascular smooth muscle cells. Thirdly an intriguing function of certain PDIs is related to homeostasis of the ER luminal calcium ion concentration ([Ca2+]ER). Ca2+ pumps and channels in the ER membrane are responsible for regulating [Ca2+]ER. In most tissues this function is carried out by sarco(endo)plasmic reticulum Ca2+ ATPase isoform 2b (SERCA2b) and by inositol 1,4,5- trisphosphate receptor type I (IP3R1). Both proteins are subject to regulation by a PDI family member interacting with an ER luminal domain (Appenzeller-Herzog & Ellgaard, 2007).

The overall architecture of PDI comprises (in order from the N-terminus) domains a-b-b_0-a_0-c, where a and a_0 are homologous thioredoxin-like domains containing a redox active site with the sequence motif WCGHCK. The b and b0 are also thioredoxin-like but lack the redox active site. The c domain is a C-terminal region rich in acidic residues. Sometimes the structure also includes an x interdomain linker comprising 19-residues linking the b_0-a_0 domains. PDI structure has a twisted U-shaped molecule where e b domain extends the hydrophobic site on b_0 and the a and a_0 domains faces each other in the central cleft. (Dellagi *et al*,2002 ; Wang & Narayan, 2008). Each domain adopts a typical thioredoxin-like fold consisting of a central five-stranded beta sheet flanked on each side by two alpha helices forming a U shaped structure which is lined with hydrophobic residues thought to constitute the ligand binding site (**Figure 5).**

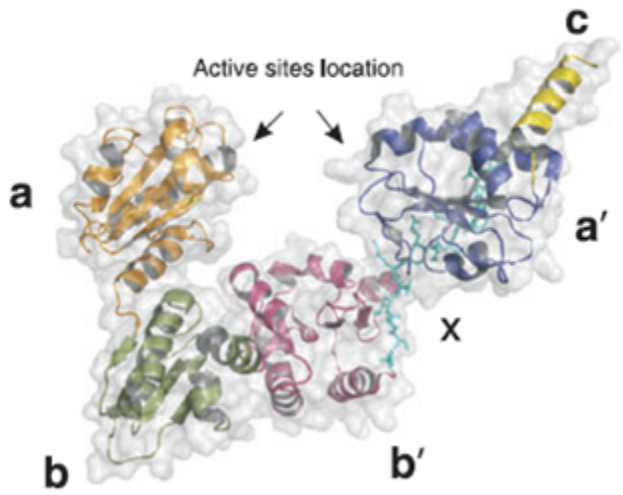


Figure 5: Crystal structure of PDI1P from S. cerevesae, showing the general structure of the four thioredoxin fold domains, a, a',b, b' and the x-linker region. The active site motifs marked with arrows (adapted from Hatahet, 2007)

The relatively weak binding of the PDI domain bb_0 to peptides and proteins and the large size of its binding pocket with a large hydrophobic surface is consistent with a low degree of specificity for ligands. High specificity is not expected because PDI acts on many substrates with different primary sequence. This site appears to be responsible for the saturable kinetics observed for RNase A folding by PDI, and blocking the site strongly inhibits the activity of PDI. (Denisov *et al, 2008)*

Crystals of yeast Ppi1p has shown flexibility in its conformation specially related to a domain orientation within the whole molecule and this flexibility would tend to prevent formation of crystals (Kozlov *et al*, 2009; Byrne *et al*,2009). The absence of good high resolution structural data for full-length human PDI (despite mammalian PDI having been abundantly available for more than 25 years) has been assumed to arise from molecular motions (e.g., independent motions of domains), but it could also arise from sample heterogeneity due to the existence of multiple conformers or multiple states of association (Wallis *et al*, 2009).

The catalytic cysteines of PDI are about 50% deprotonated at pH 7.1. Therefore redox potentials largely arises from differences in the thiol pK value of the N-terminal cysteine, which serves as the primary nucleophile that attacks the disulfide linkages of the protein substrate. As a result, the pK of the cysteine thiol provides an indication of whether these proteins will act predominantly as oxidants or as reductants under physiological conditions (**Figure 5**). The combination of redox and chaperone-like

activities allows PDI to bind to partly folded proteins and catalyze simultaneously protein folding and formation of disulfide bond (Hernandez *et al*, 2008).

Proteins with thioredoxin fold like PDI generally have two main regions that influence their function: the C*XX*C motif (“redox rheostat”) in the active site and a loop containing a *cis*-proline residue closer to the C*XX*C motif. Experiments revealed that the exchange of *X-X* dipeptide between two different PDI’s resulted in partially shifted redox potential. This is because sequence identity between PDIs for the active site is < 10%, making it difficult to compare PDI’s catalysis accurately (Ren *et a*l, 2008; Denisov *et al, 2008*). On the other hand, the loop containing the *cis*-proline residue is well conserved and biophysical studies of *E.coli* thioredoxin and DsbA indicated a significant role in the structural stability and modulation of substrate binding and activity of PDI’s (Ren *et a*l, 2008).

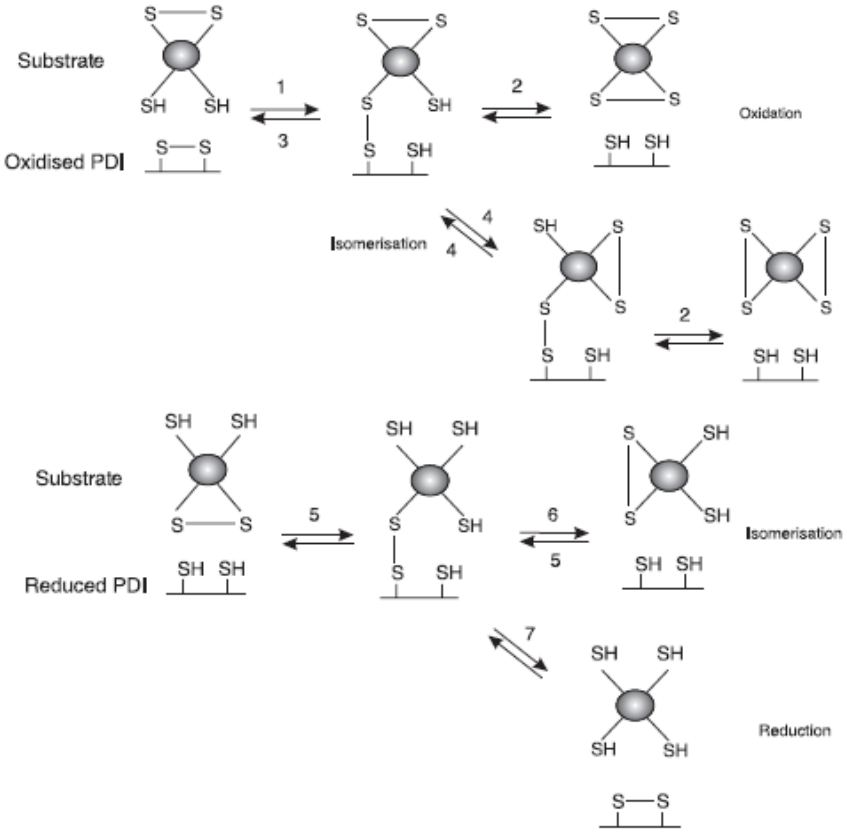


Figure 6- Reactions of oxidation, reduction and isomerization catalyzed by PDI. When PDI is in the oxidized state, the first reaction is the formation of a mixed disulfide by a nucleophilic attack of a thiol substrate (reaction 1). From the mixed disulfide, there are three possible options: First option is a nucleophilic attack by a substrate thiol on the PDI–substrate mixed disulfide (reaction 2) t resulting in oxidation of the substrate and reduction of PDI; The second option is a nucleophilic attack of PDI on the mixed disulfide (reaction 3); The third is a nucleophilic attack by a substrate thiol on a substrate disulfide (reaction 4), resulting in an isomerization reaction. When PDI is in the reduced state, the first reaction is the formation of a mixed disulfide by the nucleophilic attack of PDI on a disulfide bond in the substrate (reaction 5). Again, there

are three options: The first option is the release of reactants following nucleophilic attack by a substrate thiol on the mixed disulfide; The second option is the nucleophilic attack of a different substrate thiol on the mixed disulfide, resulting in isomerization reaction in the substrate protein (reaction 6). The third is the nucleophilic attack by PDI on the mixed disulfide, resulting in reduction of the substrate (reaction 7) (adapted from Hatahet *et al*, 2007).

Recently, several studies indicated that DsbA, a bacterial homologue of PDI, is involved in the biogenesis of the enterotoxin and the toxin-coregulated pilli of *Vibrio cholera*, making it a possible good drug target. DsbA i also catalyzes disulfide bond formation in a pili-specific chaperone, PapD in enteropathogenic *E. coli*. species, It has been shown that a PDI-mediated reduction of a disulfide bond of gp120 is required for the entry of HIV into the host cells and infection. Recently, it was demonstrated that the presence of native PDI on the surface of epithelial cells is required for infection by Chlamydia *trachomatis*, a leading bacterial agent causing sexually transmitted diseases (Raturi & Mutus, 2007).

Also, several Plasmodium sp antigens relevant for virulence and pathogenesis of malaria possess a high number of cysteine-residues and P. *falciparum* has four PDI molecules (PfPDI's) displaying a high degree of diversity in their amino acid sequences, structure and the number of thioredoxin-like domains to assist folding of those antigens (Mahajan *et al* 2008). This fact and the fact that homologous and heterologous PDI's have been used to improve the folding of foreign proteins expressed in yeast cells raised the perspective that PfPDI's could be used to improve folding of several malaria

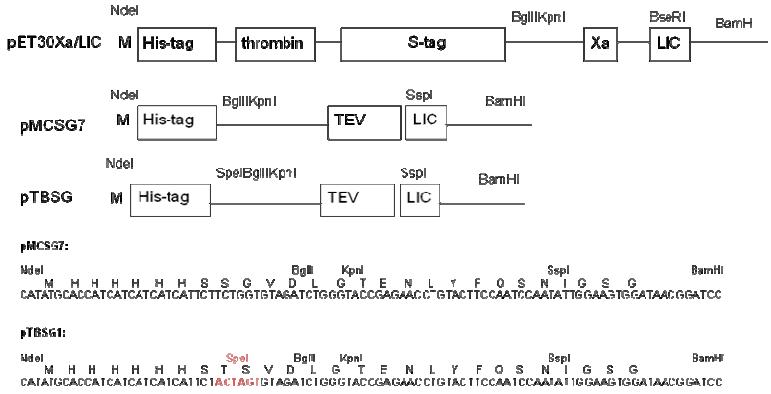
proteins produced via E.*coli* hosts and used in vaccines with biologically and immunologically inferiority to the native parasite antigens. (Mahajan *et al* 2008).

Chapter 2- Materials and Methods

Enzymes production

Aspartic protease from *S. mansoni* - E.*coli* production

In previous work by Silva, Jr (data not published), cDNA of the proform of the aspartic protease proSmCD1 was inserted into the pTBSG1 vector (built on the basis of vector pMCSG7, which is based on vector pET30Xa/LIC – Novagen, as shown in **scheme 1**), containing a six histidine tag (His6) on the N-terminus. This vector was inserted into E.*coli* strain BL21 (DE3) pLys-S for the expression of SmCD1 protein.



Scheme 1: Main features of the vector pTBSG and their precursors, pMCSG7 and pET30.

The newly transformed cells were incubated in 25 ml of Luria-Bertani (LB) media with 100 µg/ml ampicillin and 34 µg/ml of chloramphenicol in 125 ml Erlenmeyers in a shaking incubator overnight at 200 rpm, 37 ℃. The cells were inoculated into 1000 ml of liquid LB medium with the same concentration of antibiotics. The culture was incubated for 2 hours, 37 °C and shaking at 200 rpm. Readings of optical density (OD)

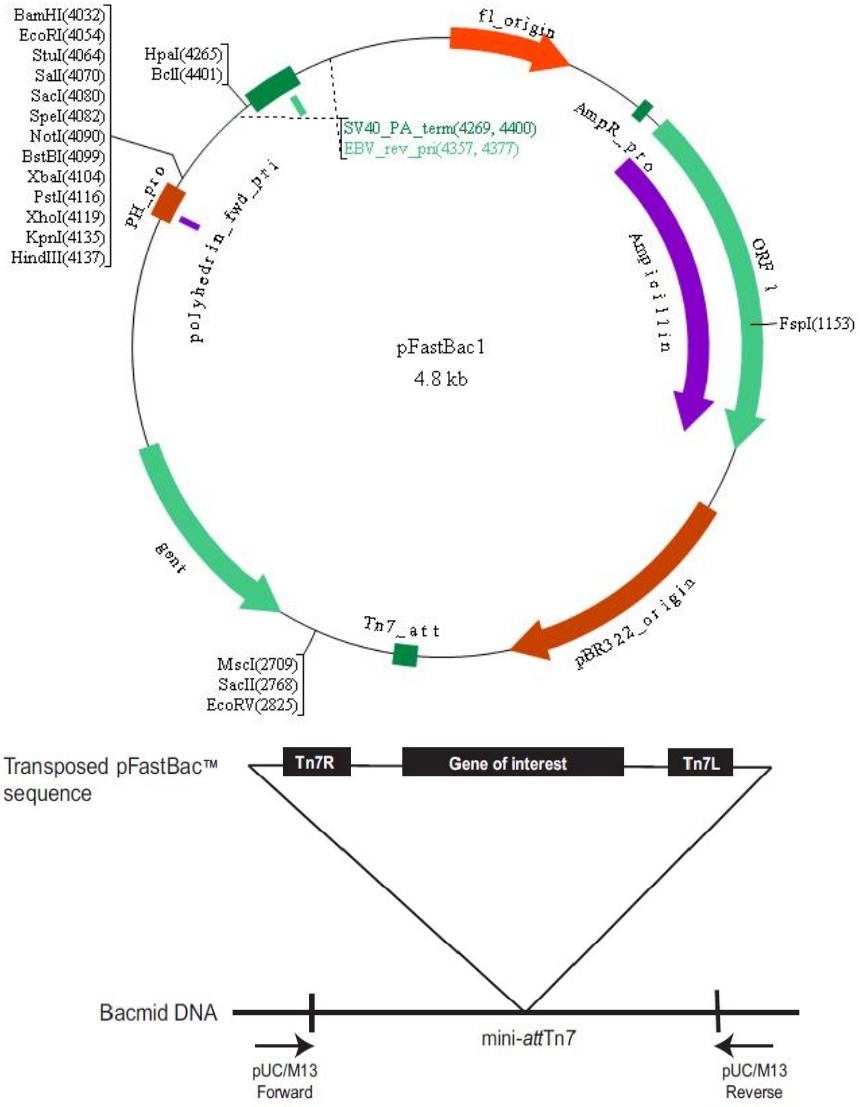
at 600 nm were taken until the OD reached 0.4-0.6. Induction of protein synthesis was initiated by adding 0.5 mM β-D Isopropyl-1-thiogalactopiranoside (IPTG) and the culture was incubated for 4 hours while shaking at 200 rpm at 37 °C.

The culture was centrifuged at 10,000xg for 10 minutes at 4 °C and the pellet was stored at -80 °C overnight. Cells were resuspended in 25 mM Tris-HCl (pH 8.0) disrupted by freezing and thawing followed by 10 min. sonication in a Sonifier 450 Branson system. Centrifugation of this material for 20 min at 4°C and 10,000xg allowed separation of inclusion bodies from the supernatant. Inclusion bodies were stored at -80 °C.

Aspartic protease from *S. mansoni* - **SF9 production**

The cDNA of the proform of aspartic protease SmCD1 was inserted into pFastBac1 LIC vector using Bac-to-Bac® Baculovirus Expression System from Invitrogen. The vector with the SmCD1 insert was transformed into E.*coli* DH10Bac™ to be integrated into the baculovirus shuttle vector (bacmid), bMON14272, containing a low-copy number mini-F replicon, Kanamycin resistance marker and a segment of DNA encoding the LacZ peptide. The bacmid propagates in E.*coli* DH10Bac™ to form colonies that are blue (Lac+) in the presence of a chromogenic substrate such as Blu-gal or X-gal and the inducer, IPTG. Recombinant bacmids (composite bacmids) are generated by transposing a miniTn7 element from a pFastBac™donor plasmid to the mini-attTn7 attachment site on the bacmid. The Tn7 transposition functions are provided

by a helper plasmid (pMON7124) which encodes a transposase and confers resistance to tetracycline.



Scheme 2: Main features of the vector pFastBac for insect cell expression, showing where ProSmCD was cloned as the gene of interest.

The recombinant E.*coli* colony is isolated and its bacmid isolated and used to transfect SF9 insect cells. An SF9 insect culture of 8×10^5 cells per ml per well with >95% viability was allowed to attach to 6 well plates for 30 minutes to achieve transfection. After attachment, 1 µg of Bacmid was diluted with 8 µl of Cellfectin (Invitrogen) in 100 µl of Grace's insect medium and kept at room temperature for 30 minutes. This mixture was dropped into the wells containing SF9 cells with medium and kept at 27 ℃ for 5 hours. Medium was removed and repl aced by Grace's insect medium and plates were kept at 27 ℃ for another 72 h ours until signs of infection were observed under a microscope. Cells were harvested to generate P0 virus stock. In order to increase multiplicity of infection (M.O.I.), a P1 recombinant virus stock was generated and used for protein expression. $2x10^7$ SF9 cells in 50 ml of medium (Expression Systems) were infected with P1 virus stock and incubated in a shaking incubator at 270 rpm at 27 °C for 7days in a 125 ml Erlenmeyer flask for protein production. Cells were harvested after 7 days for protein purification and the process repeated until an M.O.I of 0.1 was reached with P3 recombinant virus stock.

Proten Disufide Isomerase from S. *mansoni.*

The cDNA of S. *mansoni* protein disulfide isomerase, was inserted into a pTBSG1 vector containing a 6 Histidine tag on the N-terminus, and was transformed into strain BL21 (DE3) pLys-S for expression in E. *coli*. The newly transformed cells were incubated in 25 ml of Luria-Bertani (LB) media with 100 mg / ml ampicillin and 34

mg / ml of chloramphenicol in 125 ml Erlenmeyers in a shaking incubator overnight at 200 rpm, 37 °C. The cells were inoculated into 1000 ml of liquid LB medium with the same concentration of antibiotics. The culture was incubated for 2 hours at 37 °C and shaking at 200 rpm. Readings of optical density (OD) at 600 nm were taken until the OD reached 0.4-0.6. Induction of protein synthesis was initiated by adding 0.5 mM β-D Isopropyl-1-thiogalactopiranoside (IPTG) and the culture was incubated for 4 hours while shaking at 200 rpm at 37 °C.

The culture was centrifuged at 10000xg for 10 minutes at 4 °C and the cells were stored at -80 °C overnight. Cells were then disrupted by freezing and thawing followed by sonication in a system Sonifier 450 (Branson). Centrifugation of this material for 20 min at 4 °C and 20,000 rpm allowed separation of supernatant with soluble protein disulfide isomerase from the insoluble materials. Soluble protein was stored in 4 °C in 50 mM Tris-HCl buffer (pH 8.0) with 20 % glycerol.

Purification by affinity chromatography with immobilized metal IMAC

Samples of proteins were placed in His-tag affinity column (Qiagen) previously equilibrated with 25 mM Tris-HCl (pH 8.0) and washed once in the same buffer. The elution was performed with 25 mM Tris-HCl (pH 8.0) plus 300 mM imidazole. 1 ml fractions were collected and the protein concentration of each fraction was determined using the Bradford method.

Purification of Protein Disulfide Isomerase by ion-exchange in FPLC

Samples of protein disulfide isomerase purified from the His-tag affinity column were pooled and applied to an ion exchange column (Q-Sepharose-Amershan) in an AKTA FPLC system previously equilibrated with 25mm Tris-HCl buffer (pH 8.0) with 0.15 M NaCl and 10 mM beta-mercaptoethanol (Buffer A). Elution was performed with a gradient using a mixture of Buffer A and Buffer B, consisting of 25 mM Tris-HCl (pH 8.0) with 1 M NaCl and 10 mM 2-mercaptoethanol. Fractions of 1 ml each were collected and the protein concentration of each fraction was determined using the Bradford method.

Purification of Protein Disulfide Isomerase by size exclusion in FPLC

Samples of protein disulfide isomerase purified from the ion exchange column were pooled together in a final concentration of 2mg/ml and applied to a size exclusion column (Superdex™ 200 GE) in an AKTA FPLC system previously equilibrated with 25 mM Tris-HCl (pH 8.0) with 10 mM beta-mercaptoethanol and samples collected in 1 ml fractions. The protein concentration of each fraction was determined using the Bradford method and each fraction concentration were estimated to be 0.25 mg/ml and concentrated to 40 mg/ml using Amicon centrifugal filters of 3,000 MWCO.

Crystallization trials of Protein Disulfide Isomerase

A screening of crystallization conditions was performed with a sample of protein disulfide isomerase using several different conditions (presented in screening kits-Hampton Index, Wizard Crystal Screen I, II, III, IV) and the sitting drop technique. Plates of 96 wells were filled with 1 ml in each big well and with 0.1 ml in each small well with varying solution conditions as per the kits. A 0.5 µl of protein disulfide isomerase sample (40mg/ml) was added to each small well and plates were tapped and placed at either 4 °C or room temperature to allow crystal growth. Plates were periodically checked under a microscope to observe crystal growth.

Electrophoresis Sodium Dodecyl Sulfate Polyacrylamide Gel (SDS PAGE)

The polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate (SDS-PAGE) was performed in 12% polyacrylamide gels as described by Laemmli (1970). Aliquots of purified aspartic proteases were mixed with an equal volume of sample buffer (0.187 M Tris-HCl, pH 7 + 0.2% SDS (w/v) + 1% Bromophenol blue (w / v) + 27% glycerol (v/v) + 5.0% (v/v) β-mercaptoethanol and heated for 3 min at 100℃. Gels were electrophoresed at 100 V for 90 min utes using a Mini-Protean III (Bio-Rad, USA) gel apparatus. The molecular weights of proteins were determined by comparing the migration of the protein with those of molecular weight markers (Sigma Chemical, USA). Proteins were stained using Coomassie Brilliant Blue G250.

Western Blots

Proteins separated in SDS PAGE gels were transferred to Immobilon PVDF 0.45 µm membranes (Millipore) using an XCell II Blot Module overnight at 10 V. Immobilon PVDF membranes were treated with KPL Western blot kit for overnight blocking and washing.

Refolding of aspartic protease from *S. mansoni* produced in *E.coli*

Initial refolding tests were performed using a slow dialysis method. This experiment was performed using aliquots of 1 mg/ml of SmCD1 diluted one hundred fold in 25mm Tris-HCl (pH 8.0) containing different chaotropic agents such as 20% glycerol, 0.5 M sucrose, 1M LiCl, 1M NaBr, 5mM GSSH/GSH, 1M arginine. These solutions were placed inside SpectraPore Dialysis membrane MWCO 3,500 and dialysed against 25 mM Tris-HCl (pH 8.0) at room temperature or 4°C overnight.

In another experiment an attempt was made to refold the protein using an Ni-NTA affinity resin previously equilibrated with 25 mM Tris-HCl (pH 8.0) with 500 mM NaCl, 8 M urea and 2.5 mM β-mercaptoethanol, in 4 °C. SmCD1 samples were mixed up for two hours with this resin and then the resin was washed out with 25 mM Tris-HCl (pH 8.0) with 500 mM NaCl, 2.5 mM β-mercaptoethanol and decreasing urea concentrations of urea from 6 M to 0 M. After the washing step, protein was eluted in fractions of 1 ml in the same buffer containing 250 mM imidazole and 20% glycerol.

PDI activity assays

Assays was performed using a modification to Sigma's protocol, using 0.75 ml of a cocktail made of 10 mM Na-phosphate (pH 7.0), 100 mM DTT, 100 mM Na-EDTA, incubated with 0.25 ml of different concentrations of insulin (100 µg/ml, 250 µg/ml, 500 µg/ml). The mixtures were incubated for 10 min at 25℃, 37℃ and 45℃. After incubation, 10 µl of protein disulfide isomerase were added into this solution containing insulin and turbidity at 650 nm of solution was monitored at 5 minute intervalsthe until 30 min. The same assay was used for refolding of SmCDs.

SmCD1 activity assays

SmCD1 proteolysis assays were performed using the method described by Brindley *et al*, (2001) for the insect cell-derived enzyme. SmCD1 was diluted to 100 µg/ml in 0.1 M citrate (pH 3.5) and incubated at 37 ℃ with and without human hemoglobin for 1, 2, 4, 8, 12, 24, 48 and 72 hours. Results were visualized in 12% SDS PAGE gels following electrophoresis.

Chapter 3- Results and Discussion

3.1-Refolding SmCDs.

ProSmCD1 cDNA was inserted into pTBSG1 vector with a six histidine tag (His6) on the N-terminus. This vector was inserted into E.*coli* strain BL21(DE3)pLysS and expression was carried at 200 rpm at 37 °C with IPTG Induction. Cells were harvested by centrifugation at 10.000 rpm at 4°C for 15 min and resuspended in 50 ml of 25 mM Tris-HCl pH 8.0 and then submitted to sonication to be lised. Cell lisates were centrifuged in 20.000 rpm at 4°C for 20 min and pellet of inclusion bodies were collected, resuspended and stored at -80 ℃ in 25 mM Tri s-HCl pH 8.0 with 20% glycerol. Figure 7 shows the inclusion body protein solubilized in urea and SDS or SDS alone, indicating that it is obtained in a significantly purified form prior to further purification.

In this study two main strategies were used in attempts to obtain soluble, correctly folded and active proSmCD protein from heterologous expression in *E. coli*. The functional state of the folded protein was determined by examining its ability to self cleave and/or cleave human hemoglobin. The first strategy involved dissolution of inclusion bodies followed by refolding assays manipulating folding conditions. The second strategy involved covalent attachment of a solubilizing protein to the proSmCD

protein in order to confer greater solubility such as the Maltose Binding Protein (MBP) (Baneyx & Mujacic, 2004, Sahdev *et al*., 2008).

As part of the first strategy, SmCD1 inclusion bodies were subjected to a variety of ways to disrupt non-covalent forces that result in loose protein-protein interactions leading to formation of inclusion bodies (Arakawa, 2007). This involved chemical agents (chaotropes) such as urea followed by several different formulations of solutions containing combinations of compounds able to favors refolding (cosmotropic). The simplest consisted of diluting SmCD1 inclusion bodies in an appropriate buffer containing a combination of amino acids (arginine), polyols (glycerol), alcohols, / reducing agents (DTT, Glutathione), EDTA , salts from the Hofmeister series, sugars (glucose) and detergents (DPPC) (Qoronfleh *et al*., 2007).

Two folding approaches were tested: dialysis, involving the slow and gradual withdrawal of the chaotrope used to dissolve the inclusion bodies and replacing it with the chemical agents, and chromatography consisting of an affinity resin that acts as a site of nucleation in order to allow refolding of the protein while a solvent with a combination of refolding agents is passed through. However, all of these combinations lead to a non auto-activate form of the protein that is unable to show self-clevage. An example of such a trial is shown in **Figure 8.** In this experiment, proSMCD 1 inclusion bodies were solubilized in 8 M urea then dialyzed against Tris buffer containing NaBr (1 M), LiCl (1 M) or urea (8 M) to keep the folded protein from aggregating. Shifting the pH from 8.0 to 3.5 is expected to induce self cleavage of the protein to an expected 45 kDa form however no such cleavage was observed in this and similar experiments. This

suggests that either the protein is not correctly folded or that activation via a pH shift requires some additional cofactor.

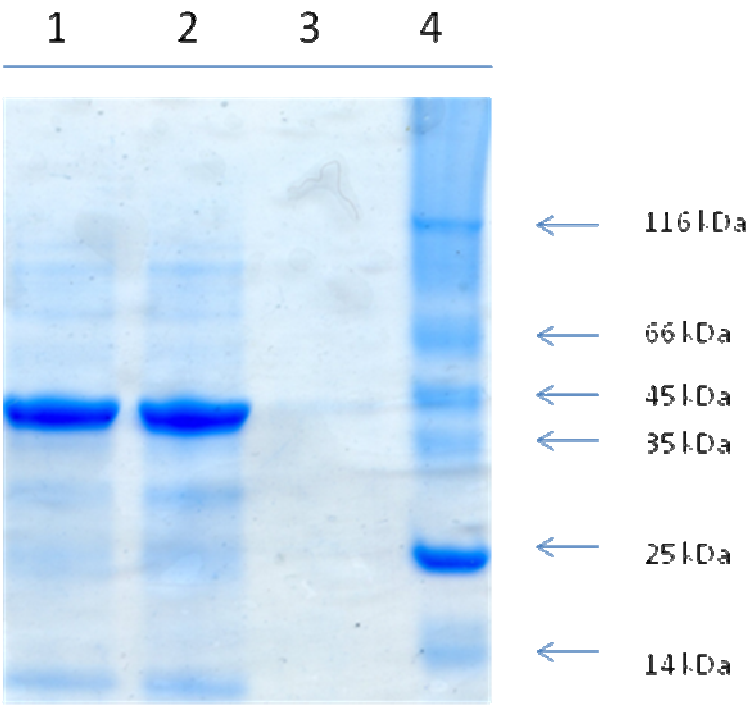


Figure 7: Pro SmCD1 inclusion bodies treated with 8 M Urea. Lanes 1: urea-solubilized proSmCD1 inclusion bodies kept at room temperature in 8 M Urea and 25 mM Tris-HCl (pH 8.0); lane 2: proSmCD1 inclusion bodies solubilized in sample buffer; Lane 4, Fisher E-z Run Protein Marker.

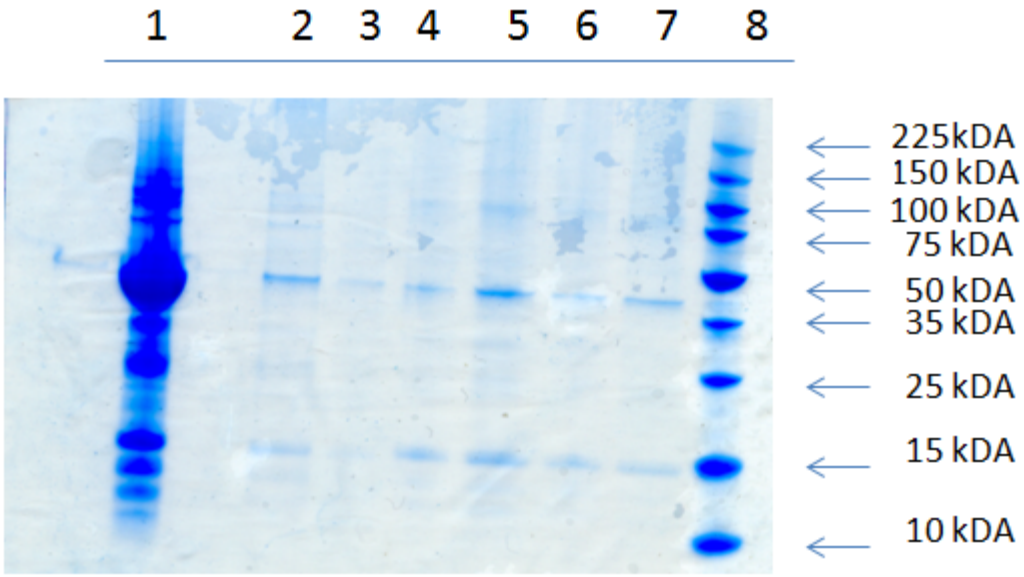


Figure 8: Pro SmCD1 inclusion bodies were solubilized in 25 mM Tris-HCl pH 8.0 with 8 M urea, dialyzed to maintain 8M Urea (Lanes 2 and 3) or 1 M NaBr (Lanes 4 and 5,) or to exchange 8M urea with 1 M LiCl (Lanes 6 and 7) to enhance solubility and dispersion. In some cases the pH was shifted to 3.5 by adding an equal volume of sodium citrate (pH 3.5) 2, 4 and 6). Lane 1: urea-solubilized proSmCD1 inclusion bodies (overloaded); Lane 28 M urea kept at pH 3.5; Lane 3: inclusion bodies in the presence of 8 M urea kept at pH 8.0; Lane 4: inclusion bodies in the presence of 1 M NaBr kept at pH 3.5; Lane 5: inclusion bodies in the presence of 1 M NaBr kept at pH 8.0; Lane 6: inclusion bodies in the presence of 1 M LiCl kept at pH 3.5; Lane 7: inclusion bodies in the presence of 1 M LiCl kept at pH 8.0; lane 8, Fisher E-z Run Protein Markers.

3.2 Use of Protein Disulfide Isomerase

The genome of *Schistosoma mansonii* encodes a protein disulfide isomerase, responsible for ensuring the correct formation of disulfide bridges as well as assisting the correct folding of proteins using an intrinsic chaperone domain (FInkenEigen *et al*, 2007). Therefore, PDI may be required for folding of SmCD proteins since these aspartic proteases have disulfide bonds that could be rearranged in a variety of ways. It is also possible that the PDI is required to assist with folding of the SmCD proteins.

To examine the potential role of PDI in SmCD folding, the cDNA of S. *mansoni* protein disulfide isomerase was cloned and inserted in a PTBSG vector and expressed in *E. coli* host cells as described in Chapter 2. PDI was expressed in soluble form and readily purified by Ni-affinity chromatography. Further purification to apparent homogeneity (as judged by SDS gel electrophoresis, not shown) was achieved by ion exchange chromatography (Figure 9) and size exclusion chromatography (Figure 10).

The recombinant PDI was active in refolding of insulin, a classical assay for this class of enzyme, as described previously and in the Chapter 2 **(Figure 11).** In these experiments, 0.75 ml of a cocktail made of 10 mM Na-phosphate (pH 7.0), 100 mM DTT, 100 mM Na-EDTA, incubated with 0.25 ml of different concentrations of insulin (100 µg/ml, 250 µg/ml, 500 µg/ml). This solution was incubated for 10 min at 25℃, 37℃ and 45 ℃. After incubation, 10 µl of protein disulfide isomerase were added into this

solution containing insulin and turbidity at 650 nm of solution was monitored from 5 to 5 minutes until 30 min.

In further experiments, folding trials to observe self-proteolysis of Pro-SmCD enzymes were undertaken in the presence of different concentrations of PDI. However, none resulted in proSmCD1 auto-activation. A typical experiment is shown in Figure 12. The lack of activation may indicate that the Schistosomal PDI is not used as a chaperone or isomerase to assist correct folding of SmCDs or there is still lacking an external factor such as another protease for SmCD activation.
Analysis of ion exchange and size exclusion chromatography led to the conclusion that Schistosoma mansoni PDI was sufficiently well purified to attempt crystallization trials. Extensive screening of crystallization conditions was performed using conditions presented in screening kits- Hampton Index, Wizard Crystal Screen I, II, III, IV- using the sitting drop technique in 96 well plates tapped and placed at either 4 ℃ or room temperature. However, PDI crystals were not obtained. One possible reason is that this enzyme has the same high degree of flexibility in the “a” domain as other members of the PDI family which make formation of crystals a difficult task (Kozlov *et al*, 2009)

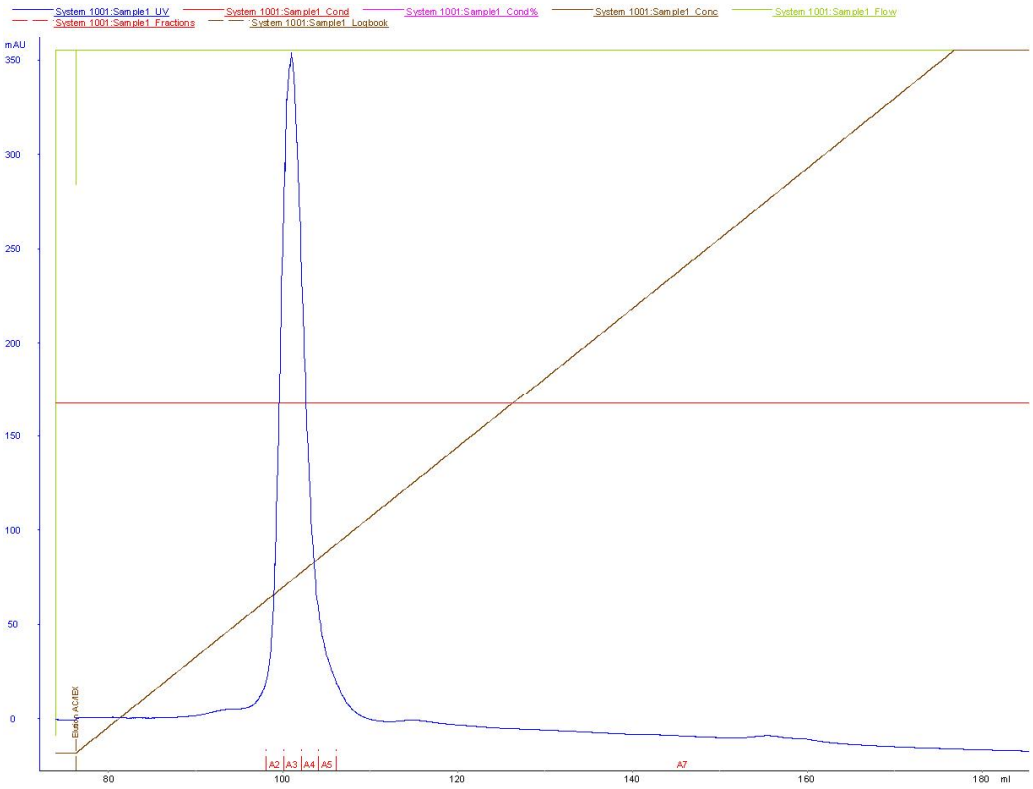


Figure 09: PDI purification by ion-exchange chromatography. A 10 ml sample of protein disulfide isomerase purified from the Ni-affinity column was applied to an ion exchange column (Q-Sepharose-Amershan) in an AKTA FPLC system previously equilibrated with 25 mM Tris-HCl buffer (pH 8.0) with 0.15 M NaCl and 10 mM beta-mercaptoethanol (Buffer A). Elution was performed with a gradient using a mixture of Buffer A and Buffer B, consisting of 25 mM Tris-HCl (pH 8.0) with 1 M NaCl and 10 mM beta-mercaptoethanol.

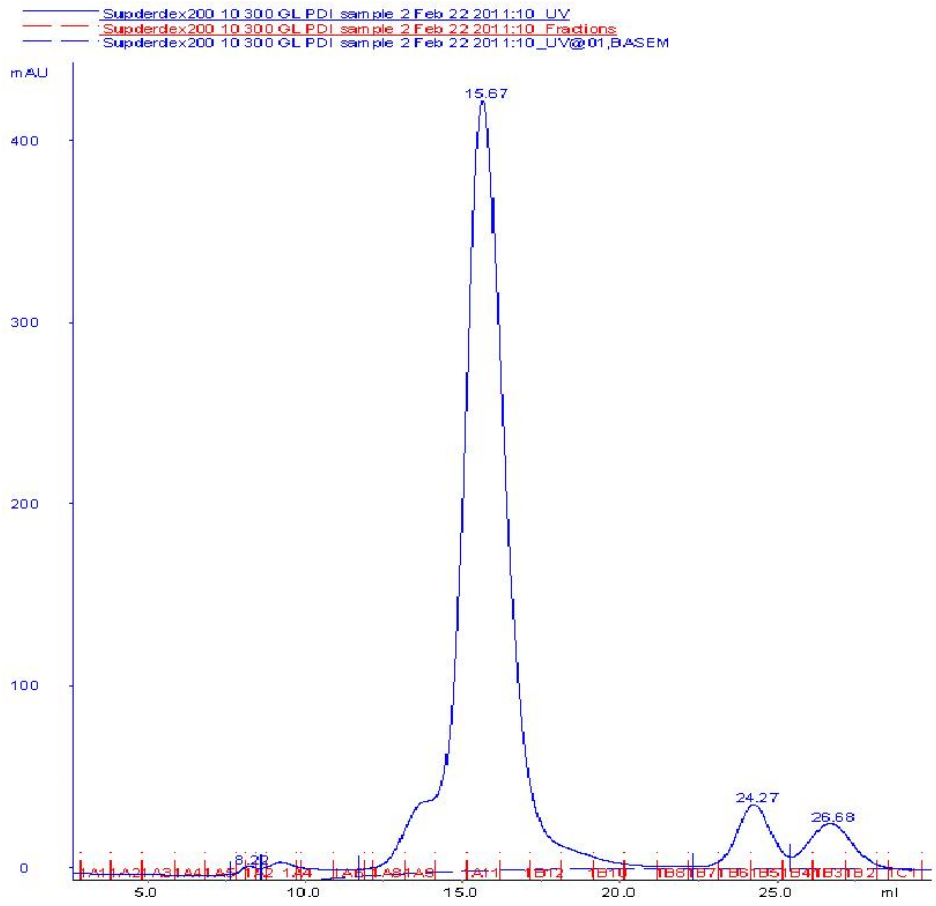


Figure 10: PDI purification by size exclusion chromatography. PDI fractions from the ion exchange column step (Figure 9) were pooled, concentrated in a Centricon 30 kDa filter column and a volume of 2 ml applied to a size exclusion column (Superdex™ 200 GE) in an AKTA FPLC system previously equilibrated with 25 mM Tris-HCl (pH 8.0) with 10 mM beta-mercaptoethanol. 1 ml fractions were collected.

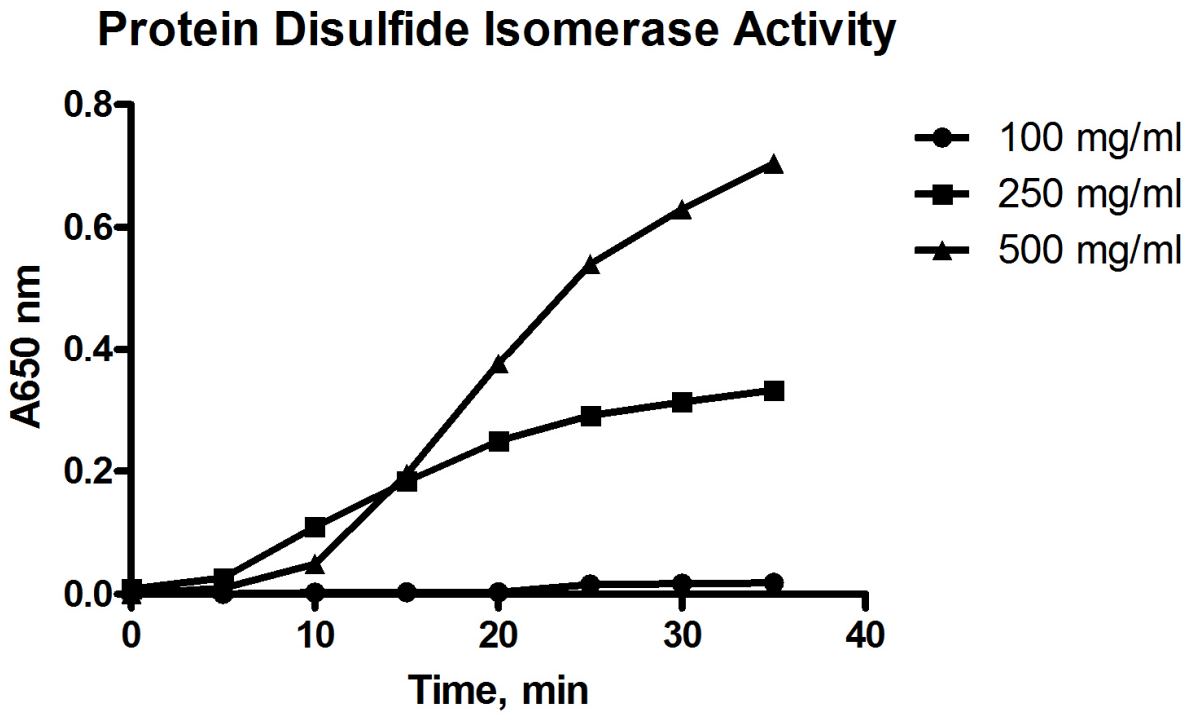


Figure 11: Protein Disulfide Isomerase activity. The assay was performed using a modification to Sigma's protocol, using 0.75 ml of a cocktail made of 10 mM Na-phosphate (pH 7.0), 100 mM DTT, 100 mM Na-EDTA, incubated with 0.25 ml of different concentrations of insulin (100 µg/ml, 250 µg/ml, 500 µg/ml). This solution was incubated for 10 min at 37℃ for 10 min. After incuba tion, 10 µl of protein disulfide isomerase were added into each solution and turbidity at 650 nm of solution was monitored at 5 minute intervals for 30 min.

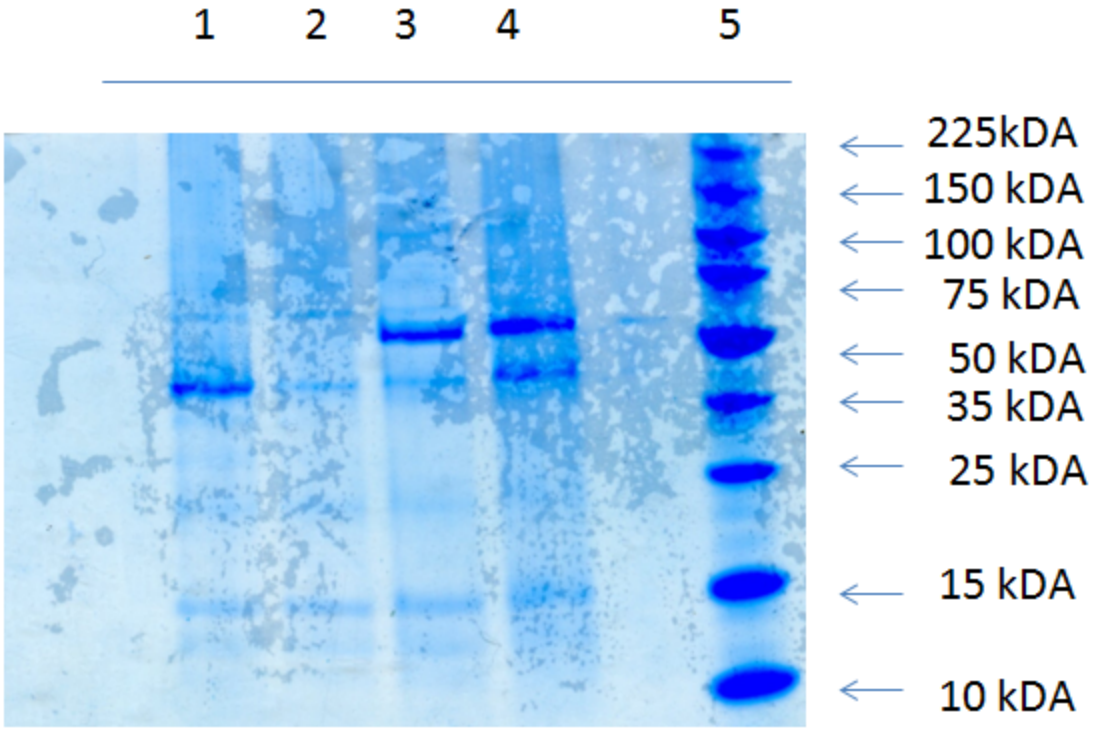


Figure 12: SmCD1 refolding assay with PDI. ProSmCD1 inclusion bodies were solubilized in urea and dialyzed against 25 mM Tris-HCl and 20% glycerol. PDI was added at the end of the dialysis step. Lane 1, without PDI (55 kDa) as a control at pH 8.0; Lane 2: proSmCD1 without PDI as a control at pH.3.5; Lane 3: with PDI at pH 8; Lane 4: Same as lane 1, but at pH 3.5; Lane 5, Fisher E-z Run Protein Marker. No apparent 35 kDa band was present indicating the lack of SmCD1 activation.

3.3- SmCD1 expression in insect cells (SF9 cells)

Previous reports of SmCDs and related parasite's proteases (as Ac-APR-1 and Na-APR-1) shows their ability to be expressed in a soluble and active manner were in insect cells (Brindley et al.2001, Verity *et al,* 2001, Williamson *et al*, 2003). In those cases, SF9 cells were used to expand the vector containing aspartic protease cDNA to be transfected in Trichoplusia ni High Five cells (HiFive TN cells) for protein production. The purification steps used in this case included as a first step a typical affinity chromatography with the most common inhibitor of aspartic proteases, pepstatin.

In this study, proSmCD1 was expressed in SF9 cells as described in Chapter 2 and purified from the soluble cell lysates using a Ni- affinity column. Analysis by Western blot revealed that a polypeptide with a larger than expected molecular weight (75 KDa) was produced at very low levels in the insect cells. Although visible during the immunodecoration procedure the image was too faint to properly capture using the available camera. The fact that the band was larger than expected suggests that the expressed protein was hyperglyscosyled. This result, which is contrary to the one achieved by Brindley *et al,* 2001, may be due to differences between the HiFive TN cells and SF9 cells used in this study. Further investigation is needed to clarify the difference and in order to obtain SmCDs in a soluble and active form from insect cells.

3.4-MBP attachment for soluble SmCD2 expression.

As a final attempt to obtain active SmCD protease a maltose binding protein (MBP) adduct was incorporated into the N-terminal end of the SmCD1 gene. A His6 tag was added to the N-terminal end of the MBP for affinity purification. In this case, the MBP tag was attached to the mature form of the SmCD2 protein rather than the pro form. This construct was prepared by Dr. Silva. Over-expression of the construct resulted in production of considerable amounts of the MBP-SmCD in soluble form as indicated in Figure 13 following purification by Ni-affinity chromatography. Examination of the gel profile indicated the presence of several lower molecular bands in addition to the 75 kDa MBP-SmCD protein. Since this system routinely results in highly purified protein the data suggest that the lower molecular weight bands may result from self cleavage by the MBP-SmCD protein. This was the first piece of evidence that could be obtained for the presence of an active recombinant SmCD protein.

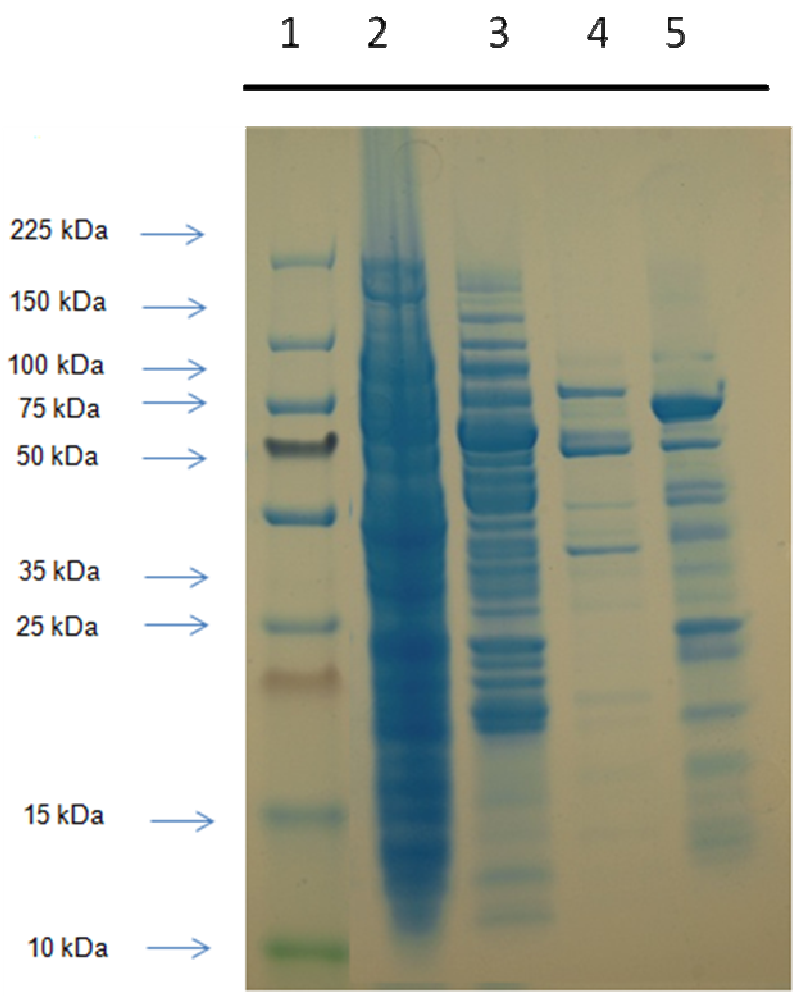


Figure 13: Purification of MBP-SmCD2. *E.coli* BL21 cells transformed with the MBP-SmCD2 DNA were grown to mid log phase and induced with IPTG as described in Chapter 2. Cells were harvested and lysed by sonication. The soluble cell fraction following lysis was applied to a Ni-affinity column. Lane1, Promega Protein Marker; Lane 2, cell lysate; Lane 3, soluble cell fraction; Lane 4, column flow through material; Lane 5: protein eluted with 300 mM imidazole.

3.5- Discussion

The expression of SmCD protein In *E.coli* resulted in formation of inclusion bodies, insoluble protein aggregates. This fact is not a surprise since the appearance of insoluble aggregates is common when *E.coli* is used as a host for helminth or mammalian proteins (Sahdev *et al*., 2008). Although inclusion body formation can be useful in terms of ease of purification, the process of solubilization and refolding can be laborious and may lead to a low percentage of active, properly folded protein (Sahdev *et al*., 2008).

Virtually all polypeptide chains can form insoluble aggregates as well as soluble forms depending on the physical factors within the environment that govern protein folding, such as temperature, pH, chemical composition, pressure and viscosity. Each cell has specific ways to modulate these factors, using organic and inorganic compounds, chaperone systems and compartmentalization. Inclusion bodie refolding processes attempt to mimic those conditions *in vitro* (Jahn & Radford, 2008; Bolen & Rose, 2008; Frauenfelder *et al*., 2006).

Extensive refolding trials aiming at recovering a self activating functional protease from ProSmCD inclusion bodies were performed under numerous different conditions. Although soluble forms of this protein were obtained, the protein was either not correctly folded for expression of its proteolysis function or this function requires an external factor or partner to be active.

A Schistosoma genome project done in early 1990's decade suggests the presence of at least one protein disulfide isomerase (PDI)in *S. mansonii*. This is important not only because it can be the partner required for SmCDs refolding (these proteins contains disulfide bridges that can be rearranged by PDI, but also because PDI can be considered a possible drug target itself, since it is involved in folding of toxins and other related infective proteins of pathogenic bacteria (enteropathogenic *E. coli* and *Vibrio cholerae*). However, PDI crystals for structural studies were not obtained in the conditions and it is possible that this enzyme have the same high degree of flexibility in the a domain as others of PDI family which turn formation of crystals a difficult task. Also, although *S. mansoni* PDI obtained is active in rearrange insulin disulfide briges, its presence did not lead to SmCDs activation which may indicate that this PDI is not used as a chaperone or isomerase to assist correct folding of SmCDs.

The external factor or partner hypothesis is also supported by the observation that the mature form of SmCD2 expressed with the MBP tag, a partner that helps increase solubility, may be active (further investigation is currently underway). However, there are other alternative explanations for the lack of function in our experiments with the pro form of SmCD1. In recent years, several authors published papers about other functions than proteolysis of human cathepsin D. Beaujouin *et al*, 2006, for example, indicates the possible role of non proteolitic (inactive for proteolysis) cathepsin D in apoptosis, which could possibly be related with similar phenomena in exchange of life stages of Schistosoma sp., Vashishta *et al***,** 2005 discuss the possible dual role of cathepsin D, as a protease and as a ligand factor crucial for cancer metabolic pathways.

Therefore, since SmCDs are cathepsin D like proteins, all the soluble forms achieved in this work may be microstates of the protein related to other functions than a proteolytic function.

On the other hand, invertebrates like the helminth Schistosoma seem to have a cascade for degradation of proteins for nutritional purposes. Although this cascade seems to occur in a non-coordinated and redundant sequence of events, in contrast to vertebrates (Delcroix *et al*., 2006), there may be some sequence of events required for activation of the SmCD proteases. An external factor required for activation of protein degradation may be the most reasonable scenario for SmCDs and their proteolysis activity. Rather than constitute a problem, the combination of this two possible explanations may increase the importance and appeal of SmCDs as possible drug targets against schistosomiasis.

In conclusion, we have obtained evidence that mature form of SmCD2 attached to the MBP may be proteolytically active. This provides a promising avenue of research to developing an assay suitable for identifying potential lead compounds, through screening of molecular libraries, as a basis for developing new drugs to target Schistosomiasis.

Chapter 4- References

1. Abdulla, M.; Lim, K.; Sajid, M.; McKerrow, J.; Caffrey, C., Schistosomiasis mansoni: Novel chemotherapy using a cysteine protease inhibitor. *Plos Medicine* **2007**, 130-138.

2. Arakawa, T.; Tsumoto, K.; Kita, Y.; Chang, B.; Ejima, D., Biotechnology applications of amino acids in protein purification and formulations. *Amino Acids* **2007,** *33* (4), 587-605.

3. Arredondo, S.; Chen, T.; Riggs, A.; Gilbert, H.; Georgiou, G., Role of Dimerization in the Catalytic Properties of the Escherichia coli Disulfide Isomerase DsbC. *Journal of Biological Chemistry* **2009**, 23972-23979.

4. Baneyx, F.; Mujacic, M., Recombinant protein folding and misfolding in Escherichia coli. *NATURE BIOTECHNOLOGY* **2004,** *22* (11), 1399-1408.

5. Beaujouin, M.; Baghdiguian, S.; Glondu-Lassis, M.; Berchem, G.; Liaudet-Coopman, E., Overexpression of both catalytically active and -inactive cathepsin D by cancer cells enhances apoptosis-dependent chemo-sensitivity. *Oncogene* **2006,** *25* (13), 1967-1973.

6. BECKER, M.; HARROP, S.; DALTON, J.; KALINNA, B.; MCMANUS, D.; BRINDLEY, P., CLONING AND CHARACTERIZATION OF THE SCHISTOSOMA-JAPONICUM ASPARTIC PROTEINASE INVOLVED IN HEMOGLOBIN DEGRADATION. *Journal of Biological Chemistry* **1995**, 24496-24501.

7. Bolen, D.; Rose, G., Structure and energetics of the hydrogen-bonded backbone in protein folding. *Annual Review of Biochemistry* **2008,** *77*, 339-362.

8. Brindley, P.; Kalinna, B.; Dalton, J.; Day, S.; Wong, J.; Smythe, M.; McManus, D., Proteolytic degradation of host hemoglobin by schistosomes. *Molecular and Biochemical Parasitology* **1997**, 1-9.

9. Capron, A.; Riveau, G.; Capron, M.; Trottein, F., Schistosomes: the road from host-parasite interactions to vaccines in clinical trials. *Trends in Parasitology* **2005**, 143-149.

10. Carginale, V.; Trinchella, F.; Capasso, C.; Scudiero, R.; Riggio, M.; Parisi, E., Adaptive evolution and functional divergence of pepsin gene family. *GENE* **2004,** *333*, 81-90.

11. Curwen, R.; Ashton, P.; Johnston, D.; Wilson, R., The Schistosoma mansoni soluble proteome: a comparison across four life-cycle stages. *Molecular and Biochemical Parasitology* **2004**, 57-66.

12. Dee, D.; Filonowicz, S.; Horimoto, Y.; Yada, R., Recombinant prosegment peptide acts as a folding catalyst and inhibitor of native pepsin. *Biochimica Et Biophysica Acta-Proteins and Proteomics* **2009**, 1795-1801.

13. Dix, J.; Verkman, A., Crowding effects on diffusion in solutions and cells. *ANNUAL REVIEW OF BIOPHYSICS* **2008,** *37*, 247-263.

14. Dunn, B., Structure and mechanism of the pepsin-like family of aspartic peptidases. *CHEMICAL REVIEWS* **2002,** *102* (12), 4431-4458.

15. Eder, J.; Hommel, U.; Cumin, F.; Martoglio, B.; Gerhartz, B., Aspartic proteases in drug discovery. *Current Pharmaceutical Design* **2007**, 271-285.

16. Fenwick, A.; Savioli, L.; Engels, D.; Bergquist, N.; Todd, M., Drugs for the control of parasitic diseases: current status and development in schistosomiasis. *Trends in Parasitology* **2003**, 509-515.

17. FinkenEigen, M.; Kunz, W., Schistosoma mansoni: Gene structure and localization of a homologue to cysteine protease ER 60. *Experimental Parasitology* **1997,** *86* (1), 1-7.

18. Fortenberry, S. C.; Chirgwin, J. M., The propeptide is nonessential for the expression of human cathepsin D. *J Biol Chem* **1995,** *270* (17), 9778-82.

19. Fowler, D.; Koulov, A.; Balch, W.; Kelly, J., Functional amyloid - from bacteria to humans. *Trends in Biochemical Sciences* **2007,** *32* (5), 217-224.

20. Frauenfelder, H.; Fenimore, P.; Chen, G.; McMahon, B., Protein folding is slaved to solvent motions. *Proceedings of the National Academy of Sciences of the United States of America* **2006**, 15469-15472.

21. Fusek, M.; Mares, M.; Vágner, J.; Voburka, Z.; Baudys, M., Inhibition of aspartic proteinases by propart peptides of human procathepsin D and chicken pepsinogen. *FEBS Lett* **1991,** *287* (1-2), 160-2.

22. Hatahet, F.; Ruddock, L., Substrate recognition by the protein disulfide isomerases. *Febs Journal* **2007**, 5223-5234.

23. Hernandez, G.; Anderson, J.; LeMaster, D., Electrostatic stabilization and general base catalysis in the active site of the human protein disulfide isomerase a domain monitored by hydrogen exchange. *Chembiochem* **2008**, 768-778.

24. Horimoto, Y.; Dee, D.; Yada, R., Multifunctional aspartic peptidase prosegments. *NEW BIOTECHNOLOGY* **2009,** *25* (5), 318-324.

25. Jahn, T.; Radford, S., Folding versus aggregation: Polypeptide conformations on competing pathways. *Archives of Biochemistry and Biophysics* **2008,** *469* (1), 100-117.

26. Jean, L.; Long, M.; Young, J.; Pery, P.; Tomley, F., Aspartyl proteinase genes from apicomplexan parasites: evidence for evolution of the gene structure. *Trends in Parasitology* **2001**, 491-498.

27. Jolodar, A.; Miller, D., Identification of a novel family of non-lysosomal aspartic proteases in nematodes. *Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology* **1998**, 13-16.

28. Kaldor, S.; Kalish, V.; Davies, J.; Shetty, B.; Fritz, J.; Appelt, K.; Burgess, J.; Campanale, K.; Chirgadze, N.; Clawson, D.; Dressman, B.; Hatch, S.; Khalil, D.; Kosa, M.; Lubbehusen, P.; Muesing, M.; Patick, A.; Reich, S.; Su, K.; Tatlock, J., Viracept (nelfinavir mesylate, AG1343): A potent, orally bioavailable inhibitor of HIV-1 protease. *Journal of Medicinal Chemistry* **1997**, 3979-3985.

29. Khan, A.; Khazanovich-Bernstein, N.; Bergmann, E.; James, M., Structural aspects of activation pathways of aspartic protease zymogens and viral 3C protease precursors. *Proceedings of the National Academy of Sciences of the United States of America* **1999**, 10968-10975.

30. Koelsch, G.; Mares, M.; Metcalf, P.; Fusek, M., Multiple functions of pro-parts of aspartic proteinase zymogens. *FEBS Letters* **1994,** *343* (1), 6-10.

31. Kozlov, G.; Maattanen, P.; Schrag, J.; Hura, G.; Gabrielli, L.; Cygler, M.; Thomas, D.; Gehring, K., Structure of the Noncatalytic Domains and Global Fold of the Protein Disulfide Isomerase ERp72. *Structure* **2009**, 651-659.

32. Lazure, C., The peptidase zymogen proregions: Nature's way of preventing undesired activation and proteolysis. *CURRENT PHARMACEUTICAL DESIGN* **2002,** *8* (7), 511-531.

33. Lazure, C., The peptidase zymogen proregions: Nature's way of preventing undesired activation and proteolysis. *Current Pharmaceutical Design* **2002**, 511-531.

34. Maeda, R.; Ado, K.; Takeda, N.; Taniguchi, Y., Promotion of insulin aggregation by protein disulfide isomerase. *Biochimica Et Biophysica Acta-Proteins and Proteomics* **2007**, 1619-1627.

35. Mahajan, B.; Noiva, R.; Yadava, A.; Zheng, H.; Majam, V.; Mohan, K.; Moch, J.; Haynes, J.; Nakhasi, H.; Kumar, S., Protein disulfide isomerase assisted protein folding in malaria parasites. *International Journal For Parasitology* **2006**, 1037-1048.

36. McManus, D.; Hu, W.; Brindley, P.; Feng, Z.; Han, Z., Schistosome transcriptome analysis at the cutting edge. *Trends in Parasitology* **2004**, 301-304.

37. Merrick, J.; Osman, A.; Tsai, J.; Quackenbush, J.; LoVerde, P.; Lee, N., The Schistosoma mansoni gene index: Gene discovery and biology by reconstruction and analysis of expressed gene sequences. *Journal of Parasitology* **2003**, 261-269.

38. Minor, D., The neurobiologist's guide to structural biology: A primer on why macromolecular structure matters and how to evaluate structural data. *Neuron* **2007**, 511-533.

39. Morales, M.; Rinaldi, G.; Gobert, G.; Kines, K.; Tort, J.; Brindley, P., RNA interference of Schistosoma mansoni cathepsin D, the apical enzyme of the hemoglobin proteolysis cascade. *Molecular and Biochemical Parasitology* **2008**, 160-168.

40. Parks, A.; Curtis, D., Presenilin diversifies its portfolio. *Trends in Genetics* **2007**, 140-150.

41. Qoronfleh, M.; Hesterberg, L.; Seefeldt, M., Confronting high-throughput protein refolding using high pressure and solution screens. *Protein Expression and Purification* **2007,** *55* (2), 209-224.

42. Rabiller, M.; Getlik, M.; Kluter, S.; Richters, A.; Tuckmantel, S.; Simard, J.; Rauh, D., Proteus in the World of Proteins: Conformational Changes in Protein Kinases. *Archiv Der Pharmazie* **2010,** *343* (4), 193-206.

43. Raturi, A.; Mutus, B., Characterization of redox state and reductase activity of protein disulfide isomerase under different redox environments using a sensitive fluorescent assay. *Free Radical Biology and Medicine* **2007**, 62-70.

44. RAWLINGS, N.; BARRETT, A., EVOLUTIONARY FAMILIES OF PEPTIDASES. *Biochemical Journal* **1993**, 205-218.

45. Renslo, A.; McKerrow, J., Drug discovery and development for neglected parasitic diseases. *Nature Chemical Biology* **2006**, 701-710.

46. Rhee, Y.; Pande, V., Solvent viscosity dependence of the protein folding dynamics. *Journal of Physical Chemistry B* **2008,** *112* (19), 6221-6227.

47. Sahdev, S.; Khattar, S.; Saini, K., Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. *Molecular and Cellular Biochemistry* **2008,** *307* (1-2), 249-264.

48. Shea, M.; Jakle, U.; Liu, Q.; Berry, C.; Joiner, K.; Soldati-Favre, D., A family of aspartic proteases and a novel, dynamic and cell-cycle-dependent protease localization in the secretory pathway of Toxoplasma gondii. *Traffic* **2007**, 1018-1034.

49. Shin, J.; Yu, M., Viscous drag as the source of active site perturbation during protease translocation: Insights into how inhibitory processes are controlled by serpin metastability. *Journal of Molecular Biology* **2006**, 378-389.

50. Shinde, U.; Inouye, M., Intramolecular chaperones: polypeptide extensions that modulate protein folding. *SEMINARS IN CELL & DEVELOPMENTAL BIOLOGY* **2000,** *11* (1), 35-44.

51. Sibley, C.; Hunt, S., Drug resistance in parasites: can we stay ahead of the evolutionary curve? *Trends in Parasitology* **2003**, 532-537.

52. Silva, F.; Ribeiro, F.; Katz, N.; Giovanni-De-Simone, S., Exploring the subsite specificity of Schistosoma mansoni aspartyl hemoglobinase through comparative molecular modelling. *Febs Letters* **2002**, 141-148.

53. Skelly, P.; Da'dara, A.; Harn, D., Suppression of cathepsin B expression in Schistosoma mansoni by RNA interference. *International Journal For Parasitology* **2003**, 363-369.

54. Tcherepanova, I.; Bhattacharyya, L.; Rubin, C.; Freedman, J., Aspartic proteases from the nematode Caenorhabditis elegans - Structural organization and developmental and cell-specific expression of asp-1. *Journal of Biological Chemistry* **2000**, 26359-26369.

55. Turk, B., Targeting proteases: successes, failures and future prospects. *Nature Reviews Drug Discovery* **2006**, 785-799.

56. Tyndall, J.; Nall, T.; Fairlie, D., Proteases universally recognize beta strands in their active sites. *CHEMICAL REVIEWS* **2005,** *105* (3), 973-999.

57. Valdivieso, E.; Bermudez, H.; Hoebeke, J.; Noya, O.; Cesari, I., Immunological similarity between Schistosoma and bovine cathepsin D. *Immunology Letters* **2003**, 81-88.

58. Vashishta, A.; Fusek, M.; Vetvicka, V., Possible role of procathepsin D in human cancer. *Folia Microbiologica* **2005,** *50* (1), 71-76.

59. Vedadi, M.; Niesen, F.; Allali-Hassani, A.; Fedorov, O.; Finerty, P.; Wasney, G.; Yeung, R.; Arrowsmith, C.; Ball, L.; Berglund, H.; Hui, R.; Marsden, B.; Nordlund, P.; Sundstrom, M.; Weigelt, J.; Edwards, A., Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proceedings of the National Academy of Sciences of the United States of America* **2006,** *103* (43), 15835-15840.

60. Verity, C.; McManus, D.; Brindley, P., Developmental expression of cathepsin D aspartic protease in Schistosoma japonicum. *International Journal For Parasitology* **1999**, 1819-1824.

61. Verity, C.; McManus, D.; Brindley, P., Cellular responses to Schistosoma japonicum cathepsin D aspartic protease. *Parasite Immunology* **2002**, 363-367.

62. Verjovski-Almeida, S.; Leite, L.; Dias-Neto, E.; Menck, C.; Wilson, R., Schistosome transcriptome: insights and perspectives for functional genomics. *Trends in Parasitology* **2004**, 304-308.

63. Wilson, R.; Coulson, P., Schistosome vaccines: a critical appraisal. *Memorias Do Instituto Oswaldo Cruz* **2006**, 13-20.

64. Wong, J.; Harrop, S.; Day, S.; Brindley, P., Schistosomes express two forms of cathepsin D. *Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology* **1997**, 156-160.

65. Zhou, H.; Rivas, G.; Minton, A., Macromolecular crowding and confinement: Biochemical, biophysical, and potential physiological consequences. *Annual Review of Biophysics* **2008**, 375-397.